

9. Mérés és értékelés

9.1. Bevezetés

A mérés¹ kifejezést ebben a fejezetben a nyelvhasználó készségszintjének mérésére vonatkozóan használjuk. Minden nyelvi teszt a mérés valamilyen formája, de számos olyan mérési forma létezik (pl. a folyamatos mérés során használt kérdőívek, informális tanári megfigyelés), amely nem nevezhető tesztnek. Az értékelés a mérésnél tágabb kategória. Minden mérés az értékelés körébe tartozik, de egy nyelvi programban a tanuló készségszintjén kívül számos más dolgot is értékelnek. Ezek közé tartozhat bizonyos módszerek vagy tananyagok hatékonysága, a program során ténylegesen létrehozott diskurzus fajtája és minősége, a tanár/tanuló elégedettsége, a tanítás hatékonysága stb. Ez a fejezet a nyelvtanuló nyelvtudásának/készségszintjének mérésével és értékelésével foglalkozik, és nem a program értékelésének tágabb kérdéseivel.

Hagyományosan három fontosnak tekintett fogalom létezik, amely alapvető a mérés és értékelés bármiféle tárgyalásakor: érvényesség, megbízhatóság és kivitelezhetőség. Az ebben a fejezetben ismertetett kérdések szempontjából jó tudni, hogy ezek a fogalmak mit jelentenek, hogyan viszonyulnak egymáshoz, és milyen szerepet töltenek be a Referenciakeretben.

Az *érvényesség* az a fogalom, amely jelentős a Referenciakeret szempontjából. Egy teszt vagy mérési eljárás olyan mértékben mondható érvényesnek, amennyire kimutatható, hogy valójában azt a felépítést (construct) méri, amit az adott kontextusban mérni *kell*, és hogy a szerzett információ az érintett vizsgázó(k) készségszintjét pontosan tükrözi.

A *megbízhatóság* viszont szakkifejezés. Alapjában véve azt a mértéket jelöli, amennyire ugyanazon mérés két különböző alkalommal történő (valódi vagy szimulált) végrehajtása során a vizsgázók ugyanolyan rangsorba kerülnek.

¹ Nincs pontos egy az egybeni megfelelés az angol és a magyar terminológia között, sőt még az angol nyelven belül is vannak értelmezési különbségek. Az ebben a fejezetben használt „mérés” és „értékelés” kifejezések a tanulók tudásának/készségszintjének/teljesítményének a megállapítására vonatkoznak. (A ford.)

Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

Ami valójában a megbízhatóságnál fontosabb, az a követelményekkel kapcsolatban meghozott *döntések helyessége*. Ha a mérés megfelelt/nem felelt meg, vagy A2+/B1/B1+ szintekben adja meg az eredményt, mennyire pontosak ezek a döntések? A döntések helyessége az adott követelménynek (pl. B1 szint) az adott kontextusra vonatkoztatott érvényességétől függ. Ezen kívül a döntés meghozatalához használt kritériumok érvényességétől és a kritériumok kialakításához alkalmazott eljárások érvényességétől is függ.

Ha két eltérő szervezet vagy régió ugyanazon követelményeken alapuló kritériumokat használ az ugyanazon készséggel kapcsolatos mérési döntései meghozatalához, ha maguk a követelmények érvényesek és megfelelőek mind a két érintett kontextusban, és ha a követelményeket a mérési feladatok tervezése és a teljesítmény értelmezése során következetesen értelmezik, akkor a két rendszer eredményei egymással összhangban vannak. Hagyományosan a két ugyanazt mérő teszt közötti korrelációt „megegyező érvényességnek” (*cuncurrent validity*) nevezzük. Ez a fogalom természetesen kapcsolódik a megbízhatósághoz, hiszen a nem megbízhatóan mérő tesztek nem korrelálnak egymással. Azonban ennél központibb kérdés a két teszt közötti összhang a *mérés tárgyával* és a *teljesítmény értelmezésének módjával* kapcsolatban.

A Közös Európai Referenciakeret ezzel a két kérdéssel foglalkozik. A következő rész ismerteti a Referenciakeret használatának három fő módját.

1. Tesztek és vizsgák részletes *mit mérünk* tartalmi követelményeihez:
2. A tanulási cél elérésének *hogyan értelmezzük a teljesítményt* meghatározásához használt kritériumok megfogalmazásához:
3. A létező tesztek és vizsgák *hogyan lehet összehasonlításokat tenni* készségszintjének leírásához, ami lehetővé teszi a minősítési rendszerek közötti összehasonlításokat:

Ezek a kérdések különbözőképpen kapcsolódnak a mérés és értékelés különböző módjaihoz. A mérésnek és értékelésnek számos fajtája és hagyománya van. Hiba feltételezni, hogy egy megközelítés (pl. egy külső értékelésű vizsga) szükségszerűen felette áll egy másik megközelítés oktatási hatásának (pl. tanári mérés és értékelés). A közös követelményrendszereknek - mint például a Közös Referenciaszinteknek - hatalmas előnye, hogy lehetővé teszik a mérés és értékelés különböző formáinak egymáshoz való hasonlítását.

A fejezet harmadik részében különböző mérési és értékelési módokat ismertetünk. A választási lehetőségeket ellentétpárok formájában mutatjuk be. Minden esetben meghatározzuk a használt kifejezéseket, valamint a mérés és értékelés oktatási kontextusban betöltött célja alapján ismertetjük a viszonylagos előnyöket és hátrányokat. Megvitatjuk az egyik vagy másik alternatív választás gyakorlati következményeit is. Ezután bemutatjuk a Referenciakeret ilyenfajta mérésben és értékelésben betöltött szerepét.

Egy mérési folyamatnak gyakorlatiasnak, *kivitelezhetőnek* is kell lennie. A kivitelezhetőség különösen fontos a teljesítmény tesztelésénél. Az értékelők ilyenkor az idő szorításában dolgoznak. Csupán egy korlátozott mintát látnak a teljesítményből, és erősen korlátozott azoknak a kategóriáknak a típusa és száma, amelyeket kritériumként kezelni tudnak. A Referenciakeret viszonyítási pontot és nem gyakorlati értékelési eszközt kíván nyújtani. Ezért a Referenciakeretnek átfogónak kell lennie, a felhasználóknak viszont szelektálniuk kell. A szelekció jelentheti akár valamilyen egyszerűbb alkalmazási rendszer használatát is, amelyben a Referenciakeret által különválasztott kategóriákat összevonják. Például a 4. és 5. fejezetben használt, a szöveget kísérő szemléltető szintleíró skálák gyakran jóval egyszerűbbek, mint a magában a szövegben megvitatott kategóriák és magyarázatok. A fejezet befejező része ezt a kérdést tárgyalja példákkal illusztrálva.

9.2. A Referenciakeret mint mérési és értékelési háttér

9.2.1. A tesztek és vizsgák részletes tartalmi követelményei

Valamilyen kommunikatív mérés és értékelés feladat-meghatározásának elkészítéséhez használható a 4. fejezetben található „nyelvhasználat és a nyelvhasználó” című leírás, valamint a „kommunikatív nyelvi tevékenységek” című 4.4. fejezet. Egyre inkább elterjedőben van az a nézet, miszerint az érvényes méréshez a diskurzus különböző megfelelő fajtáiból vett minták szükségesek. Például egy, a beszéd tesztelésére vonatkozó, nemrégiben kidolgozott teszt alátámasztja ezt a feltételezést. Az első rész egy szimulált *társalgás*, ami bemelegítésre szolgál, a második egy *informális eszmecsere* olyan aktuális témákkal kapcsolatban, amelyeket a vizsgázó érdekesnek ítél. Ezt követi a *tranzakciós* fázis, amely vagy szemtől szemben lejátszódó vagy szimulált telefonos információkérésre irányuló tevékenység formájában történik. Ezután következik az írott *beszámolón* alapuló *produkciós* szakasz, amelyben a *vizsgázó* szakterületének és terveinek *leírását* adja. Végül következik egy *célirányos együttműködés*, amelynek során a vizsgázóknak egyetértésre kell jutniuk.

Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

Mindent egybevetve, a Referenciakeret által a kommunikatív tevékenységek leírásához használt kategóriák a következők:

	Interakció (Rögtönzött, rövid megszólalások)	Produkcó (Előkészített, hosszú megszólalások)
Szóbeli:	<i>Társalgás</i> <i>Informális eszmecsere</i> <i>Célirányos együttműködés</i>	A szakterület <i>leírása</i>
Írásbeli:		<i>Beszámoló/leírás</i> a vizsgázó szakterületéről

220

A feladat-meghatározások részleteinek kidolgozásakor a felhasználó folyamaodhat a 4.1. fejezethez a nyelvhasználat kontextusaival kapcsolatban (tartományok, feltételek és korlátozások, mentális kontextus), a 4.6. fejezethez a szövegekkel kapcsolatban, és a 7. fejezethez a feladatok nyelvoktatásban betöltött szerepével kapcsolatban, különösképpen a 7.3. fejezethez a feladatok nehézségi szintjeit illetően.

Az 5.2. fejezet, amely a kommunikatív nyelvi kompetenciákról szól, segíti a teszt-itelemek létrehozását vagy egy szóbeli teszt szakaszainak kialakítását, a megfelelő nyelvi, szociolingvisztikai és pragmatikai kompetenciák meglétének vizsgálata érdekében. Az Európa Tanács által készített, több mint 20 nyelvre vonatkozó részletes tartalmi követelmények (lásd a Bibliográfiát) - a **B1** szintre, valamint az angol nyelv esetében az **A2** szintre és a **B2** szintre vonatkozóan is, továbbá a más nyelvekre és szintekre a későbbiekben kidolgozandó ekvivalens részletes követelmények - a fő Referenciakeret mellékleteinek tekinthetők. A tesztkészítés elősegítésére további részleteket biztosítanak az A1, A2, B1 és B2 szinten.

9.2.2. A tanulási cél elérésének kritériumai

A skálák segítséget adnak egy bizonyos tanulási cél elérésének mérését szolgáló pontozó skálák elkészítéséhez, a szintleírások pedig segíthetnek a kritériumok megfogalmazásában. A cél lehet az általános nyelvtudás egy tág szintje, ami az egyik Közös Referenciaszintnek felel meg (pl. B1). Másrészt lehet a tevékenységek, készségek és kompetenciák sajátos keveréke, amint azt a „Részleges kompetenciák és a célkitűzések különbözősége a Referenciakeret viszonylatában” tárgyú 6.1.4. fejezetben ismertettük. Az ilyen moduláris célkitűzést a 2. táblázatban bemutatott szintezett kategória-rendszer alapján lehet leírni.

A szintleírások használatának elemzésekor fontos különbséget tenni:

1. A 4. fejezetben ismertetett kommunikatív tevékenységek szintleírásai
2. és a nyelvtudás aspektusainak az 5. fejezetben ismertetett egyes kompetenciákra vonatkozó szintleírásai között.

Az előbbiek nagyon alkalmasak a valós világ feladataira vonatkozó tanári és önértékeléshez. Az ilyen tanári, illetve önértékelés a tanuló kurzus folyamán kialakított nyelvismeretének részletes képe alapján történik. Azért előnyös, mert elősegíti, hogy mind a tanár, mind a tanuló a cselekvésközpontú megközelítésre összpontosítson.

Általában nem tanácsos azonban a kommunikatív tevékenységek meghatározásainak használata egy adott szóbeli vagy írásbeli teszt értékelési kritériumai között, amikor az eredményeket az elért készségszint alapján szeretnénk kifejezni. Ez azért van így, mert a készségszint megállapításához a mérésnek nem elsősorban egy bizonyos teljesítményre kell összpontosítania, hanem inkább az adott teljesítmény alapján kimutatott általánosítható kompetenciák megítélése a feladata. Természetesen lehet jó pedagógiai oka annak, hogy egy adott tevékenység elvégzésének sikerét értékeljük, különösen a fiatalabb minimumszintű és alapszintű nyelvhasználók esetében (A1 és A2 szintek). Az ilyen eredmények kevésbé általánosíthatóak, de a nyelvtanulás korai szakaszában általában nem fontos az eredmények általánosíthatósága.

Ez megerősíti azt a tényt, hogy a mérésnek különböző funkciói lehetnek. Ami az egyik mérési célnak megfelel, az nem felel meg a másiknak.

9.2.2.1. A kommunikatív tevékenységek szintleírásai

A kommunikatív tevékenységek szintleírásai (4. fejezet) három különböző módon használhatók a célok eléréséhez kapcsolódóan.

1. **Tesztkészítés:** ahogy ezt a 9.2.1. fejezetben ismertettük, a kommunikatív tevékenységek skálái segíthetnek a mérés feladatainak tervezéséhez szükséges részletes követelmények meghatározásában.
2. **Megjelenítés:** a kommunikatív tevékenységek skálái az eredmények megjelenítésében is nagyon hasznosak lehetnek. Az oktatási rendszer produktumainak felhasználói, például a munkáltatók, gyakran inkább az átfogó eredmények, mintsem a kompetencia részletes leírása iránt érdeklődnek.
3. **Önértékelés vagy tanári értékelés:** végül a kommunikatív tevékenységek szintleírásai többféleképpen használhatók önértékelésre és tanári értékelésre. Például:
 - *Lista:* Folyamatos értékeléskor vagy a kurzus végén összegző értékelésre. Az adott szint leírásai felsorolhatóak. Más esetben a szint-

Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

leírások tartalma „felbontható”. Például a szintleírás *Tud személyes információt kérni és megadni, felbontható a bele értendő alkotóelemekre: Be tudok mutatkozni; Meg tudom mondani, hol lakom; Meg tudom mondani a címem franciául; Meg tudom mondani, hány éves vagyok stb. és Meg tudom mástól kérdezni, hogy hívják; Meg tudom kérdezni mástól, hogy hol lakik; Meg tudom mástól kérdezni, hány éves stb.*

- **Értékelő rács:** Folyamatos vagy összegző értékeléskor a különböző szinteken meghatározott (B1+, B2, B2+), kiválasztott kategóriák (pl. *Társalgás, Eszmecsere, Információcsere*) táblázata alapján készíthető el a tanulói profil.

222

Az utóbbi tíz évben gyakoribbá vált a szintleírások ilyen jellegű használata. A gyakorlat azt mutatja, hogy növelhető a tanárok és tanulók következetesége a szintleírások értelmezésében, ha a szintleírások nem csupán azt írják le, hogy MIT képes megtenni a tanuló, hanem azt is, hogy ezt MILYEN JÓL teszi.

9.2.2.2. A nyelvtudás aspektusainak bizonyos kompetenciákra vonatkozó szintleírásai

A nyelvtudás aspektusainak szintleírásai a célok elérésére vonatkozóan két fő módon használhatók.

1. **Önértékelés vagy tanári értékelés:** Amennyiben a szintleírások pozitív, önálló állítások, akkor alkalmazhatóak az önértékelés és tanári értékelés listáiban. Azonban a meglévő skálák többségének az a hibája, hogy alacsonyabb szinteken gyakran negatív a megfogalmazása, és a skála közepe táján normafüggő. Gyakran csupán verbális megkülönböztetést tesznek a szintek között, vagyis a szomszédos szinteknél egy vagy két szót változtatnak meg, ami a skálán kívüli szövegben kevés jelentéssel bír. Az „A” Függelék tárgyalja röviden az olyan szintleírások kialakítási módjait, amelyek elkerülik ezeket a problémákat.
2. **Teljesítményértékelés:** Az 5. fejezetben ismertetett kompetenciák aspektusaira vonatkozó szintleíró skálák magától értetődőbb használata az, hogy kiindulási pontot ad a mérési kritériumok kialakításához. Azáltal, hogy a személyes, nem szisztematikus benyomásokat átgondolt döntésekké alakítják, az ilyen leírások segítenek egy közös referenciakeret létrehozásában az érintett értékelők számára.

Alapvetően háromféleképpen jelenhetnek meg a szintleírások a mérési kritériumokban:

- A szintleírások megjeleníthetők *skálaként* - gyakran a különböző kategóriák szintleírásainak szintenként egyetlen holisztikus bekezdésbe való összefoglalásával. Ez nagyon gyakori megközelítés.
- Megjeleníthetők *lista* formájában, általában lényeges szintenként egy listában, és a szintleírások gyakran alcímek, azaz kategóriák köré csoportosulnak. A listák kevésbé gyakoriak a szóbeli értékelésnél.
- Megjeleníthetők kiválasztott kategóriák *táblázatában* is, gyakorlatilag a különböző kategóriák párhuzamos skáláinak formájában. Ez a megközelítés lehetővé teszi a diagnosztikus profil létrehozását, azonban az értékelők csak korlátozott számú kategóriával tudnak dolgozni.

Két teljesen eltérő módon lehet az alskálák táblázatát megjeleníteni:

- *Készség szint-skálával*: profiltáblázat létrehozásával, amely meghatározza az egyes kategóriák megfelelő szintjeit, például az **A2** szinttől a **B2** szintig. A mérés ezután közvetlenül azokon a szinteken történik; ha nagyobb mértékű differenciálás szükséges, akkor lehetőség szerint további finomítások alkalmazásával, például egy második számjegy vagy pluszjelek segítségével. Tehát még akkor is lehetséges az erősebb tanulók esetében a **B1+**, **B1++**, vagy **B1.8** szint elismerése, ha a készség szint teszt a **B1** szintet célozta is meg, és még ha egyik tanuló sem érte el a **B2** szintet.
- *Vizsgáztatói pontozó skálával*: minden lényeges kategóriában olyan szintleírás kiválasztásával vagy meghatározásával, amely az adott modulnak vagy vizsgának megfelelő kategóriában leírja a megkívánt megfelelési szintet vagy normát. Ekkor a szintmeghatározás elnevezése „megfelelt” vagy „3”, és a skála ahhoz a követelményhez viszonyítva normaorientált (nagyon gyenge teljesítmény = „1”, kiváló teljesítmény = „5”). Az „1” és „5” megfogalmazása történhet más szintmeghatározások segítségével, amelyek lehetnek az 5. fejezet megfelelő szakaszából vett skála szomszédos szintjeiből származó és adaptált szintleírások, illetve a szintleírás meghatározható a „3” megfogalmazása alapján is.

9.2.3. A készségszintek leírása tesztekben és vizsgákon az összehasonlíthatóság érdekében

A Közös Referenciaszintek skáláinak célja a meglévő minősítésekben elért készségszint leírásának megkönnyítése, és ezáltal a különböző rendszerek összehasonlíthatóságának előmozdítása. A méréssel kapcsolatos szakirodalom a különböző mérések összekapcsolásának öt módját különbözteti meg: 1. megfeleltetés, 2. bemérés, 3. statisztikai moderálás, 4. szintmeghatározás és 5. társadalmi moderálás.

Az első három módszer hagyományos: 1. ugyanazon teszt alternatív változatainak létrehozása (megfeleltetés), 2. különböző tesztek eredményeinek egy közös skálához való viszonyítása (bemérés), és 3. a tesztek nehézségének és a vizsgáztatók szigorúságának ellensúlyozása (statisztikai moderálás).

Az utolsó két módszer egyetértést teremt megbeszélés segítségével (társadalmi moderálás), illetve standardizált definíciók és példák segítségével mintákat hasonlít össze (szintmeghatározás). A Referenciakeret egyik célja, hogy előmozdítsa ezt az egyetértést eléréséhez vezető folyamatot. Ezért szigorú kidolgozási módszerek alkalmazásával standardizáltuk az e célból használandó szintleíró skálákat. Az oktatásban ez a megközelítés egyre inkább standard-szemponturnak szerepel. Általánosan elismert, hogy egy standard-szemponturnak megközelítés kidolgozása időbe telik, mivel a felek a szemléltetés és a véleménycsere folyamata során értik meg a követelmények jelentését.

Feltételezhetjük, hogy ez a megközelítés a különböző mérések összekapcsolásának lehetséges legjobb módszere, mivel a felépítés (construct) közös értelmezésének a kialakításával és érvényesítésével jár. Alapvetően azért annyira nehéz a különböző nyelvi mérések összehasonlítása - a hagyományos módszerek statisztikai ördögössége ellenére -, mert a mérések, még ha az eredeti szándék szerint ugyanarra a mérendő területre vonatkoznak is, általában teljesen eltérő dolgokat tesztelnek. Ez egyrészt a) a felépítés (construct) nem teljesen megfelelő értelmezésének, valamint gyakorlatba való átültetésének, másrészt b) a tesztelési módszerek kihatásának a következménye.

Az európai kontextusban az élő nyelvek tanulására vonatkoztatva a Referenciakeret megkísérli elvi alapokon álló megoldást adni az első problémára, és arra, ami amögött meghúzódik. A 4-től a 7. fejezetig egy leírórendszert ismertetünk, amely megkísérli a nyelvhasználat, a kompetenciák valamint a tanítás és tanulás folyamatának gyakorlati megfogalmazását, amely segítséget nyújt a különböző feleknek, hogy az itt javasolt kommunikatív nyelvvismeretet át tudják ültetni a gyakorlatba.

A szintleíró skálák olyan fogalmi hálót alkotnak, amely felhasználható:

- a) a Közös Referenciakeret segítségével a nemzeti és intézményi keretprogramok egymáshoz való viszonyításához;
- b) a kategóriák és skálaszintek felhasználásával az adott vizsgák és oktatási modulok céljainak feltérképezéséhez.

Az „A” jelű függelék áttekintést nyújt az olvasóknak a szintleíró skálák kialakítási módszereiről, valamint azoknak a Referenciakeret skálájához való viszonyításáról.

Az ALTE által elkészített Használati útmutató vizsgáztatóknak [CC-Lang (96) 10. átdolgozott dokumentum] részletes útmutatást ad a felépítés (construct) tesztekbe való átültetéséhez és a tesztelési módszerek hatásaként fellépő szükségtelen torzulás elkerüléséhez.

9.3. Mérési és értékelési típusok

A méréssel és értékeléssel kapcsolatban számos fontos megkülönböztetést lehet tenni. A =. táblázatban bemutatott lista semmiképpen sem kimerítő részletességű. Semmilyen jelentőséggel nem bír az a tény, hogy a megkülönböztetésben egy elnevezés a jobb vagy a bal oldalra került.

9.3.1. Teljesítményt mérő értékelés / tudásszintet mérő értékelés

A *teljesítményt mérő értékelés* bizonyos célok elérésének értékelése, vagyis a tanított anyag elsajátításának értékelése. Következésképpen a heti/félévi munkára, a tankönyvre, a tanmenetre vonatkozik. A teljesítményt mérő értékelés az adott kurzust veszi figyelembe, és annak a belső szemszögét tükrözi.

A *tudásszintet mérő értékelés* viszont a tantárgy valós életben való alkalmazásával kapcsolatban értékeli a tudást/ismereteket, és külső szemszöget tükröz.

A tanárokat általában jobban érdekli a teljesítményt mérő értékelés, mert a tanítással kapcsolatban szeretnének visszajelzést kapni. A munkáltatókat, az oktatásügyi hatóságokat és a felnőtt tanulókat jobban érdekli a tudásszintet mérő értékelés: az eredmények, a szerzett képességek értékelése. A teljesítményt mérő megközelítés előnye, hogy közel áll a tanulók tapasztalatához. A tudásszintet mérő megközelítés előnye, hogy mindenkinek segít annak megítélésében, hogy hol tart; az eredmények átláthatóak.

Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

7. táblázat. Mérési és értékelési típusok

1	Teljesítményt mérő értékelés (Achievement assessment)	Tudásszintet mérő értékelés (Proficiency assessment)
2	Normaorientált értékelés (Norm-referencing, NR)	Kritériumorientált értékelés (Criterion-referencing, CR)
3	„Megfelelt”/”Nem felelt meg” kritériumorientált értékelés (Mastery learning CR)	Kontinuum viszonylatú kritériumorientált értékelés (Continuum CR)
4	Folyamatos értékelés (Continuous assessment)	Rögzített időpontú értékelés (Fixed assessment points)
5	Formáló értékelés ² (Formative assessment)	Összegző értékelés (Summative assessment)
6	Közvetlen értékelés (Direct assessment)	Közvetett értékelés (Indirect assessment)
7	Teljesítményalapú értékelés (Performance assessment)	Tudásalapú értékelés (Knowledge assessment)
8	Szubjektív értékelés (Subjective assessment)	Objektív értékelés (Objective assessment)
9	Listán alapuló mérés (Checklist rating)	Skálán alapuló mérés (Performance rating)
10	Benyomáson alapuló véleményalkotás (Impression)	Irányított véleményalkotás (Guided judgement)
11	Holisztikus értékelés (Holistic assessment)	Analitikus értékelés (Analytic assessment)
12	Sorozatalapú értékelés (Series assessment)	Kategóriákat felállító értékelés (Category assessment)
13	Mások által végzett értékelés (Assessment by others)	Önértékelés (Self-assessment)

A kommunikatív tesztelés során, szükségletszemponitú tanítási és tanulási kontextusban azt feltételezhetnénk, hogy ideális esetben kicsi a különbség a (kurzus tartalmára irányuló) teljesítmény és a (valós élet képességeire irányuló) készségszint között. Amennyiben a teljesítményt mérő értékelés megfelelő helyzetekben való gyakorlati nyelvhasználatot tesztel, és célja a kirajzolódó kompetenciáról alkotott kiegyensúlyozott kép létrehozása, valóban van szintmérő jellege. Amennyiben a tudásszintet mérő típusú értékelés olyan világos és megfelelő tanmeneten alapuló nyelvi és kommunikatív feladatokból áll, amelyek lehetőséget adnak a tanulónak arra, hogy megmutassa, mit ért el, valóban van teljesítményt mérő jellege.

A szemléltető szintleíró skálák a tudásszintet mérő értékeléstípusra vonatkoznak: a valós életben felhasználható képesség folytonosságát adják.

² A szakirodalomban a formatív értékelés és a szummatív értékelés kifejezések is használatosak.

A teljesítményt mérő tesztelésnek a tanulás megerősítésében játszott szerepét a 6. fejezetben vitattuk meg.

9.3.2. Normaorientált / kritériumorientált értékelés

A *normaorientált értékelés* a tanulók sorrendbe helyezése, vagyis társaikhoz képest való értékelése és rangsorolása.

A *kritériumorientált értékelés* a normaorientált értékelés elleni reakció, amikor a tanulót az adott tantárgyból társai képességétől függetlenül, teljesen saját képességei alapján értékelik.

Normaorientált tesztelés végezhető az osztály viszonyában („Te vagy a 18.”), vagy egy demográfiai csoport („Te vagy a 21 567., benne vagy a felső 14%-ban”) vagy a tesztet megíró csoport viszonylatában. Az utóbbi esetben a nyers tesztpontok kiigazíthatók a teszteredmények eloszlás görbéjének a tavalyihoz való hasonlításával a „tisztesseges” eredmény, a színvonal fenntartása és annak biztosítása érdekében, hogy a teszt nehézségétől és a tanulók képességeitől függetlenül minden évben ugyanannyi tanuló kapjon „ötöst”. A normaorientált értékelés gyakran használatos a tanulói csoportok kialakításához alkalmazott besorolási tesztekben.

A kritériumorientált értékelés magában foglalja a készség szint folyamatoságának (függőleges dimenzió) és a megfelelő tartományok körének (vízszintes dimenzió) feltérképezését annak érdekében, hogy az egy tesztben elért egyéni eredmények a teljes kritériumsorban elhelyezhetőek legyenek. Ez magában foglalja *a)* az adott tesztben/modulban szereplő megfelelő tartomány(ok) meghatározását, és *b)* a „határpontok” azonosítását: vagyis a tesztben a megállapított készség szint követelményeinek eléréséhez szükséges pontszámot.

A szemléltető szintleíró skálák a leírórendszer kategóriáihoz készített kritérium-meghatározásokból állnak. A Közös Referenciaszintek közös követelményrendszert nyújtanak.

9.3.3. „Megfelelt”/„Nem felelt meg” kritériumorientált értékelés / Kontinuum viszonylatú kritériumorientált értékelés

A *„Megfelelt”/„Nem felelt meg” kritériumorientált* megközelítés olyan megközelítés, amelyben egyetlen „minimum kompetenciakövetelményt” vagy „határpontot” állítanak fel, ami a tanulókat „megfelelt” és „nem felelt meg” csoportokra osztja, a cél elérésének milyenségét minősítő fokozatok nélkül.

A *kontinuum viszonylatú kritériumorientált* megközelítés olyan megközelítés, amelyben az egyén képességét az összes kérdéses terület megfelelő képességi szintjének egy meghatározott kontinuumához viszonyítják.

Valójában számos kritériumorientált megközelítés létezik, de legtöbbjük alapvetően a „megfelelt”/”nem felelt meg” vagy a „kontinuum” megközelítés alapján értelmezhető. Sok félreértéshez vezet a kritériumorientált értékelés téves, kizárólagos azonosítása a „megfelelt”/”nem felelt meg” megközelítéssel. A „megfelelt”/”nem felelt meg” megközelítés teljesítményorientált, a kurzus/modul tartalmára irányuló megközelítés. Kevesebb hangsúlyt fektet a modul (és így az annak során elért teljesítmény) nyelvismereti kontinuumon való elhelyezésére.

A „megfelelt”/”nem felelt meg” megközelítés alternatívája, ha minden egyes teszt eredményét a megfelelő nyelvismereti kontinuumhoz viszonyítjuk, általában osztályzatok sorával. Ebben a megközelítésben a kontinuum a „kritérium”, a külső valóság, ami biztosítja, hogy a teszteredmények jelentéssel bírjanak. Ehhez a külső kritériumhoz skálaelemzéssel (pl. Rasch-moddellel) viszonyíthatunk, hogy az összes teszt eredményét összehasonlíthassuk egymással, és az eredményeket közvetlenül, egy közös skálán helyezzük el.

A Referenciakeret használható mind a „megfelelt”/”nem felelt meg”, mind a kontinuum viszonylatú megközelítés során. A kontinuum viszonylatú megközelítés szintskálái összevethetők a Közös Referenciaszintekkel, és a „megfelelt”/”nem felelt meg” megközelítés szerint elérendő célkitűzés ábrázolható a Referenciakeret által nyújtott kategóriák és szintek szempontrendszerének táblázatában.

9.3.4. Folyamatos értékelés / rögzített időpontú értékelés

A *folyamatos értékelés* az órai teljesítmény, produktumok és projektek értékelése a tanár és lehetőleg a diák által a kurzus egésze alatt. A végső jegy tehát az egész kurzust/évet/szemesztert tükrözi.

Rögzített időpontú értékelésről akkor beszélünk, amikor egy bizonyos napon folyó vizsga vagy más értékelés alapján jegyek és döntések születnek – általában egy kurzus végén vagy a kezdete előtt. Nem számít, hogy mi történt korábban, az a fontos, amit a tanuló az adott pillanatban teljesíteni tud.

Ezt az értékelést gyakran a döntéshozatal kurzuson kívüli, egy adott időben végzett módjának tekintik. A folyamatos értékelés olyan értékelést jelent, amely a kurzus szerves részét képezi, és amely folyamatos gyarapodás révén járul hozzá a kurzus végén történő értékeléshez. A házi feladat osztályozásán és az alkalmanként vagy rendszeres jelleggel íratott rövid teljesítmény mérő teszteken kívül a tanulás előmozdítása érdekében a folyamatos értékelés a tanár és/vagy a tanulók által kitöltött listákat/táblázatokat is jelenthet, valamint speciális feladatsorok értékelése, formális tantárgyi értékelés, és/vagy a lehetőleg a készítés különböző szakaszaiból vett, és/vagy a kurzus különböző időpontjaiban írt munkákat tartalmazó portfólió létrehozása révén is történhet.

Mindkét megközelítésnek vannak előnyei és hátrányai. A rögzített időpontban történő értékelés biztosítja, hogy a tanulók még mindig meg tudnak olyan dolgokat csinálni, amelyek esetleg két évvel korábban szerepeltek a tanmenetben. Ugyanakkor egyeseknél vizsgatraumához vezethet, míg más típusú tanulókat előnyhöz juttathat. A folyamatos értékelés jobban figyelembe veszi a nyelvtanuló kreativitását és különböző erős pontjait, viszont nagyon erősen függ a tanár objektivitásától. Ha túlzásba viszik, a tanuló életét folyamatos, soha véget nem érő tesztté, a tanárét pedig bürokratikus rémálommá változtathatja.

Folyamatos értékelés során hasznosnak bizonyulhatnak a kommunikatív tevékenységekre vonatkozó, képességeket leíró meghatározások kritériumlistái (4. fejezet). A kompetencia aspektusainak szintleírásaihoz (5. fejezet) kialakított értékelő skálák felhasználhatók a rögzített időpontú értékelés során történő osztályozáshoz.

9.3.5. Formáló értékelés / összegző értékelés

A *formáló (formatív) értékelés* a tanulás mennyiségére, az erős és a gyenge pontokra vonatkozó folyamatos információgyűjtési folyamat, amelynek eredményét a tanár a kurzustervezés folyamatába visszaforgathatja, valamint a tanulók felé jelezheti. A formáló értékelés gyakran nagyon tág jelentésben használatos, és kérdőívek alapján, illetve konzultációk során szerzett, mennyiségileg nem kifejezhető információt is tartalmaz.

Az *összegző (szummatív) értékelés* a kurzus végén jegyben fejezi ki a haladást. Ez nem szükségszerűen szintező típusú értékelés. Tulajdonképpen az összegző értékelés gyakran normaorientált, rögzített időpontú teljesítményt mérő értékelés.

A formáló értékelés előnye, hogy a javuló tanulást tűzi ki célul, hátránya pedig a visszajelzés területén található. A visszajelzés csak akkor működik, ha az, aki kapja olyan helyzetben van, hogy a) *észreveszi*, vagyis figyel, motivált, és ismeri a kapott információ kifejezési módját, b) *felfogja*, azaz nem árasztják el túlzott mennyiségű információval, képes azokat rögzíteni, megszerezni és személyére vonatkoztatni, c) *értelmezi*, azaz elégséges előzetes ismerettel és tudatossággal rendelkezik ahhoz, hogy megértse a kérdéses pontot, és ne cselekedjen a fejlődést gátló módon, továbbá d) *integrálja* az információt, vagyis rendelkezik azzal az idővel, irányultsággal és azokkal a megfelelő forrásokkal, amelyek az új információ átgondolásához, integrálásához és észben tartásához szükségesek. Ez már egymagában is feltételezi az önirányítást, amihez szükség van az önirányítás fejlesztésére, az egyén tanulási folyamatának önellenőrzésére, valamint arra, hogy az egyén a visszajelzés alapján meg tudja tenni a szükséges lépéseket.

Az ilyen tanulási tréning vagy tudatosítás elnevezésére a *formáló értékelés* (*évaluation formative*) kifejezés használatos. Ehhez a tudatosításhoz számos különböző eljárást lehet alkalmazni. Alapvető elv azonban a szubjektív elképzelés (pl. a listán szereplő követelmények közül melyikről gondoljuk azt, hogy megfelelünk neki) és a valóság összevetése (pl. a lista követelményeit tükröző anyag tényleges meghallgatása alapján megállapítjuk, hogy valóban megfelelünk-e a követelménynek). A DIALANG ilyen módon viszonyítja az önértékelést a teszt eredményéhez. Egy másik fontos eljárás különböző feladatminták - általános példaminták vagy tanulói minták - megbeszélése és ennek segítségével a tanulók ösztönzése a minőségi szempontokkal kapcsolatos, személyre szabott metanyelv kialakítására. Ezután a tanulók ezt a metanyelvet felhasználhatják erős és gyenge pontjaik ellenőrzésére, és egy önállóan irányított „tanulási szerződés” megfogalmazására.

A legtöbb formáló vagy diagnosztikus értékelés a közelmúltban tanult vagy hamarosan sorra kerülő egyes nyelvi jelenségek vagy készségek nagyon részletezett szintjén működik. A diagnosztikus értékelés esetében az 5.2. fejezetben ismertetett tartalmak listája túl általános a gyakorlati felhasználáshoz, utána kell nézni az adott szint megfelelő részletes követelményeiben is (pl. „*Waystage 1990*”, „*Threshold Level 1990*”). Azonban a kompetencia különböző szintjeinek különböző aspektusait meghatározó szintleírásokat tartalmazó táblázatok (4. fejezet) hasznosnak bizonyulhatnak egy szóbeli értékelésről adott formáló visszajelzésnél.

A Közös Referenciaszintek látszólag az összegző értékelésnél bizonyulnak a leghasznosabbnak. Azonban, ahogy ezt a DIALANG kutatási program alátámasztja, még az összegző értékelés során kapott visszajelzésnek is lehet diagnosztikus, és így formáló jellege.

9.3.6. Közvetett értékelés / közvetlen értékelés

A *közvetlen értékelés* a vizsgázó által ténylegesen végrehajtott tevékenység értékelése. Például az értékelő megfigyeli, amikor egy kis csoport valamit megbeszél, összehasonlítja egy kritériumtáblázattal, a táblázatban található legmegfelelőbb kategóriákkal összeveti a teljesítményt, és így ad értékelést.

A *közvetett értékelés* viszont általában írásbeli tesztet használ, ami gyakran a mögöttes készségeket értékeli.

A közvetlen értékelés a beszédre, az írásra és az interakciós beszédértésre korlátozódik, hiszen receptív tevékenységet soha nem lehet közvetlenül megfigyelni. Az olvasás például csak közvetett módon értékelhető, ha megkérjük a tanulókat, hogy a megértést kipipálással, mondatok befejezésével, kérdések megválaszolásával stb. bizonyítsák. A nyelvi formák választékossága és helyes használata értékelhető közvetlenül, a kritériumoknak való megfeleltetéssel vagy közvetett módon, a teszt kérdéseire adott válaszok értelmezé-

sével és általánosításával. Az interjú a közvetlen teszt klasszikus formája, a közvetett teszt klasszikus formája pedig a szövegkiegészítés (cloze).

Az 5. fejezetben megadott, különböző szintű kompetenciák különböző aspektusait meghatározó szintleírások felhasználhatók a közvetlen tesztek értékelési kritériumainak kialakításához. A 4. fejezetben ismertetett paraméterek segítséget nyújthatnak a produktív készségeket mérő közvetlen tesztek, valamint a beszédértést és az olvasást mérő közvetett tesztek témáinak, szövegeinek és tesztfeladatainak kiválasztásához. Az 5. fejezetben megadott paraméterek ezen kívül segíthetnek a nyelvtudás közvetett tesztelésében szereplő főbb nyelvi kompetenciák meghatározásában, valamint abban, hogy mely főbb pragmatikai, szociolingvisztikai és nyelvi kompetenciákra összpontosítsanak a négy készséget mérő tesztkérdések megfogalmazásában.

9.3.7. Teljesítményalapú értékelés / tudásalapú értékelés

A *teljesítményalapú értékelés* során a tanuló egy közvetlen teszt révén beszédben vagy írásban nyelvi mintát nyújt.

A *tudásalapú értékelés* során a tanuló nyelvismeretének és nyelvhasználatának bizonyítása érdekében különböző típusú kérdéseket válaszol meg.

Sajnos soha nem lehetséges a kompetenciák közvetlen tesztelése. Mindig a teljesítmények sorával kell megelégedni, amelyekből a nyelvi készségszintre vonatkozó általánosító következtetést vonunk le. A készségszint alkalmazott kompetenciaként fogható fel. Tehát ebben az értelemben a tesztek csupán a teljesítményt értékelik, bár ebből lehet következtetni a teljesítmény mögött meghúzódó kompetenciákra.

Egy interjú típusú feladat során ugyanakkor nagyobb „teljesítményre” van szükség, mint mondatkiegészítéskor, de a szövegkiegészítés nagyobb „teljesítményt” igényel, mint egy feleletválasztós teszt. Ebben az értelemben a „teljesítmény” szó nyelvi produkciót jelent. De a „teljesítmény” korlátozottabb jelentésben szerepel a „teljesítményalapú értékelés” szóban. Itt a „teljesítmény” szó egy (viszonylag) autentikus és gyakran munkához vagy tanuláshoz kapcsolódó helyzetben nyújtott megfelelő teljesítményt jelent. A „teljesítményalapú értékelés” kifejezés kissé lazább értelmezésében a szóbeli értékelési eljárások teljesítményalapú tesztnek tekinthetők abban az értelemben, hogy általános következtetést vonunk le a nyelvtudásról a tanulási kontextusban és a tanulók szükségleteinek megfelelően lényegesnek tekintett diskurzusfajták alkalmazásával nyújtott teljesítmény alapján. Egyes tesztek kiegyensúlyozzák a teljesítményalapú értékelést a nyelv mint rendszer ismeretének értékelésével, míg mások ezt nem teszik.

Ez a megkülönböztetés nagyon hasonló a közvetlen és közvetett tesztek közöttihez. A Referenciakeretet hasonlóképpen lehet felhasználni. Az Európa Tanács különböző szintekhez kapcsolódó részletes követelményei („*Waystage*

1990", „Threshold Level 1990”, „Vantage Level”) ezen kívül megfelelő részleteket kínálnak a célnyelvi tudásról azokon a nyelveken, amelyek rendelkezésre áll ilyen.

9.3.8. Szubjektív értékelés / objektív értékelés

A *szubjektív értékelés* az értékelő véleményén alapul. Ez a vélemény általában a teljesítmény minőségét fejezi ki.

Az *objektív értékelés* az értékelés olyan fajtája, amelyből eltávolították a szubjektivitást. Ezen általában olyan közvetett tesztet értünk, amelyben a kérdésekre csupán egyetlen jó megoldás létezik, pl. a feleletválasztós tesztben.

A szubjektivitás/objektivitás kérdése azonban ennél jóval összetettebb.

A közvetett teszteket gyakran „objektív tesztnek” hívják, amikor az értékelő egy adott megoldókulcs alapján dönti el, hogy elfogadja vagy elutasítsa-e a választ, és azután az eredmény megadásához összeszámolja a helyes válaszokat. Egyes tesztfajták ezt a folyamatot még egy lépéssel továbbfejlesztik azáltal, hogy csak egy helyes válasz létezik egy adott kérdésre (pl. feleletválasztós teszt és C-teszt, amit ezért fejlesztettek ki a szövegkiegészítési feladatból [cloze]), és a javítási hibák kiküszöbölésére gyakran alkalmaznak gépi javítást. Valójában az „objektív”-nak nevezett tesztek objektivitása ilyen értelemben némiképpen túlzó, hiszen valaki eldöntötte az értékelés olyan eljárásokra való korlátozását, amelyek lehetővé teszik a teszthelyzet szorosabb ellenőrzését (ez maga is egy szubjektív döntés, amivel mások nem feltétlenül értenek egyet). Azután valaki megírta a teszt részletes követelményeit, és lehet, hogy valaki más írta meg a részletes követelmények egy bizonyos részét a gyakorlatban tükröző itemet. Végül valaki kiválasztotta ezt a bizonyos itemet a tesztben felhasználható összes lehetséges item közül. Mivel ezek a döntések mind tartalmaznak szubjektív elemeket, az ilyen tesztek megjelölésére talán megfelelőbb az objektív pontozású teszt elnevezés.

A közvetlen teljesítményértékelés során az érdemjegyek általában valamilyen vélemény alapján születnek. Ez azt jelenti, hogy a tanuló teljesítményével kapcsolatos döntést szubjektív módon hozzák meg, a megfelelő tényezők figyelembevételével, valamint irányelvekre, kritériumokra vagy tapasztalatra való hivatkozással. A szubjektív megközelítés előnye, hogy a nyelv és a kommunikáció nagyon összetett jelenségek, nem darabolhatók fel alkotóelemekre, és részeik összegénél nagyobbak. Nagyon gyakran nehéz megállapítani, hogy egy adott item pontosan mit tesztl. Ezért a teszt-itemeknek a kompetencia vagy teljesítmény adott aspektusainak való megfeleltetése sokkal kevésbé egyértelmű, mint gondolnánk.

Mégis az igazságosság érdekében minden értékelésnek a lehető legobjektívebbnek kell lennie. Amennyire csak lehetséges, csökkenteni kell a szemé-

lyes értékítéletek szerepét a tartalom kiválasztására és a teljesítmény minőségének értékelésére vonatkozó szubjektív döntésekben, különösen összegző értékelés esetén. Ennek az az oka, hogy a teszteredményeket gyakran használják kívülálló felek az értékelt személyek jövőjére vonatkozó döntések meghozatalára.

Az értékelés szubjektivitása csökkenthető, az érvényesség és a megbízhatóság következképpen növelhető, ha megtesszük a megfelelő lépéseket, melyek a következőkhöz hasonlóak lehetnek:

- az értékelés **tartalmára vonatkozó részletes követelmények** létrehozása, például az érintett kontextus közös referenciakeretére alapozva;
- **különböző vélemények együttes figyelembevétele** a tartalom kiválasztásánál és/vagy a teljesítmények értékelésénél;
- az értékelés elvégzésére vonatkozó **egyezményes eljárások** alkalmazása;
- **határozott javítókulcsok** készítése a közvetlen tesztekhez és **specifikus meghatározott kritériumokon** alapuló döntések meghozatala a közvetett tesztek esetében;
- **több bírálótól származó vélemény** és/vagy a **különböző tényezők súlyozásának** megkövetelése;
- az **értékelési irányelvekkel** kapcsolatos megfelelő **kiképzés** megszervezése;
- az értékelés minőségének (érvényesség, megbízhatóság) ellenőrzése az **értékelési adatok elemzésével**.

Ahogy azt a fejezet elején megtárgyaltuk, a vélemények szubjektivitásának csökkentése érdekében tett első lépés az értékelési folyamat minden szakaszában az adott felépítés közös értelmezésének, vagyis egy közös referenciakeretnek a létrehozása. A Referenciakeret ilyen alapot kíván nyújtani a közvetlen tesztelés **tartalmi követelményeihez** és a tesztek **specifikusan meghatározott kritériumai** kifejlesztéséhez.

9.3.9. Listán alapuló mérés / skálán alapuló mérés

Skálán alapuló mérés: annak megítélése, hogy egy személy egy több szintből és sávból álló skálán egy bizonyos szinten vagy sávban található-e.

Listán alapuló mérés: egy személy megítélése egy bizonyos szint vagy modul esetében megfelelőnek ítélt szempontokból álló lista alapján.

A skálán alapuló mérés esetében az értékelt személyt egy sávrendszerben helyezük el. A hangsúly a függőleges tengely menti megoszláson van: milyen magasan helyezkedik el a személy a skálán? A különböző sávok/szintek jelentését a szintezett leírásoknak világosan kell meghatározniuk. A különböző kategóriák esetében számos skála létezik, és ezek megjeleníthetők

táblázat formájában ugyanazon az oldalon vagy különböző oldalakon. Létezhet definíció minden sáv/szint vagy minden második esetében, vagy a sáv/szint tetején, alján és közepén.

Az alternatív megoldást a lista jelenti, amelyben a hangsúly egy lényeges terület elsajátításának kimutatásán fekszik, vagyis a megközelítés a vízszintes tengelyen való elhelyezkedésen alapszik: a modul tartalmának mekkora részét teljesítette sikeresen a tanuló? A lista pontok sorát tartalmazhatja kérdőívhez hasonló formában. Megjeleníthető ugyanakkor körkörös alakban vagy valamilyen más formában is. A válasz lehet igen/nem, de lehet differenciáltabb is, (lehetőleg megnevezett) lépések sorozatával (pl. 0-4), és az elnevezések értelmezését megadó definíciókkal.

Mivel a szemléltető szintleírások az adott szinteknek megfelelően kalibrált különálló kritériumokat tartalmaznak, felhasználhatók *mind* egy adott szintnek megfelelő lista készítéséhez, mint például a Nyelvtanulási Napló egyes változataiban, *mind* az összes megfelelő szintre vonatkozó értékelő skála vagy táblázat készítéséhez, mint azt a 3. fejezet 2. táblázatában láttuk az önértékelés és a 3. táblázatban a vizsgáztatói értékelés vonatkozásában.

9.3.10. *Benyomáson alapuló véleményalkotás / irányított véleményalkotás*

Benyomáson alapuló véleményalkotás: a tanuló órai teljesítménye alapján kialakított, teljes mértékben szubjektív vélemény, amely független egy adott mérés adott kritériumaitól.

Irányított véleményalkotás: olyan vélemény, amelynek megalkotásakor az értékelő szubjektivitását csökkenti, hogy a benyomásokat specifikus kritériumok alapján történő tudatos értékeléssel egészítik ki.

Itt a benyomás szón azt értjük, amikor egy tanár vagy tanuló teljesen az órai munkában, házi feladatban stb. nyújtott teljesítménnyel kapcsolatos tapasztalata alapján értékeli. A szubjektív értékelés számos formája, különösen folyamatos értékelés során, az adott személy hosszabb idő alatt történő - lehetőleg tudatos - megfigyelése során képződő vélemény vagy emlék révén kialakított benyomás alapján történő értékelést jelent. Nagyon sok iskola-rendszer ilyen alapon működik.

Az irányított véleményalkotás itt olyan helyzetet jelent, amelyben a benyomások egy értékelési megközelítés segítségével irányított, megfontolt véleménnyé alakulnak. Az ilyen megközelítés feltételez a) egy valamilyen eljárás alapján történő értékelési tevékenységet, és/vagy b) egy meghatározott kritériumrendszert, amely különbséget tesz az eltérő pontszámok és érdemjegyek között, és c) a normarendszer alkalmazására irányuló kiképzést. Az irányított megközelítés előnye az, hogy ha ilyen módon létrehozzák az érintett értékelők csoportjának közös referenciakeretét, nagymértékben megnövelhető a döntések következetessége. Ez különösen igaz abban az

esetben, ha a „határpontok” teljesítményminták formájában jelennek meg, és kötött kapcsolódási pontjuk van más rendszerekhez. Az ilyen irányítás fontosságát megerősíti az a tény is, hogy - amint azt több tudományágban folyó kutatások ismételten kimutatták - a gyakorlatlan döntések miatt az értékelők szigora majdnem akkora eltérést felelős, mint a tanulók tényleges képességei közötti eltérések, és így az eredmények szinte teljesen a véletlennek köszönhetőek.

A közös referenciaszintek leíró skálái felhasználhatók a fent b) pont alatt bemutatott meghatározott kritériumrendszer létrehozásához vagy a már meglévő kritériumok által képviselt követelményeknek a közös referenciaszinteknek való megfeleltetéséhez. A jövőben esetleg kidolgozhatók olyan teljesítményminták, amelyek jól tükrözik a közös referenciaszintek egyes szintjeit, hogy segítséget nyújtsanak a normarendszer alkalmazására irányuló kiképzéshez.

9.3.11. Holisztikus / analitikus értékelés

A *holisztikus értékelés* átfogó, összefoglaló (szintetizáló) értékelést jelent. Az értékelő intuitív módon súlyozza a különböző szempontokat.

Az *analitikus értékelés* eltérő szempontok különálló vizsgálatát jelenti.

Ezt a megkülönböztetést kétféleképpen tehetjük meg: a) a vizsgálandó dolog szempontjából, és b) abból a szempontból, hogyan jön létre egy sáv, érdemjegy vagy pontszám. A rendszerek gyakran kombinálják az egyik szinten végzett analitikus megközelítést egy másik szinten végzett holisztikus megközelítéssel.

- a) *Mit értékeljünk:* Egyes megközelítések olyan átfogó kategóriákat értékelnek, mint a „beszéd” vagy az „interakció”, és mindegyik mellé pontszámot vagy érdemjegyet rendelnek. Más analitikusabb megközelítések megkövetelik az értékelőtől, hogy a teljesítmény több különálló szempontját külön eredményekkel értékelje. Ismét más megközelítések az átfogó benyomás megjegyzését, majd annak különböző kategóriák alapján történő elemzését és végül egy átgondolt holisztikus értékelés kialakítását várják el az értékelőtől. Az analitikus megközelítés különálló kategóriáinak az az előnye, hogy az értékelőt alapos megfigyelésre ösztönzik. Olyan metanyelvet nyújtanak, amelyet az értékelők az egymás közötti egyeztetéskor, valamint a tanulónak adott visszajelzéskor használhatnak. Hátránya, a gyakorlat tükrében, hogy az értékelők nem tudják könnyen elválasztani a kategóriákat a holisztikus értékeléstől. Ezenkívül több mint négy vagy öt kategória kezelésekor kognitív túlterhelés lép fel.

b) *Az eredmény kiszámítása:* Egyes megközelítések a megfigyelt teljesítményt holisztikus módon kapcsolják az értékelési skála szintleírásaihoz, akár holisztikus a skála (egy átfogó skála), akár analitikus (3-6 kategória egy táblázatban). Az ilyen megközelítések nem járnak számolással. Az eredményeket egyetlen szám vagy több kategória pontszámát egymás mellé írva, egy „telefonszám” formájában fejezik ki. Más analitikusabb megközelítések során több különböző szempontra egyenként bizonyos osztályzat adandó, amelyeket azután össze kell adni, és az így kapott végső pontszámot érdemjeggyé kell alakítani. Erre a megközelítésre jellemző a kategóriák súlyozása, vagyis a különböző kategóriák nem ugyanannyi pontot érnek.

A 3. fejezet 2. és 3. táblázata példát ad az önértékelésben és vizsgáztatói értékelésben használt analitikus kritériumskálákra (azaz táblázatokra), amelyek *holisztikus* pontozás során használhatók (azaz össze kell kapcsolni a teljesítmény alapján levont következtetést a definíciókkal, és így kell döntést hozni).

9.3.12. Sorozat-alapú értékelés / Kategóriákat felállító értékelés

A *kategóriákat felállító értékelés* egyszeri értékelési feladatot jelent (amelynek különböző szakaszai lehetnek, hogy különböző diskurzusmintákat vételesszenek, ahogy ezt a 9.2.1. fejezetben bemutattuk), amelynek során a teljesítményt egy értékelési táblázat kategóriái alapján ítélik meg a 9.3.1.1. fejezetben kifejtett analitikus megközelítés szerint.

A *sorozatalapú értékelés* egy sor különálló értékelési feladatot jelent (gyakran szerepjátékot más tanulókkal vagy a tanárral), amelyeket egyetlen egyszerű holisztikus jeggyel értékelnek egy 0-3-ig, vagy 1-4-ig terjedő skálán.

A sorozatalapú értékelés arra a kategóriákat felállító értékelés során megfigyelhető tendenciára ad egyfajta megoldást, hogy a kategóriákat felállító értékelésben az egyik kategóriában elért eredmény befolyásolja a többi kategória eredményét. Alacsonyabb szinteken a feladat-végrehajtáson van a hangsúly; a cél egy olyan szempontlista kitöltése, ami a tényleges teljesítmények tanári/tanulói értékelése alapján, és nem egyszerű benyomás alapján tartalmazza, hogy mit tud elvégezni a tanuló. Magasabb szinteken a feladatokat lehet úgy tervezni, hogy a teljesítmény során a készségszint egyes aspektusait mutassák ki. Az eredmények profil formájában jelennek meg.

Az 5. fejezetben ismertetett, a nyelvi kompetencia különböző kategóriáira vonatkozó skálák és a kísérő szöveg a kategóriákat felállító értékelés kritériumainak kialakításához nyújt segítséget. Mivel az értékelők csupán kevés kategóriát tudnak használni, a folyamat során engedményeket kell tenni. A 4.4. fejezetben kifejtett kommunikatív tevékenységi típusok és az 5.2.3.2.

fejezetben felsorolt különböző típusú funkcionális kompetenciák segíthetnek a sorozatalapú értékelésre alkalmas feladatok azonosításában.

9.3.13. Mások által végzett értékelés / önértékelés

Mások által végzett értékelés: a tanár vagy vizsgáztató értékelése.

Önértékelés: a tanuló saját készség szintjéről alkotott véleménye.

A tanulók több, fent vázolt értékelési eljárásba is bevonhatók. A kutatások azt mutatják, hogy amennyiben nem „nagy a tét” (pl. bekerül-e a tanuló egy kurzusra), az önértékelés hatékony kiegészítője lehet a teszteknek és a tanári értékelésnek. Az önértékelés pontosságát növeli, a) ha az értékelés a készség szint követelményeit világosan meghatározó szintleírások alapján történik, és/vagy b) amikor az értékelés egy bizonyos tapasztalathoz kapcsolódik. Ez akár egy teszt is lehet. Valószínűleg akkor is pontosabb az önértékelés, ha a tanulókat erre felkészítik. Az ilyen strukturált önértékelés akár olyan mértékben is összhangban lehet a tanári értékeléssel és tesztekkel, ami megfelel a gyakran előforduló tanárok közötti összhangnak, a tesztek közötti összhangnak, valamint a tanárok értékelése és a tesztek közötti összhangnak (megegyező érvényesség).

Az önértékelés legfőbb haszna azonban motivációs és tudatosságnövelő eszközként való alkalmazásában rejlik: abban segíti a tanulókat, hogy értékeljék erős pontjaikat, felismerjék gyenge pontjaikat, és hatékonyabban tanuljanak.

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

- *a fent felsorolt értékelési formák közül*
- *melyek felelnek meg jobban a rendszerükben tanulók szükségleteinek,*
- *melyek megfelelőbbek és kivitelezhetőbbek rendszerük pedagógiai kultúrájában,*
- *melyek eredményesebbek a „visszahatás” alapján történő tanári fejlődés szempontjából;*
- *milyen egyensúly áll fenn rendszerükben a teljesítmény (iskolaszempon-
tú; tanulásszempon-
tú) és a készségszint (valósvilág-szempon-
tú; eredmény-
szempon-
tú) mérése között, illetve ezek hogyan egészülnek ki
más mérési módszerekkel, és milyen mértékben értékelik a nyelvi tu-
dáson kívül a kommunikatív teljesítményt is;*
- *a tanulás eredményét milyen mértékben értékelik meghatározott köve-
telmények és kritériumok alapján (kritériumorientált értékelés), és
milyen mértékben kapcsolódnak az érdemjegyek és értékelések a tanu-
ló csoportjához (normaorientált értékelés);*
- *a tanárok milyen mértékben:*
 - *ismerik a követelményeket (pl. közös szintleírások, teljesítményminták),*
 - *kapnak ösztönzést, hogy megismerkedjenek többféle mérési eljárással,*
 - *képzettek az eljárások és az eredmények értelmezése terén;*
- *milyen mértékben kívánatos és megvalósítható a kurzuson végzett
munka folyamatos értékelésének és a rögzített időpontú értékelés
integrált megközelítésének létrehozása a vonatkozó követelmények és
kritérium-meghatározások alapján;*
- *milyen mértékben kívánatos és megvalósítható a tanuló bevonása az
önértékelésbe a feladatok meghatározásai és a különböző készség-
szin-
tek aspektusai alapján, valamint - például sorozatalapú értékelés ese-
tén - azon szintleírások alkalmazása révén;*
- *a Referenciakeret által megadott részletes követelmények és skálák
kontextusukra vonatkozó jelentőségét, valamint azt, hogyan lehet azo-
kat kiegészíteni vagy tovább részletezni.*

Önértékelési és vizsgáztatói értékelő táblázatok láthatók a 3. fejezet 2. és 3. táblázatában. A kettő között tapasztalható legszembeszökőbb különbség - a szemmel látható megfogalmazások mellett, mint például „Tudok...” vagy „Tud...” - az, hogy míg a 2. táblázat a kommunikatív tevékenységekkel foglalkozik, a 3. táblázat a bármely szóbeli teljesítmény során megnyilvánuló kompetencia általános aspektusait ismerteti. Azonban könnyen elképzelhető a 3. táblázat kissé egyszerűsített változata önértékelési célokra. A gyakorlat

azt mutatja, hogy legalábbis a felnőtt tanulók képesek kompetenciáik ilyen minőségi szempontú értékelésének az elvégzésére.

9.4. Megvalósítható értékelés és metarendszer

A 4. és 5. fejezetben bemutatott skálák példát adnak a 4. és 5. fejezet szöveges részében ismertetett, átfogóbb leírórendszerhez kapcsolódó, de leegyszerűsített kategóriarendszerre. Nem az a szándékunk, hogy bárki is az összes szint valamennyi skáláját használja egy gyakorlati értékelési megközelítésben. Az értékelők nehezen birkóznak meg nagyszámú kategóriával, és ráadásul az ismertetett szintek teljes skálája nem feltétlenül helytálló az adott kontextusban. A skálák inkább referenciaeszközként használhatók.

Bármilyen megközelítést választunk is, mindenfajta gyakorlati értékelési rendszer esetében a lehetséges kategóriákat kezelhető számúra kell csökkenteni. A tapasztalat alapján megállapíthatjuk, hogy több mint 4 vagy 5 kategória kognitív túlterhelést eredményez, és pszichológiai szempontból 7 kategória a felső határ. Ebből következik, hogy szelektálni kell. Szóbeli értékeléskor, ha az interakciós stratégiákat a kommunikáció szóbeli értékelés szempontjából lényeges minőségi tényezőjének tekintjük, a szemléltető skálák a szóbeli értékelés szempontjából 12 fontos minőségi kategóriát tartalmaznak, melyek:

Szóátvételi stratégiák
Együttműködési stratégiák
Magyarozatkérés
Folyékonyság
Rugalmasság
Koherencia
Témakifejtés
Pontosság
Szociolingvisztikai kompetencia
Általános nyelvi készségek
Szókincs terjedelme
Grammatikai helyesség
Szókincs alkalmazása
Fonológiai tudás

Nyilvánvaló, hogy míg egy általános szempontlistában ezek közül több szintleírás is szerepelhetne, 12 kategória túl sok bármely teljesítmény értékelése esetén. Ezért a kategóriák ilyen listáit minden gyakorlati megközelítés esetében szelektív módon kell kezelni. A jellemző tulajdonságokat össze kell vonni, átnevezni, és az adott tanulók szükségleteinek, az adott értékelési feladat követelményeinek és az adott pedagógiai kultúrának megfelelő,

Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

kevesebb számú értékelési kritériumra csökkenteni. Az így keletkező kritériumokat lehet egyenlően súlyozni, vagy lehet egyes tényezőket az adott feladat esetében kiemelt fontosságúnak tekinteni, és nagyobb súlyt helyezni rájuk.

A következő négy példában szemléltetjük, hogyan lehetséges ez. Az első három példa rövid bemutatása annak, hogyan lehet a kategóriákat létező értékelési megközelítések tesztkritériumaiként használni. A negyedik példa bemutatja, hogyan történt a Referenciakeret skála leírásainak összevonása és átfogalmazása egy bizonyos alkalomra, egy adott célnak megfelelő értékelési táblázat létrehozása érdekében.

1. példa:

Cambridge Certificate in Advanced English (CAE), Paper 5: Értékelési kritériumok (1991)

Tesztkritériumok	Szemléltető skálák	Egyéb kategóriák
Folyékonyság	Folyékonyság	
Nyelvhelyesség és választékosság	Általános nyelvi készségek Szókincs terjedelme Grammatikai helyesség Szókincs alkalmazása	
Kiejtés	Fonológiai tudás	
Feladatteljesítés	Koherencia és kohézió Szociolingvisztikai megfelelés	A feladat sikeressége A beszélőpartner segítségének szükségessége
Interaktív kommunikáció	A szó átvétele Együttműködés Témakifejtés	A beszélgetéshez való hozzájárulás mértéke és könnyedsége

*Megjegyzés az egyéb kategóriákkal kapcsolatban: A szemléltető skálákban a feladat sikerességével kapcsolatos állítások a *Kommunikatív tevékenységek* címszó alatt szereplő tevékenység típusra vonatkoztatva találhatóak. Azokban a skálákban a *beszélgetéshez való hozzájárulás mértéke és könnyedsége* a *Folyékonység*hoz tartozik. A *beszélgetőpartner segítségének szükségességét* a skálarendszerben meghatározó szintleírások megírására és bemérésére tett kísérlet sikertelennek bizonyult.*

2. példa:

International Certificate Conference (ICC): Certificate in English for Business Purposes, 2. Teszt: Üzleti beszélgetés (1987)

Tesztkritériumok	Szemléltető skálák	Egyéb kategóriák
1. skála (elnevezés nélkül)	Szociolingvisztikai megfelelés Grammatikai helyesség Szókincs alkalmazása	A feladat sikeressége
2. skála (a diskurzus jegyeinek használata a beszélgetés kezdeményezésére és a beszélgetés menetének fenntartására)	A szó átvétele Együttműködés Szociolingvisztikai megfelelés	

3. példa:

Eurocentres - Kiscsoportos interakció értékelése (RADIO) (1987)

Tesztkritériumok	Szemléltető skálák	Egyéb kategóriák
Választékosság	Általános nyelvi készségek Szókincs terjedelme	
Nyelvhelyesség	Grammatikai helyesség Szókincs alkalmazása Szociolingvisztikai megfelelés	
Előadásmód	Folyékonyság Fonológiai tudás	
Interakció	A szó átvétele Együttműködés	

241

4. példa:

Svájci Nemzeti Kutatási Tanács: Videóra rögzített teljesítmények értékelése

Kontextus: A szemléltető szintleírásokat egy svájci kutatás során az „A” jelű függelékben ismertetett módon helyezték el a skálákban. A kutatás befejezésekor a részt vevő tanárokat meghívták egy konferenciára, hogy ismertessék az eredményeket, és elindítsák Svájcban az Európai Nyelvtanulási Nappalóval végzett kísérleteket. Ezen a konferencián a megvitattott témák közé tartozott a) a folyamatos értékelés és az önértékelés során használt szempontlisták egy átfogó referenciakerethez való viszonyításának szükségessége, és b) a projekt során skálákba rendezett szintleírások felhasználásának módjai az értékelésben. A vita folyamatának részeként a 3. fejezet 3. táblázatában bemutatott értékelési táblázat alapján értékelték néhány, a felmérésben részt vevő tanulóról készült videofelvételt. Ez a szemléltető szintleírások összevont, átírt válogatását mutatja be.

Tesztkritériumok	Szemléltető skálák	Egyéb kategóriák
Választékosság	Általános nyelvi készségek Szókincs terjedelme	
Nyelvhelyesség	Grammatikai helyesség Szókincs alkalmazása	
Folyékonyság	Folyékonyság	
Interakció	Általános interakció Szó átvétele Együttműködés	
Koherencia	Koherencia	

A különböző rendszerek különböző tanulók esetében, különböző kontextusban, eltérő módokon egyszerűsítik le, választják ki és kombinálják a jellemzőket a különböző értékelési típusokhoz. Valójában a 12 kategóriából álló

Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

lista nem hosszú, sőt valószínűleg nem képes az emberek által választott valamennyi variációs lehetőség bemutatására, ezért kiegészítésre van szükség ahhoz, hogy teljesen átfogó legyen.

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

- *milyen módon egyszerűsítik rendszerükben az elméleti kategóriákat alkalmazható megközelítéssé;*
- *a rendszerükben értékelési kritériumként használt fő tényezőket milyen mértékben lehet elhelyezni az 5. fejezetben ismertetett kategóriák rendszerében - amelyhez skálamintákat adtunk meg a Függelékben - és kontextusuknak megfelelően bővebben kifejteni úgy, hogy a nyelvhasználat specifikusabb tartományait is figyelembe vegyék.*