

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

Ebben a függelékben a nyelvelsajátítás szintjeinek leírásával kapcsolatos technikai kérdésekkel foglalkozunk. Ismertetjük a szintleírások kidolgozásának kritériumait. Bemutatjuk a skálák kialakításának módszereit, majd végül annotált bibliográfiát adunk meg.

253

Szintleírások kialakítása

A nyelvi tesztelésben a skálák létrehozásának gyakorlata, az alkalmazott pszichológia átfogóbb területén a színtezés elmélete, valamint a konzultációkban¹ részt vevő tanárok véleménye alapján a szintleírások kialakításának alábbi irányelvei javasolhatók:

- **Pozitív irányultság:** Az értékelő-szemponitú készségszintskálák és a vizsgákon használt pontozó skálák alacsonyabb szinteken gyakran negatív értelmű megfogalmazást használnak. Alacsonyabb szinteken nehezebb a készségszint megfogalmazása abból a szempontból, hogy a tanuló mit tud megtenni, mint abból a szempontból, hogy mit nem tud. Azonban ha a készségszintek nem csupán a vizsgáztatás eszközeként szolgálnak, hanem célkitűzést is jelentenek, akkor kívánatos a pozitív megfogalmazás. Néha ugyanazt a szempontot pozitív és negatív formában is meg lehet fogalmazni, például a nyelvhasználat választékossága esetében:

¹ pl. az Egyesült Királyságban létrehozott színtezett célrendszerek és a svájci kutatási program esetében.

„A” Függelék: A készség szintek leírására szolgáló szintleírások kialakítása

A1. táblázat: *Értékelés: pozitív és negatív kritériumok*

Pozitív	Negatív
<ul style="list-style-type: none"> a nyelvismeret és stratégiák alapvető készségével rendelkezik, ami lehetővé teszi számára, hogy kiszámítható, mindennapi helyzetekben elboldoguljon. (Eurocentres 3. szint: bizonyítvány) a nyelvismeret és stratégiák alapvető készségével rendelkezik a mindennapi szükségletek nagy részének kifejezéséhez, de általában le kell egyszerűsíteni mondanivalóját, és keresgélnie kell a szavakat. (Eurocentres 3. szint: pontozási táblázat) 	<ul style="list-style-type: none"> szűk nyelvi készséggel rendelkezik, állandóan újra kell fogalmaznia mondanivalóját, és keresgéli a szavakat. (ESU 3. szint) nem begyakorolt helyzetekben korlátozott nyelvtudása gyakran a kommunikáció megszakadásához és félreértéshez vezet. (Finn 2. szint) a kommunikáció megszakad, amikor a nyelvi korlátok kihatnak a mondanivaló megformálására. (ESU 3. szint)
<ul style="list-style-type: none"> a szókincs az alapvető tárgyakat, helyeket és a leggyakoribb rokonságot jelölő szavakra terjed ki. (ACTFL Novice) 	<ul style="list-style-type: none"> csak korlátozott szókincssel rendelkezik (Holland 1. Szint) a szavak és kifejezések korlátozott köre gátolja a gondolatok és elképzelések közvetítését. (Gothenburg U)
<ul style="list-style-type: none"> Használja és felismeri a szavak és rövid fordulatok betanult készletét. (Trim, 1978: 1. szint) 	<ul style="list-style-type: none"> csak formulákat, felsorolásokat és egyszerű megnevezéseket tud használni. (ACTFL Novice)
<ul style="list-style-type: none"> Használni tud rövid mindennapos kifejezéseket konkrét típusú egyszerű szükségletek kielégítésének érdekében (üdvözlés, információ stb. területén). (Elviri; Milánó 1. Szint, 1986) 	<ul style="list-style-type: none"> csak a legalapvetőbb nyelvi készséggel rendelkezik, a nyelv funkcionális használatának kevés vagy semmi jele sincs. (ESU 1. szint)

A negatív irányú megfogalmazás elkerülésekor további bonyodalmat okoz, hogy a kommunikatív nyelvi készség szintnek vannak olyan jellemzői, amelyek nem összeadó jellegűek. Minél kevesebb, annál jobb. Ennek a legnyilvánvalóbb példája, amit olykor önállóságnak neveznek, vagyis hogy milyen mértékben függ a tanuló attól, hogy *a*) a beszélgetőpartner módosítsa beszédét, *b*) van-e lehetőség az elmondottak tisztázására, és *c*) kap-e segítséget mondanivalója megfogalmazásában. Az ilyen kérdéseket gyakran a pozitív módon megfogalmazott szintleírásokhoz mellékelt kikötésekben lehet megfogalmazni, például:

Általában meg tudja érteni a hozzá intézett, ismerős témákkal kapcsolatos világos, köznyelven megfogalmazott beszédet, amennyiben időnként ismétlést vagy átfogalmazást kérhet. Egyszerű, mindennapi társalgás során meg tudja érteni, amit világosan, lassan és

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

közvetlenül mondanak neki; meg lehet értetni vele, hogy miről van szó, ha a beszélő veszi a fáradságot.

vagy:

Strukturált helyzetekben és rövid társalgásokban viszonylag könnyedén tud kommunikálni, ha a másik személy szükség esetén segít.

- **Meghatározottság:** A szintleírásoknak konkrét feladatokat és/vagy a feladat végrehajtásához alkalmazott készségek konkrét fokozatait kell leírniuk. Ehhez két dolgot kell kiemelniük. Először is, kerülni kell a szintleírások megfogalmazásakor a pontatlanságot, mint például: „Használni tudja a stratégiák megfelelő körét.” Mit értünk stratégián? Mihez megfelelő? Hogyan értelmezzük a „kör” szót? Az a baj a pontatlan szintleírásokkal, hogy lehet, hogy jól hangzanak, de az elfogadhatóság látszólagos egyszerűsége elfedheti azt a tényt, hogy mindenki másképpen értelmezi őket. Másodsor, az 1940-es évek óta elfogadott elv, hogy egy skála fokozatai közötti különbségek ne egy módosító szó másakra való kicserélésén alapuljanak, mint például az „egyes” vagy „néhány” helyettesítése a „sok” és „legtöbb” szavakkal, vagy a „meglehetősen széles” helyettesítése a „nagyon széles” kifejezéssel, illetve a „mérsékelt” helyettesítése a „jó” szóval az eggyel feljebb következő szinten. A megkülönböztetés legyen valós, ne csak fogalmazásbeli, ez viszont hézagokat eredményezhet ott, ahol nem lehet jelentéssel bíró, konkrét megkülönböztetéseket tenni.
- **Világosság:** A szintleírások legyenek átláthatóak és a szakzsargonról mentesek. A megértési korlátoktól eltekintve, néha a szaknyelv elhagyása esetén egy látszólag lenyűgöző szintleírás szinte semmitmondóvá válhat. Másrészt megfogalmazásuk legyen egyszerű mondatszerkezetű, világos és logikus felépítésű.
- **Rövidség:** Van olyan elméleti felfogás, amelyet a holisztikus skálákkal azonosítunk, ezt a felfogást különösen Amerikában és Ausztráliában követik. Ennek a felfogásnak megfelelően egy hosszabb bekezdésben átfogó jelleggel ismertetik a főbb jellemzőket. Az ilyen skálák a „meghatározottság” kritériumát nagyon átfogó felsorolással érik el, amelynek célja egy részletes portré megalkotása arról, ami az értékelők számára megjeleníti az adott szinten lévő tipikus tanulót, és így ezek a skálák a leírás nagyon gazdag forrását nyújtják. Azonban az ilyen megközelítés két hátránnyal jár. Az egyik az, hogy valójában egy tanuló sem „tipikus”. A részletes jellemzők eltérő módokon keverednek egymással. Másrészt a két tagmondatnál hosszabb szintleírások reálisan nem kezelhetők az értékelési folyamat során. A tanárok következetesen a rövid szintleírások-

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

kat kedvelik. A szemléltető szintleírásokat létrehozó kutatási program során a tanárok elvetették vagy felbontották a körülbelül 25 szónál hosszabb szintleírásokat (normál betűmérettel körülbelül két sornyi).

- **Önállóság:** A rövid szintleírásoknak két másik előnye is van. Először is, valószínűbb, hogy olyan viselkedést írnak le, amiről elmondható: „Igen, ez az ember ezt meg tudja csinálni.” Tehát a rövidebb, konkrét szintleírások használhatók önálló kritérium-állításokként listákban vagy kérdőívekben a folyamatos tanári értékelés és/vagy önértékelés során. Ez a fajta önálló integritás jelzi, hogy egy adott szintleírás szerepelhet célkitűzésként, és nem csak a skála többi szintleírásának viszonylatában bír jelentéssel. Ezt sokféleképpen lehet kiaknázni a különböző fajta értékelések terén (lásd 9. fejezet).

A Referenciakeret használói gondolkozzanak el, és ahol lehet, fogalmazzák meg, hogy

- *a felsorolt kritériumok közül melyek a legfontosabbak, és saját kontextusukban milyen más kritériumokat használnak explicit vagy implicit módon;*
- *milyen mértékben kívánatos és kivitelezhető rendszerükben a fent említett kritériumoknak megfelelő szintleírások kialakítása.*

A skálák kialakításának módszerei

Az a tény, hogy a szintek egymásra épülő sort alkotnak, azt a feltételezést sugallja, hogy bizonyos dolgok inkább az egyik, mint a másik szinten helyezhetők el, és hogy egy bizonyos készség adott fokának a leírása inkább az egyik szintre vonatkozik, mint egy másikra. Ez egy fajta következetesen alkalmazott szintezési eljárást jelent. Számos különböző módja van a nyelvtudás készségszintekben kifejezett leírásának. A létező módszerek három csoportba sorolhatók: intuitív módszerek, kvalitatív módszerek és kvantitatív módszerek. A nyelvi készségszintek legtöbb létező skáláját és más szintrendszereket az első, intuitív csoportba tartozó három módszer egyike alapján fejlesztették ki. A legjobb megközelítések viszont mindhárom módszert kombinálják az egymást kiegészítő halmazati folyamat során. A kvalitatív módszerek alkalmazásához szükség van az anyag intuitív előkészítésére és kiválasztására, valamint az eredmények intuitív értelmezésére. A kvantitatív módszereknek kvalitatív módon előtesztelt anyagot kell számszerűsíteniük, az eredményeket pedig intuitív módon kell értelmezni. Ezért a Közös

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

Referenciaszintek kialakításakor kombináltuk az intuitív, kvalitatív és kvantitatív megközelítéseket.

Kvalitatív és kvantitatív módszerek alkalmazásakor két lehetséges kiindulópont van: szintleírások vagy teljesítményminták.

Szintleírásokból való kiindulás: Az egyik kiindulópont annak meghatározása, hogy mit kívánunk leírni, majd ezután történik az adott kategóriákhoz tartozó szintleírások megfogalmazása, összegyűjtése vagy szerkesztése. Ezek a szintleírások a kvalitatív fázis bemeneteként szolgálnak. A 4. és 9. módszer - a lent megadott felsorolásban a kvalitatív csoport első és utolsó tagja - erre a megközelítésre ad példát. Ez a megközelítés különösen alkalmas olyan tantervhez kötődő kategóriák szintleírásainak kialakítására, amelyek például a kommunikatív nyelvi tevékenységekre vonatkoznak, de a kompetencia különböző aspektusait jellemző szintleírások kialakítására is alkalmas. A kategóriákból és szintleírásokból való kiindulás előnye, hogy így elméletileg megalapozott módon lehetséges egy kiegyensúlyozott vonatkozási terület meghatározása.

Teljesítménymintákból való kiindulás: A másik alternatíva - ami csak a teljesítmény értékeléséhez használható szintleírások kialakítására alkalmas - a teljesítmény reprezentatív mintáiból való kiindulás. Ebben az esetben meg lehet kérdezni a reprezentatív pontozókat, hogy mire figyelnek, amikor a mintákkal dolgoznak (kvalitatív). Az 5.- 8. módszer ennek az elgondolásnak különböző változatait képviseli. Egy másik eljárás szerint csak arra kéri a pontozókat, hogy értékeljék a mintákat, majd a megfelelő statisztikai eljárás alkalmazásával meg lehet állapítani, hogy ténylegesen milyen főbb jellemzők alapján döntenek a pontozók (kvantitatív). A 10. és a 11. módszer erre a megközelítésre ad példát. A teljesítményminták elemzésének előnye, hogy az adatok alapján nagyon konkrét leírásokat lehet létrehozni.

Az utolsó, 12. módszer az egyetlen, amelyik matematikai értelemben véve ténylegesen *szintezi* a szintleírásokat. Ezt a módszert használtuk a Közös Referenciaszintek és a szemléltető szintleírások kialakításakor, a 2. módszert (intuitív), valamint a 8. és 9. módszert (kvalitatív) követően. Azonban a skála kialakítása után ugyanez a statisztikai eljárás felhasználható a skála gyakorlati alkalmazásának validálására (a skála érvényességének a megállapítására), valamint az átdolgozás szükségességének a megállapítására.

„A” Függelék: A készség szintek leírására szolgáló szintleírások kialakítása

Intuitív módszerek:

Ezek a módszerek semmiféle strukturált adatgyűjtést nem igényelnek, csupán a gyakorlat elvi alapokon történő értelmezését.

1. **Szakértői:** Valakit felkérnek egy skála megírására, ami megtörténhet létező skálák, tantervi dokumentumok és más megfelelő forrásanyagok felhasználásával, valószínűleg az adott célcsoport szükségleteinek elemzését követően. Ezután kipróbálják és átdolgozzák a skálát, valószínűleg adatközlők segítségével.
2. **Bizottsági:** Hasonló az előzőhöz, de egy nagyobb tanácsadói csoport mellett egy kis fejlesztési csoportot hoznak létre. A tanácsadók véleményezik a különböző verziókat. A tanácsadók intuitív módon tapasztalataikra alapozva dolgozhatnak, és/vagy a tanulókkal vagy teljesítménymintákkal való összehasonlítás alapján is. Gipps (1994) és Scarino (1996; 1997) tárgyalja az Egyesült Királyságban és Ausztráliában bizottság által az élő idegen nyelvek középiskolai tanulásához kialakított tantervi skálarendszer gyenge pontjait.
3. **Tapasztalati:** Mint a bizottság esetében, de a folyamat meglehetősen hosszú ideig tart, amíg kialakul egy intézményen és/vagy egy adott értékelési kontextuson belül egy „belső megegyezés”. A döntéshozóknak egységes értelmezésre kell jutniuk a szinteket és a kritériumokat illetően. A megfogalmazás finomítása érdekében szisztematikus kipróbálás és visszajelzés következhet. Az értékelők csoportosan megbeszélhetik a teljesítményeket a definíciók alapján, és a definíciókat a teljesítményminták alapján. Ez a készség szintrendszerek kialakításának hagyományos módja (Wilds 1975; Ingram 1985; Liskin-Gasparro 1984; Lowe 1985, 1986).

Kvalitatív módszerek:

Ezek a módszerek az adatközlők kis csoportjaival dolgozó műhelymunkára építenek, és a szerzett információ értelmezése inkább kvalitatív, mint statisztikai módszerekkel történik.

4. **Kulcsfogalmak: megfogalmazás:** Amikor már létrejött egy skálatervezet, a skála feldarabolásának egyszerű eljárását alkalmazzák, és a skála tipikus felhasználói csoportjának reprezentatív adatközlőit megkérlik arra, hogy a) tegyék a definíciókat a szerintük helyes sorrendbe, b) magyarázzák meg, hogy miért ezt tartják a helyes sorrendnek, és -

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

miután az ő sorrendjük és a tervezett sorrend közötti különbségek megállapítást nyertek - c) az adatközlők meghatározzák azokat a kulcsfontosságú pontokat, amelyek a sorba rendezésnél segítségükre voltak, vagy összezavarták őket. A finomítás néha egy szint elmozdítását jelenti, ami azt a másodlagos feladatot eredményezi, hogy meg kell állapítani azt a két szint közötti hézagot, ahonnan hiányzik egy szint.

5. **Kulcsfogalmak: teljesítmények:** A leírás és a tényleges előfordulás közti következetesség biztosítása érdekében a szintleírásokat az adott szint sávjában tipikus teljesítményekkel vetik össze. Egyes cambridge-i vizsgaleírások végigvizik a tanárokat ezen a folyamaton oly módon, hogy a skálák megfogalmazását egyes dolgozatokra adott jegyekkel társítják. Az IELTS (International English Language Testing System) szintleírásainak létrehozásakor tapasztalt pontozók néhány csoportját megkérték arra, hogy minden szinten azonosítsanak „mintadolgozatokat”, majd egyezzenek meg minden dolgozat „főbb tulajdonságait” illetően. Azután megbeszélés során azonosítják azokat a tulajdonságokat, amelyeket a különböző szinteken jellemzőnek ítélnék, majd ezeket belefoglalják a szintleírásokba (Alderson 1991; Shohamy et al. 1992).
6. **Legfőbb jellemző (primary trait):** Az (általában írásbeli) teljesítményeket az adatközlők egyéni munka során rangsorolják. Ezt követően megegyeznek egy közös rangsorban. Ezután minden szintre meghatározzák és leírják azokat az elveket, amelyek alapján a dolgozatokat ténylegesen besorolták - figyelmet fordítva az adott szinten fontos tulajdonságok kiemelésére. Amit így leírtak, az a legfőbb jellemző (tulajdonság, felépítés), amely a rangsort meghatározza (Mullis 1980). Ennek gyakori változata, amikor a rangsor helyett egy adott számú csoportra osztják szét a teljesítményeket. A klasszikus megközelítésnek van egy érdekes multidimenzionális változata is. Ebben a változatban a kulcsfontosságú vonások azonosítása alapján (fent az 5. számú) először azt döntenek el, hogy melyek a legjelentősebb vonások. Ezt követően a mintákat minden egyes vonás esetében külön sorrendbe állítják. Tehát végül inkább egy analitikus, többjellemezős skála jön létre, mint egy holisztikus, a legfőbb jellemzőt megjelenítő skála.
7. **Kétpólusú döntések:** A legfőbb jellemzőn alapuló módszer egy másik változatában a reprezentatív mintákat először szintek szerint csoportosítják. Ezt követően a szintek határait összpontosító megbeszélés során azonosítják a főbb tulajdonságokat (mint a fenti 5. számú megközelítés). Az adott tulajdonságot meghatározó kritériumot azonban rövid eldöntendő kérdés formájában fogalmazzák meg. Ily módon két-

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

pólusú választások ágrajza hozható létre. Ez az értékelők számára a követendő döntések algoritmusát képezi (Upshur és Turner 1995).

8. **Vélemények összehasonlítása:** teljesítménypárok csoportos megbeszélése, hogy megfogalmazzák, hogy melyik jobb, és miért. Ily módon azonosítják a pontozók által használt metanyelvi kategóriákat, valamint az egyes szintekre vonatkozó fontos jellemzőket. Ezt követően ezek a jellemzők szintleírásokká alakíthatók (Pollitt és Murray 1996).
9. **Csoportosítás:** Amikor már léteznek a szintleírások tervezetei, az adatközlőket meg lehet kérni arra, hogy csoportosítsák őket azoknak a kategóriáknak megfelelően, amelyek leírására szolgálnak és/vagy a szintek alapján. Az adatközlőket arra is meg lehet kérni, hogy véleményezzék, átírjanak/kijavítsanak és/vagy kiselejtezzenek egyes szintleírásokat, valamint arra, hogy határozzák meg, mely szintleírások különösen világosak, hasznosak, megfelelőek stb. Így alakították ki és szerkesztették meg azt a szintleírás-gyűjteményt, amelyből a szemléltető skálákat hozták létre (Smith és Kendall 1963; North 1996/2000).

Kvantitatív módszerek:

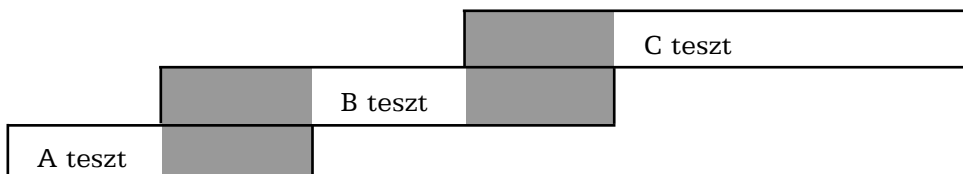
Ezek a módszerek jelentős mértékű statisztikai elemzést alkalmaznak, és az eredmények gondos értelmezését igénylik.

10. **Diszkrimináns elemzés:** Először részletes diskurzuselemzésnek vetnek alá már (lehetőleg csoport által) értékelt teljesítménymintákat. Ez a kvalitatív elemzés azonosítja és megszámlolja a különböző kvalitatív jellemzők gyakoriságát. Ezt követően többszörös regresszió segítségével meghatározzák, hogy az azonosított jellemzők melyike játszik láthatóan jelentős szerepet az értékelők által adott pontokban. Ezeket a fontos jellemzőket végül minden szinten felhasználják a szintleírások kialakításánál (Fulcher 1996).
11. **Multidimenzionális szintezés:** A név ellenére leíró eljárásról van szó, amelynek feladata a főbb jellemzők és a közöttük fennálló kapcsolat azonosítása. A teljesítményeket számos kategória analitikus szintrendszere alapján pontozzák. Az elemzési eljárás eredménye megmutatja, hogy mely kategóriák játszottak ténylegesen döntő szerepet a szint meghatározásában, és olyan diagramot nyújt, amely feltérképezi a különböző kategóriák egymáshoz viszonyított közelségét vagy távolságát. Ez tehát egy kutatási eljárás a fontos kritériumok meghatározására és érvényességének megállapítására (Chaloub-Deville 1995).

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

12. Valószínűségi tesztelmélet (Item Response Theory, IRT) vagy „Látens jellemző” elemzés: A valószínűségi tesztelmélet mérési vagy szintezési modellek családját alkotja. A legegyszerűbb és legerőteljesebb ezek közül a dán matematikus, George Rasch után elnevezett Rasch - modell. A valószínűségi tesztelmélet a valószínűség-elméletből fejlődött ki, és elsősorban arra használják, hogy egy adott feladatbank egyes teszt-iteleinek nehézségi szintjét meghatározzák. Ha valaki haladó szinten van, akkor nagy esélye van arra, hogy megválasszon egy alapszintű kérdést, ha viszont valaki alapszinten van, akkor nagyon kevés esélye van egy haladó szintű kérdés megválaszolására. Ezt az egyszerű tényt fejlesztették szintezési módszerré a Rasch - modellben, amely felhasználható az egyazon skála itemeinek beméréséhez. A megközelítés egyik továbbfejlesztett változata használható mind a kommunikatív készségszint, mind a teszt-itelek szintleírásainak szintezésekor is.

Rasch - elemzés során a különböző tesztek vagy kérdőívek egymást részlegesen átfedő láncolattá alakíthatók ún. „horgony-itelek²” segítségével, amelyek mind a két szomszédos tesztverzióra jellemzők. Az alábbi diagramon szürkével jelöltük a „horgony-iteleket”. Ily módon a különböző tesztverziók adott tanulói csoportokat célozhatnak meg, és mégis egy közös skálához kapcsolhatók. Ennél a folyamatnál azonban elővigyázatosnak kell lennünk, mivel a modell minden tesztverzió esetében eltorzítja a magas és alacsony pontszámokat.



A Rasch - elemzés előnye, hogy mintafüggetlen, skálamentes mérést tesz lehetővé, vagyis olyan szintezést, ami független az elemzés során használt mintáktól vagy tesztek-től/kérdőívektől. Olyan skálaértékeket biztosít, amelyek az eljövendő csoportok esetén is állandóak maradnak, amennyiben azok az új csoportok ugyanazon statisztikai populációhoz tartozónak tekinthetők. Az idővel bekövetkező szisztematikus értékmódosulások (pl. a tanterv változása vagy az értékelők kiképzése miatt) számszerűsíthetők és kiigazíthatók. A tanulói vagy értékelői típusok közötti szisztematikus variáció szintén számszerűsíthető és kiigazítható (Wright és Masters 1982; Lincare 1989).

² Jelentése: „közös elem”. (A ford.)

„A” Függelék: A készségszintek leírására szolgáló szintleírások kialakítása

A Rasch elemzés sokféleképpen használható a szintleírások színtezésében:

- a) A 6., 7. és 8. kvalitatív eljárások során nyert adatok a Rasch - elemzés segítségével számskálán helyezhetők el.
- b) A tesztek gondosan összeállíthatók úgy, hogy az egyes teszt-ítemek a készségszint leírásainak megfeleljenek. Ezt követően a teszt-ítemek a Rasch - elemzés segítségével színtezhetők, és a skálán elfoglalt értékük jelzi a szintleírások relatív nehézségi szintjét (Brown et al. 1992; Carroll 1993; Masters 1994; Kirsch és Mosenthal 1995).
- c) A szintleírások felhasználhatók a tanári értékelés során a kérdőív részeként (Meg tudja tenni X dolgot?). Ily módon a szintleírások közvetlenül a számskálára mérhetők be ugyanúgy, ahogy a teszt-ítemeket színtezik a feladatbankban.
- d) A 3., 4. és 5. fejezetben használt szintleíró skálákat így fejlesztették ki. A „B”, „C” és „D” jelű függelékben ismertetett mindhárom kutatási program a Rasch - módszert használta a leírások színtezésére és a létrejövő szintleíró skálák egymáshoz igazítására.

A skálák kifejlesztésében nyújtott haszna mellett a Rasch - módszer felhasználható annak elemzésére, hogyan használják az értékelési skála sávjait a valóságban. Ez segíthet a pontatlan megfogalmazás, illetve egy sáv túl sok vagy túl kevés használatának felismerésében és az átdolgozásban (Davidson 1992; Milanovic et al. 1996; Stansfield és Kenyon 1996; Tyndall és Kenyon 1996).

A Referenciakeret használói gondolkozzanak el, és ahol lehet, fogalmazzák meg, hogy

- *a rendszerükben adott érdemjegyek milyen mértékben alapulnak a közös meghatározások egységes értelmezésén;*
- *a fent ismertetett módszerek melyikét - illetve milyen más módszereket - használnak az ilyen meghatározások kialakításához.*

„A” Függelék: A készség szintek leírására szolgáló szintleírások kialakítása

Annotált bibliográfia-válogatás: a nyelvtudás szintezése

- Alderson, J. C. (1991): Bands and scores. In Alderson, J. C. - North, B. (szerk.): *Language testing in the 1990s*. London: British Council/Macmillan, Developments in ELT, 71-86.
- Brindley, G. (1991): Defining language ability: the criteria for criteria. In Anivan, S. (szerk.) *Current developments in language testing*, Singapore, Regional Language Centre.
- Brindley, G. (1998): Outcomes-based assessment and reporting in language learning programmes, a review of the issues. *Language Testing* 15 (1), 45 - 85.
- Brown, Annie - Elder, Cathie - Lumley, Tom - McNamara, Tim és McQueen, J. (1992): *Mapping abilities and skill levels using Rasch techniques*. Paper presented at the 14th Language Testing Research Colloquium, Vancouver. Reprinted in: *Melbourne Papers in Applied Linguistics* 1/1, 37-69.
- Carroll, J. B. (1993): Test theory and behavioural scaling of test performance. In Frederiksen, N., Mislevy, R. J. - Bejar, I. I. (szerk.) *Test theory for a new generation of tests*. Hillsdale N.J.: Lawrence Erlbaum Associates: 297 - 323.
- Chaloub-Deville M. (1995): Deriving oral assessment scales across different tests and rater groups. *Language Testing* 12 (1), 16 - 33.
- Davidson, F. (1992): Statistical support for training in ESL composition rating. In Hamp-Lyons (szerk.): *Assessing second language writing in academic contexts*. Norwood N.J.: Ablex: 155 - 166.
- Fulcher (1996): Does thick description lead to smart tests?
A data-based approach to rating scale construction. *Language Testing* 13 (2), 208 - 38.
- Gipps, C. (1994): *Beyond testing*. London: Palmer Press.
- A cél és az irányultság összekeveréséből adódó problémákat tárgyalja, valamint az IELTS beszédtevékenységet leíró skáláit ismerteti.*
- Elméleti alapokon nyugvó kritika a készség szintrendszer azon állításáról, hogy a kritériumorientált értékelési eljárást képviselik.*
- Kritizálja, hogy nagy figyelmet fordítanak az olyan eredményekre, amelyek azt hangsúlyozzák, mit tud megtenni a tanuló, ahelyett hogy a kialakuló kompetencia aspektusait állítanák középpontba.*
- A Rasch - módszer klasszikus használata a teszt-itekek szintezésében, a különböző itemekben tesztelt szövegolvasási feladatok készség szintrendszerének létrehozásához.*
- Nagy jelentőségű cikk, amely a Rasch - módszer használatát javasolta teszt-itekek szintezéséhez és ezáltal a készség szintrendszer létrehozásához.*
- Ebben a tanulmányban a tanulók teljesítményének véleményezésekor az arab anyanyelvű beszélők által használt kritériumokat elemzik. Valójában ez a multidimenzionális szintezés egyetlen alkalmazása a nyelvi tesztelésben.*
- Nagyon világos beszámoló egy pontozó skála érvényességének a Rasch elemzés segítségével, ciklikus folyamattal történő bizonyításáról. A szintezésben inkább „szemantikai” mint „konkrét” megközelítést javasol, például a szemléltető szintleírások esetében.*
- A szintleírás és a skála kialakításának szisztematikus megközelítése, amely annak az elemzéséből indul ki, hogy tulajdonképpen mi is megy végbe performancia közben. Nagyon időigényes módszer.*
- A tanárok egymással való kapcsolattartásának segítségével létrehozott közös referenciapontokhoz kapcsolódó tanári „normaorientált értékelés” mellett foglal állást. Az Angol Nemzeti Alaptanterv ködös szintleírásai által okozott problémák megvitatása. Átfogó tantervek.*

„A” Függelék: A készség szintek leírására szolgáló szintleírások kialakítása

Kirsch, I. S. (1995): Literacy performance on three scales: definitions and results. In *Literacy, economy and society: Results of the first international literacy survey*. Párizs, Gazdasági Együttműködési és Fejlesztési Szervezet (OECD): 27 - 53.

Kirsch, I. S. - Mosenthal, P. B. (1995): Interpreting the IEA reading literacy scales. In Binkley, M. - Rust., K. - Winglee, M. (szerk.) *Methodological issues in comparative educational studies: The case of the IEA reading literacy study*. Washington D.C.: US Department of Education, National Center for Education Statistics: 135 - 192.

Linacre, J. M. (1989): *Multi-faceted Measurement*. Chicago: MESA Press.

Liskin-Gasparro, J. E. (1984): The ACTFL proficiency guidelines: Gateway to testing and curriculum. In *Foreign Language Annals* 17/5, 475-489.

Lowe, P. (1985): The ILR proficiency scale as a synthesising research principle: the view from the mountain. In James, C. J. (szerk.): *Foreign Language Proficiency in the Classroom and Beyond*. Lincolnwood (Ill.): National Textbook Company.

Lowe, P. (1986): Proficiency: panacea, framework, process? A Reply to Kramersch, Schulz, and particularly, to Bachman and Savignon. In: *Modern Language Journal* 70/4, 391-397.

Masters, G. (1994): Profiles and assessment. *Curriculum Perspectives* 14,1: 48 - 52.

Milanovic, M. - Saville, N. - Pollitt, A. - Cook, A. (1996): Developing rating scales for CASE: Theoretical concerns and analyses. In Cumming, A. - Berwick, R. *Validation in language testing*. Clevedon, Avon, Multimedia Matters: 15 - 38.

Mullis, I. V. S. (1981): *Using the primary trait system for evaluating writing*. Manuscript No. 10-W-51. Princeton N.J.: Educational Testing Service.

Egyszerű, statisztikával nem túlterhelt beszámoló a Rasch - modell komplex felhasználásáról tesztek adatai alapján készített szintskálák létrehozásához. Azért fejlesztették ki ezt a módszert, hogy megjósolják és megmagyarázzák új teszt-itelemek nehézségi szintjét a vonatkozó feladatok és kompetenciák alapján - vagyis egy referenciakerethez viszonyítva.

A fent említett mű részletesebb és technikaibb jellegű változata, a módszer fejlődését három egymáshoz kapcsolódó projekten át követi nyomon.

A statisztika területén áttörést jelentő mű, amely lehetővé teszi a vizsgázatók szigorúságának figyelembevételét egy mérés eredményének megjelenítésekor. A projekt során az iskolai évfolyamok és a szintek egymáshoz való viszonyításának ellenőrzésére alkalmazható szintleírások kialakítására használták.

Az amerikai ACTFL skálának a Foreign Service Institute (FSI) skálából való kialakításának céljait és kidolgozását ismerteti.

Az amerikai Interagency Language Roundtable (ILR) skála FSI skálából való kialakításának részletes leírása. A skála funkciói.

Egy adott kontextusban jól működő rendszer megvédése az elméleti kritikával szemben, amit a skála nagy elterjedése és az interjú típusú módszer váltott ki (az ACTFL-lel).

Rövid ismertetése annak, hogyan a Rasch - módszert a teszteredmények és a tanári értékelések szintezésére használták Ausztráliában egy tantervi profilkészítő rendszer létrehozásához.

Klasszikus beszámoló a Rasch - módszer alkalmazásáról egy szóbeli vizsga pontozó skálájának finomításához - annyira csökkenti a skálán található szintek számát, amennyit az értékelők hatékonyan tudnak kezelni.

Klasszikus ismertetés a legfőbb jellemzőn alapuló módszer alkalmazásáról egy értékelési skála kialakításához az anyanyelvi írás területén.

„A” Függelék: A készség szintek leírására szolgáló szintleírások kialakítása

North, B. (1993): *The development of descriptors on scales of proficiency: perspectives, problems, and a possible methodology*. NFLC Occasional Paper, National Foreign Language Center, Washington D.C., April 1993.

North, B. (1994): *Scales of language proficiency: a survey of some existing systems*, Strasbourg, Európa Tanács CC-LANG (94) 24.

North, B. (1996/2000): *The development of a common framework scale of language proficiency*. PhD - disszertáció, Thames Valley University. Újranyomás 2000, New York, Peter Lang.

North B. (megjelenés előtt): *Scales for rating language performance in language tests: descriptive models, formulation styles and presentation formats*. TOEFL Research Paper. Princeton NJ; Educational Testing Service.

North, B. - Schneider, G. (1998): *Scaling descriptors for language proficiency scales*. *Language Testing* 15/2: 217 - 262.

Pollitt, A. - Murray, N. L. (1996): *What raters really pay attention to*. In Milanovic, M. - Saville, N. (szerk.) (1996): *Performance testing, cognition and assessment*. Studies in Language Testing 3. Selected papers from the 15th Language Testing Research Colloquium, Cambridge and Arnhem, 2 - 4 August 1993. Cambridge: University of Cambridge Local Examinations Syndicate: 74 - 91.

Scarino, A. (1996): *Issues in planning, describing and monitoring long-term progress in language learning*. In Proceedings of the AFMLTA 10th National Languages Conference: 67-75.

Scarino, A. (1997): *Analysing the language of frameworks of outcomes for foreign language learning*. In Proceedings of the AFMLTA 11th National Languages Conference: 241 - 258.

Schneider, G - North, B. (1999): *„In anderen Sprachen kann ich...“ Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeiten*. Bern / Aarau: NFP 33 / SKBF (Umsetzungsbericht).

A hagyományos készség szintskálák tartalmának és kialakítási módszerének kritikája. Olyan projekt szükségességének kimutatása, amely során a szemléltető skálákat tanárok közreműködésével dolgozzák ki, majd a tanári értékelések alapján a Rasch - modell segítségével szintezik.

A tantervi skálák és pontozó skálák átfogó áttekintése, majd elemzése és kiindulópontként való felhasználása a szemléltető skálákat kialakító projektben.

A készség szintskálák ismertetése, illetve, annak bemutatása, hogyan viszonyulnak a skálákhoz a kompetencia és nyelvhasználati modellek. A projekt fejlesztési szakaszainak részletes ismertetése, amelyek során megszülettek a szemléltető szintleírások - felmerülő problémák, megoldások.

A szóbeli és írásbeli tesztek során használt pontozó skálák típusainak részletes elemzése és történeti áttekintése: előnyök, hátrányok, buktatók stb.

A szemléltető szintleírásokat létrehozó kutatási program ismertetése. Bemutatja az eredményeket és a skála stabilitását. Egy függelékben példát ad az eszközökre és az eredményekre.

Érdekes módszertani cikk, amely egyszerű színtezési eljárással kapcsolja össze a gyűjteményes rácselemzést annak megállapítása érdekében, hogy mire összpontosítanak az értékelők a különböző készség szinteken.

A homályos megfogalmazás kritikáját adja, továbbá kritikával illeti az Egyesült Királyságban és Ausztráliában tipikus tantervi profilok tanári értékelésről szóló leírásaiban tapasztalható információhiányt azzal kapcsolatban, hogy a tanulók milyen jól teljesítenek.

Mint fent.

Rövid ismertetés arról a kutatási programról, amely létrehozta a szemléltető skálákat. A Nyelvtanulási Napló svájci változatát is bemutatja (40 oldal A5).

„A” Független: A készség szintek leírására szolgáló szintleírások kialakítása

Schneider, G. - North, B. (2000): „Dans d'autres langues, je suis capable de...” Echelles pour la description, l'évaluation et l'auto-évaluation des compétences en langues étrangères. Berne / Aarau PNR33/CSRE (rapport de valorisation).

Schneider, G. - North, B. (2000): Fremdsprachen können - was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit. Chur/Zürich, Verlag Rüegger AG.

Skehan, P. (1984): Issues in the testing of English for specific purposes. In Language Testing 1/2, 202-220.

Shohamy, E. - Gordon, C. M. - Kraemer, R. (1992): The effect of raters' background and training on the reliability of direct writing tests. Modern Language Journal 76: 27 - 33.

Smith, P. C. - Kendall, J. M. (1963): Retranslation of expectations: an approach to the construction of unambiguous anchors for rating scales. In: Journal of Applied Psychology, 47/2.

Stansfield, C. W. - Kenyon, D. M. (1996): Comparing the scaling of speaking tasks by language teachers and the ACTFL guidelines. In Cumming, A. and Berwick, R. Validation in language testing. Clevedon, Avon, Multimedia Matters: 124 - 153.

Takala, S. - F. Kaftandjieva (megjelenés előtt). Council of Europe scales of language proficiency: A validation study. In J. C. Alderson (szerk.) Case studies of the use of the Common European Framework. Európa Tanács.

Tyndall, B. - Kenyon, D. (1996): Validation of a new holistic rating scale using Rasch multifaceted analysis. In Cumming, A. and Berwick, R. Validation in language testing. Clevedon, Avon, Multimedia Matters: 9 - 57.

Upshur, J. - Turner, C. (1995): Constructing rating scales for second language tests. English Language Teaching Journal 49 (1), 3 - 12.

Wilds, C. P. (1975): The oral interview test. In: Spolsky, B. and Jones, R. (szerk.): Testing language proficiency. Washington D.C.: Center for Applied Linguistics, 29 - 44.

Mint fent.

A szemléltető skálákat kialakító kutatási program teljes körű bemutatása. Lényegre törő angol nyelvű fejezet a szintezésről. Szintén ismerteti a Nyelvtanulási Napló svájci változatát.

Az ELTS skálák normaorientált jellegének és relatív megfogalmazásának kritikáját adja.

Egy analitikus íráskészségre vonatkozó skála kialakításához használt alapvető kvalitatív módszer egyszerű ismertetése. A módszer meglepően magas értékelők közötti megbízhatóságot eredményezett olyan értékelők esetében, akik nem tesztelési szakértők voltak, és akik külön képzésben nem részesültek.

Az első olyan megközelítés, amely a szintleírások szintezésére és nem csupán skálák írására irányult. Úttörő jellegű munka. Nagyon nehéz olvasmány.

A Rasch - módszer használata az ACTFL irányelvekben megjelenő feladatok rangsorolásának megerősítéséhez. Érdekes módszertani tanulmány, amely a projekt során alkalmazott módszert inspirálta a szemléltető skálák kialakításához.

A Rasc - modell továbbfejlesztett változatának használatát mutatja be a nyelvi önértékelések szintezésében a szemléltető skálák adaptációjának vonatkozásában. Kontextus: DIALANG projekt: kísérletek a finn nyelvvel kapcsolatban.

Egyetemi felvételnél alkalmazott angol mint második nyelv (ESL) besorolási elbeszélgetéseknél alkalmazott skála érvényesség bizonyításának egyszerű ismertetése. A sokoldalú Rasch - modell klasszikus használata a képzési szükségletek azonosítására.

A legfőbb jellemzőn alapuló eljárás kifinomult továbbfejlesztése kétpólusú döntések diagramjainak létrehozásához. Nagyon fontos az iskolai szektorban.

Az eredeti nyelvi készség szintet pontozó skála eredeti megjelenése. Erdemes alaposan elolvasni olyan árnyalatok miatt, amelyek a legtöbb, azóta kialakult interjú típusú megközelítésben elvesztek.