



Assessing Learning Achievement

– A tanulmányi eredmények változásának értékelése –

Az értékeléssel foglalkozó, az alábbiakban ismertetett publikáció előkészítésére az UNESCO, alapanyagként, 1991. évi világgjelentése számára az *International Association for Educational Assessment-et* (IAEA), Az oktatási eredmények értékelésének nemzetközi szervezetét (rövidítése nálunk: IEA) kérte fel.

Ennek az anyagnak a kidolgozásakor az értékelés elméletének és gyakorlatának három jellemző iránya volt. Az IEA tagjait felkérték, hogy példákkal illusztrálják, illetőleg egészítsék ki tartalmukat. Mindhárom fejlődési trend jellemzőit, illetőleg várható, jövőbeni alakulását megtárgyalták. A vitákon a következő fő kérdések merültek fel: Miért szükséges a tanulmányi eredmények értékelése? Miként értékelhetők a tanulmányi eredmények? Melyek az értékelés eljárásai? Melyek az értékelés jelenlegi és a jövőben várható trendjei? A viták a felsorolt kérdésekre megadták a választ. Ezek jelentés formájában való összegezése az UNESCO-publikáció számára igényes feladat volt, s annak teljesítésében komoly szerepet játszott az IEA-tagok közreműködési készsége. Az értékelés témában kiadott UNESCO-kiadvány anyagát az IEA nevében *John Izard*, a pedagógiai kutatás ausztráliai tanácsának tagja foglalta össze. Tartalma a szerző álláspontját fejezi ki, s nem képezi szükségképpen az UNESCO véleményét.

A kiadvány kitűnő áttekintést ad a témával foglalkozó kutatók széles körének a legfontosabb vitakérdésekben elfoglalt álláspontjáról, a publikáció előkészítésében részt vevő intézményekről országonként, és tartalmazza a hivatkozott művek bibliográfiai adatait.

Az UNESCO-publikáció ismertetését különösen aktuálissá teszi az a körülmény, hogy a pedagógiai értékelés fogalma változóban van. A hagyományos didaktika az értékelést az oktatási folyamat mozzanataként, tanítási módszerként tartotta számon. Fogalma a 60-as évektől világszerte kitágul. Növekszik szerepe a pedagógiai eljárások hatékonyságának megítélésében, a tanítási-tanulási folyamat tervezésében, irányításában. Hazánkban a „tanterv törvény” szemlélet meghaladása, a 80-as évektől a centralizált tanterv létjogosultságának és a központi irányítási rendszernek a megkérdőjelezése együtt járt az eredményre orientált irányítás és iskolafejlesztés térnyerésével. A pedagógiai értékelés, a tanulmányi eredmények alakulásának figyelemmel kísérése, a monitor típusú eredménykövető ciklikus tudásszintmérés mind nagyobb szerepet kap a tanterv értékelésében, a curriculum céltudatos, megalapozottabb fejlesztésében.¹

¹ A monitor típusú vizsgálatok rendszerét az Egyesült Államokban alakították ki. A National Assessment of Educational Progress, az 1968-ban megindult NAEP-program, az oktatásügyi eredmények fejlődésének az országos felmérése „azt tűzte ki célként, hogy a legfontosabb tantárgyi területeken azonos rendszerben, meghatározott időpontokban visszatérően, azaz ciklikusan végezzenek feltáró vizsgálatokat”. (Vajthó Erik: Monitor típusú vizsgálatok az Amerikai Egyesült Államokban – NAEP program. = Pedagógiai Szemle, 1985. 6. sz. 579. o.)

MIÉRT SZÜKSÉGES A TANULMÁNYI EREDMÉNYEK ÉRTÉKELÉSE?

A publikáció a tanulmányi eredmény fogalmát átfogóan, a tantárgyi tartalom elsajátításához szükséges tudás-, képesség- és készségrendszerként értelmezi, amelyben a tanuló saját véleményét felépíti, a tanultakat beilleszti a már megszerzett ismeretbázisának szélesebb összefüggésébe azzal a céllal, hogy ha szükséges van rá az iskolával, a munkával, a mindennapi élettel kapcsolatban várhatóan jelentkező problémák megoldásához, visszaidézhesse azt (*Palincsar, A. S.-Winn, J.*).

Az értékelésnek az a *funkciója*, hogy érvényes bizonyítékát adja a tanulmányi eredmények változásának,² s egyben informálja annak jellemzőiről a tanulókat, a tanárokat és a szülőket, s ezzel kedvező hatást gyakoroljon mindhárom, a tanulmányi munka eredményességét közvetlenül befolyásoló, abban leginkább érdekelt tényezők tevékenységére.

Az értékelés a *tanuló* számára tanúsítja, hogy elérte-e a megkövetelt tanulmányi szintet vagy nem, és befolyásolja további ismeretszerzésének folyamatát. Az eredménymérésnek a tanulmányi tevékenységre gyakorolt sokféle hatása *Crooks, T. J.* egy átfogó – 14 specifikus területre kiterjedő – vizsgálatának a tanúsága szerint eltérő az *osztály szintjén*, az *oktatási kurzus folyamán*, illetve *hosszabb időtávban*. – Az *osztályban* az értékelés a tanulók figyelmét ráirányítja a tantárgy fontos aspektusaira, ösztönzi az aktív tanulási stratégiák alkalmazását, korrekatív visszacsatolást jelent, amellyel segíti a tanulókat saját tudásuk elbírálásában: fejleszti önértékelési készségeiket. – A *kurzus során* befolyásolja az értékelt tantárgy tanulásának motivációját és az egyéni fejlődés lehetőségeinek a felismerését az adott tárgyban. Közvetíti a tanulók felé és megerősíti a helyi, valamint a nemzeti curriculum tágabb céljait, beleértve a megkövetelt standard teljesítményeket. Leírja és tanúsítja a diákoknak a kurzus folyamán elért fejlődését, és ezzel hatást gyakorol jövőbeni tanulmányi aktivitásukra. – *Hosszabb időtávra* is kiterjed az értékelés hatása: növekszik a gyakorlottság a tanult anyag megőrzésében, összefüggéseinek felismerésében és eltérő módokon való alkalmazásában. Befolyásolja továbbá a motiváció tartósságát nemcsak az adott tantárgyban, hanem általánosságban is.

Az értékelés jelentős hatása következtében érthető, hogy a tanulók gyakorta követelik: tájékoztassák őket a tananyagnak azokról a részleteiről, témáiról, amelyeket az értékelők – legyenek azok belső vagy külső személyek – leginkább elismernek. Ennek érdekében arra is vállalkoznak, hogy elemezzék a megelőző értékelések jellegzetességeit, ha azok anyaga: tesztjei, felmérő kérdései, feladatai elérhetőek számukra.

Az értékelés feltétlenül szükséges mozzanata a tanítási-tanulási folyamatnak. Elengedhetetlen a *tanárok* számára megelőző pedagógiai munkájuk szakmai mérlegelése és az oktatási folyamat további tervezése, vezetése, esetleg annak korrekciója szempontjából. A tanulmányi eredmény mérésének a tanárok mindennapi gyakorlati munkájára kifejtett hatása szintén jelentős. Befolyásolja azt az értékelt tananyag specifikálásával, leírásával, rámutat súlypontjaira, felhívja a figyelmet az elsajátításához szükséges készségekre. Hatást gyakorol az oktatómunkára azokkal az explicit és implicit eljárásokkal is, amelyekkel mérte az eredményeket,

² A tudásintenzitás és változása mérésének alapmodelljei a következők: a) kategóriaváltás a kétfokozatú nominálskálán: két különböző időpontban elvégzett mérés alkalmával a megoldás helytelenről helyesre vált; b) rangsori hely váltása ordinálskálán: például osztályelső pozícióból a második rangsori helyre kerülés; c) a mért érték változása intervallumskálán: például „A” tanuló és az osztályelső diák közötti, a helyes megoldások számában kifejezett távolságnak a változása. (*Wilfried Echterhoff: Lern- und Veränderungsmessung. In: Handbuch der Pädagogischen Diagnostik. Hrg.: Klausner, K. J. Band 1. Pädagogischer Verlag Schwamm, Düsseldorf, 1978. 157. o.*)

illetve végül, de nem utolsósorban azzal, hogy a felmérték vonatkozásában pontos képet ad arról, hogy az egyes gyerekek teljesítették-e, felül- vagy alulmúlták-e az irántuk támasztott követelményeket. A tanárok tehát növendékeik tudásának, teljesítőképeségének teljesebb ismeretében céltudatosabban fejleszthetik az egyes tanulókat.

Az értékelés kérdéseire adott tanulói válaszok a tanárok oktatómunkája szempontjából is informatívak, hiszen a jó felelethez szükséges készségek rangsorára, típusára lehet belőlük következtetni. További információs forrást is jelentenek: impliciten segítik azoknak a készségeknek a meghatározását, amelyeket az értékelők/az állami adminisztráció fontosaknak tart.

Abban az esetben, ha a tanárok nem is ismerik pontosan az értékelés kérdéseit, azok formája tekintetében nem bizonytalankodnak, hiszen az általában állandó. Ezért nem ritka, hogy a tanárok inkább a többnyire tesztekben szokványos kérdések megválaszolását gyakoroltatják a tanulókkal, mintsem a természetes összefüggésben jelentkező feladatok megoldását. Gyakorta adják például feladatul egyes, mintegy „szótárból kiemelt” szavak meghatározásának a memoralizálását tekintet nélkül arra, hogy azok teljes értelme a beszédösszefüggésben tárul fel. Ekképpen ahelyett, hogy növendékeiket rávezetnék a szó jelentésére, művi tanulást hoznak létre, amelynek jobbra az a szerepe, hogy a tanulók jól teljesítsenek az értékelés/vizsgáztatás alkalmával (*Windham, D.-Levin, H.-Báthory, Z.*). A tanárok oktatási gyakorlatát az egymást követő értékelések nagymértékben el is torzíthatják. Vannak közöttük, akik végül tesztmegoldásra tanítanak színvonalasabb oktatás helyett: a jó teszteredmények elérését fontosabbnak tartják a tanulmányi előrehaladásnál.

Természetes, hogy növendékeikhez hasonlóan, főként külső értékelés esetén a pedagógusokat is érdeklik a várható tudásszintmérés részletei. A hasonló témájú régebbi értékelések feladataira adott pontértékek, az egyes témák körében megadott kérdések száma alapján informálnak arról, hogy az értékelők/vizsgáztatók miként minősítenek, mit tartanak fontosnak, mit értékelnek magasra.

A kutatók figyelmeztetnek egy csapdára: az értékelés túlburjánzása, a jobb teszteredmények elérése érdekében reá fordított időtartam megnövelése rossz hatású: kevesebb idő jut a tanításra. Mivel azonban a tanításra fordított idő a tanulás lehetőségeként is kifejezhető, a tanítás kedvéért bekövetkezik az értékelés időtartamának kényszerű csökkentése, ami viszont káros gyakorlatra vezet: a felhasznált értékelési technikák a „ceruza-papír” tesztfeladatokra korlátozódnak, amelyek valamennyi tanulótól ugyanazt a választ várják el. Nem tartalmaznak többféle megoldásra lehetőséget kínáló, a problémamegoldást különböző irányokból megközelíthető feladatokat (*de Lange, J.-Verhage, H.; Izard J. F.*).

Mivel legkönnyebb a tények ismeretét (főként az elszigetelt tényekét) inkább, mint azok összefüggését vagy az elméletek, általánosítások, definíciók szóbeli kifejezését) tudakoló kérdések megválaszolása (és kidolgozása), az értékelés/vizsgák kérdései jelentős arányban inkább ismereteket tesztelő feladatokat tartalmaznak, mint a megértést, az ismeretek új helyzetben való alkalmazását követelőket. Ez még akkor is kedvezőtlen, ha a tényszerű felidézés a magasabb rendű készségek indirekt mértékű szolgálhat annak a feltételezésnek az alapján, hogy meglehetősen magas a korreláció a megkövetelt készségek teljesebb köre és az aktuálisan indirekt módon mérték között.

Az értékelés eredménye fontos információt jelent a szülő számára is, aki annak ismeretében gyermeke otthoni tanulására céltudatosabb befolyást gyakorolhat, s tudásának, képességeinek realisabb megítélése alapján hatékonyabban segítheti szakma- és pályaválasztását.

A tudásszintmérésnek az említettekén túl terjedő jelentősége is van. Tájékoztatja az *iskolai körzeteket*, amelyek az értékelés adatainak, azok gondos elemzésének, valamint a tanárok tanácsainak alapján továbbfejlesztik az adott iskolafokozat oktatási terveit. Az országos szintű reprezentatív mérések fontos *közoktatás-politikai célok*at szolgálnak szerzte a világban. A Koreai Köztársaságban például egy négyéves, 1986-tól 1989-ig végzett monitor vizsgálattal a tanulmányi eredmények változásait kívánták feltárni; Indiában pedig még a közszolgálati állások betöltése során is figyelembe veszik az értékelési eredményeket.

Bármilyen szűk vagy tág körben, bármiféle speciális céllal is történjen a tudásszintmérés, a tanulmányi előrehaladás értékelése, az *információkat szelektív módon kell gyűjteni*, akárha a tanulók által elért színvonalnak a kitűzött szinttel, a standarddal való összehasonlítására, akár annak tisztázására szolgál, hogy milyen oktatásra van szükség ahhoz, hogy hatékonyan segítse a tanulókat az iskolázás következő, magasabb szintre jutásában, vagy éppen azt várják tőle, hogy felhívja a figyelmet olyan hiányosságokra, az erőforrások várható szűkösségére, amelyek ismerete fontos lehet a nemzet érdeke szempontjából.

MIKÉNT ÉRTÉKELHETŐK A TANULMÁNYI EREDMÉNYEK?

A fenti kérdés megválaszolása nem egyszerű, mivel az értékelés alkalmával sokféle tényezőt kell szem előtt tartani. Tekintetbe kell venni a kapott információk felhasználásának módját, továbbá azokat a szakértő személyeket, akik felelősek a curriculum szándékait kifejező feladatok típusainak a kiválasztásáért, az információk összegyűjtéséért. Számba kell venni, milyen időbeli és anyagi kihatásai vannak a feladatok tanulókhoz való eljuttatásának, a feleletek pontozásának, valamint a tesztfeladatoknak a curriculum további részeire gyakorolt befolyását.

Az értékelés eltérő szempontú megközelítése

Az iskolázás céljáról, az értékesnek tartott ismeretekről és készségekről, valamint a tanulás preferált módjáról alkotott nézetek az értékelés különböző szempontú megközelítésére vezetnek. Eszerint az értékelés háromféle: *kompetitív, nem kompetitív és kooperatív* perspektívából valósítható meg (Suggett, D.).

A kompetitív szempontú értékelés

Az értékelésnek ebben a perspektívájában a versengés szelleme érvényesül. Az értékelés célja egyfelől annak megállapítása, hogy a tanuló a standardot elérte-e, másfelől, hogy lehetővé váljék a különbségtétel a tanulók között, létrejöjjön a szelekció bázisa. Ennek megfelelően az értékelési technikák a tanulókat minősítik, fokozatokba sorolják, kategorizálják. Az értékelés a sorrendjük és a megoldásukhoz szükséges idő szempontjából ellenőrzött feladatokra épül. A tanárnak az a szerepe, hogy közvetítse a tantárgyi ismeretet, megállapítsa, hogy ki tanult növendékei közül, hogy különbséget tegyen a tanulók között képességeik szerint, és megfelelően differenciáljon közöttük. Ebben a perspektívában a tanulóknak az a szerepe, hogy készüljön fel az értékelés kérdéseire, „adja vissza” azt, amit neki megtanítottak, az értékelés során önállóan dolgozzék és reménykedjék a sikerben. Az értékelők feltételezik, hogy a tanulás a diáktól megkövetel bizonyos képességeket, elvárják tőle, hogy tanúsítson erőfeszítést iskolai munkája során. És fordítva! Feltételezik, hogy a tanulmányi kudarc a képességek nem kielégítő voltából és/vagy a tanulás iránti alacsony motivációs szintből fakad. Az egyes tanulók

képességeinek eltéréseit adottságaik különbözőségéből eredőnek tekintik, és a jobb adottságúakat a kevésbé rátermettek elé sorolják be.

A nem kompetitív szemléletű értékelés

Ilyen megközelítésben az értékelés fő célja a tanulók erőfeszítéseinek jutalmazása, bátorításuk a további tanulásra, a diákok egyéni elbírálása. Az értékelési technikák a feladatokhoz/kérdésekhez írásos és szóbeli megjegyzéseket/magyarázásokat fűznek, és elkerülik a tanulók összehasonlításának olyan önkényes eljárásait, mint a kategóriákba sorolás. Az értékelésnek ebben a változatában a tanárnak az a szerepe, hogy kijelölje a tanulók szükségleteinek és érdeklődésének megfelelő, az intellektuális és szociális fejlődésüket előmozdító feladatokat/kérdéseket, hogy lehetővé tegye számukra a közülük való választást, továbbá, hogy rokonszenvvel viszonyuljon a tanulók munkájához, és ösztönözze egyéni fejlődésüket. A tanuló szerepe a részvétel, az értékelés elfogadása és az, hogy tanuljon a hozzáfűzött magyarázatokból. Az értékelők feltételezik, hogy a tanulás megköveteli a diák érdeklődését és motiváltságát; a kudarc pedig a „próbálkozás” hiányából, inkább az eltérő szükségletekből és érdeklődésből, mintsem a képességbeli különbségekből ered. Az értékelt tanulók csoportosítása nem a képességek alapján történik.

A kooperatív szemléletű értékelés

Ebben a szemléletben az értékelés fő célja az értelmes tanulás ösztönzése, valamennyi tanuló előrehaladásának az előmozdítása és egyesek megjutalmazása. Az értékelési technikát jellemzi, hogy a diákok munkája elé már megelőzően célokat tűz ki, amelyek elérését a kielégítő befejezés meghatározó alapjának tekinti. Az értékelés a tanulók végzett munkájáról leíró megállapításokat fogalmaz meg. Ebben az értékelési szemléletben a tanár szerepe az, hogy olyan feladatokat adjon, amelyek a tanulók számára értékesek, lényegesek, továbbá, hogy bevonja őket a tanulásukért, annak eredményesebbé tételéért érzett felelősség vállalásába. A tanárok hajlanak arra, hogy az adott kurzus folyamán végzett oktatómunkát, illetve minden egyes tanuló teljesítményének ismételt értékelését két diszkrét folyamatnak tekintsék. Arra törekszenek, hogy elérhető követelmények állításával valamennyi diákjukat képessé tegyék a sikeres előrehaladásra. A tanulók szerepe a részvétel, illetve további tanulmányaik megtervezésének megbeszélése a tanárral, a kemény munkálkodás valamennyi felmérő feladat megoldása érdekében és a törekvés tanulmányi tevékenységük folyamatos fejlesztésére.

A kooperatív szemléletű értékelés a diákoktól azt várja továbbá, hogy gondolkodjanak el a tanulatokon, és segítsék az együttműködés szellemének kialakítását az osztályban. A különböző képességű tanulókból álló heterogén osztályokban az elért teljesítményszintek nem determinálják a diákok csoportosítását. A képességeket szorosan az adott kurzus minőségéhez, a feladatok összefüggéséhez és természetéhez, valamint a tanuló motivációjához kapcsolódónak tekintik.

Az értékelés típusai

Az ausztráliai iskolákban végzett értékelések tapasztalata alapján *Withers* és *Batten* arra a meggyőződésre jutottak, hogy szükség van az *értékelés kulcstípusai*-nak a megállapítására. Ezek a következők: a *leíró értékelés*, a *kritériumorientált értékelés*, a *standardra vonatkoztatott értékelés*, valamint a *normatív értékelés*.

A leíró értékelés informálja a tanulót és szüleit a tanulmányi szakasz folyamán azokról az aspektusokról, amelyek a diák tanulmányi munkáját értelmesebb teszik minőségének emelésével vagy új irányba terelésével. A tananyagegység befejezését követően pedig tájékoztatja őket a végzett munka során tapasztalt tanulói szándékokról, a munka tartalmáról és minőségéről.

Suggett, D.-nek a leíró értékeléshez kidolgozott irányelvei mellett érvelnek, hogy a diákok tevékenysége az értékelés folyamán legyen változatos formájában (szóbeli, írásbeli vagy multimediális), típusában (leíró, kifejező, elemző vagy elmélkedő), terjedelmében (csoportos vagy egyéni). Az értékelésről készített beszámolót szintén jellemezza a változatoság (például írásbeli tanári megjegyzés/magyarázat, önértékelés, tanuló-tanár interjú, tanulói napló a végzett munkáról, társ értékelése). A legfontosabb az, hogy az információ pontosan, korrekten, igazságosan, adekvát módon tükrözze a diák tanulmányi tevékenységének minőségét.

A beszámolóknak a következő hét elemet kell tartalmaznia: a tanulói munka és az értékelés összefüggését; a megkövetelt munka típusának világos részletezését és fő hangsúlyát; a befejezett munka pontos leírását mindazon technikai részletekkel együtt, amelyek értékelik a diákokat; annak a felfedhető összefüggésnek a részletezését, amely miatt nem készült el a munka; a tanulmányi előrehaladás bizonyítékát (beleértve a váratlan fejlődést); amennyiben szükséges, útmutatást a tanulónak arra, miként fejlődhetne tovább.

A beszámolóknak el kell kerülniük a következő kilenc csapdát: a képességek, attitűdök, a személyiség globális megítélését; a megfigyelt teljesítmény okaira vonatkozó feltevéseket; a határozatlan általánosításokat; az értékelési kliséket, a szakmai zsargonjt; a szélsőséges megállapításokat, amelyek minden bizonnyal fékezik a tanulók erőfeszítéseit; a sértő vagy destruktív kritikát; a tanár érzelmait kifejező megjegyzéseket a végzett munkáról vagy a tanulóról; a nemi, faji vagy egyéb diszkriminációt kifejező előítéleteket; a tanuló összehasonlítását másokkal.

A leíró értékelést nem használják fel a tanulók rangsorolására. *Withers, G.* és *Batten, M.* azt a tudósítás egyik formájának tekintik.

A kritériumorientált értékelés

Az értékelés céljának egy, az előzőtől eltérő felfogását képviseli a kritériumorientált értékelés,³ amely tudósítja a tanulót azokról a meghatározott célokról/kritériumokról, amelyeket elért (vagy nem ért el). Az értékelésnek ez a típusa specifikált feladatok kombinációja azzal a céllal, hogy bizonyítsa és a kritériumra vonatkoztatottan határozza meg az elsajátítást vagy annak hiányát, illetőleg megállapítsa a teljesítmény kategorizálásának a szintjeit. Amennyiben a tanuló megfelel a kritériumoknak, akkor elismerik róla, hogy rendelkezik az értékelt és az ahhoz hasonló feladatok teljesítéséhez szükséges képességekkel. Mivel a kritériumorientált értékelés esetében nem korlátozott a sikeresen teljesítők száma, ez jelenti az oktatási programok hatékonyság szerinti összehasonlításának legobjektívabb típusát. Az a

³ A kritériumorientált értékelésre a klasszikus, a normaorientált tesztelmélet hiányosságai vezettek. A kritériumorientált metodika a vizsgált tanulók teljesítményét az adott populáció teljesítményétől, az „átlagtól” független vonatkoztatási rendszerben állapítja meg. Előírt szintekhez (például tantervi követelményekhez), a lehetőséges maximumhoz, többnyire azonban operacionálisan feladatokhoz, kritériumokhoz viszonyítja a tanulók teljesítményét. Az eljárás eredményeként megállapítja, hogy a tanuló(k) az elvárt szintet képviselő feladatokból mennyit oldottak meg, hogy a mért ismeretlem, tudás, tulajdonság, képesség milyen mértékben alakult ki. A kritériumorientált tesztelés jellemző funkciója a minősítés, a kategorizálás (l. *Csapó Benő*: A kritériumorientált értékelés. =Magyar Pedagógia, 1987. 3. sz. 247–266. o.; *Báthory Zoltán*: Tanulók, iskolák – különbségek. Tankönyvkiadó, Bp., 1992. 235–236 o.).

program jobb a másiknál, amelyiknél az oktatott csoport tanulói nagyobb arányban sikeresek a mérés alkalmával. Ha a kritériumorientált értékelés mérészközeiben a direkt feladatok sora megoldásuk nehézsége szerint rendezett, akkor a pedagógusok számára az ilyen mérések eredményeként kapott pontszámoknak az interpretálása a legkönnyebb.

A standardra vonatkoztatott értéklés

Amennyiben a kritériumorientált értékelést külső standardok rendezett sorozatához vagy teljesítményfokokozatokhoz kötött deskriptorokhoz (minőségi leírásokhoz) viszonyítva értelmezik, azt standardra vonatkoztatott értékelésnek tekintik.⁴ A standardok vonatkozhatnak ismeretekre, folyamatokra, érzelki-mozgásos és emocionális tárgyakra. Például Nyugat-Ausztrália iskoláinak 8. osztályaiban a matematika tantárggyal kapcsolatban végzett eredményvizsgálat során az „általánosítás” folyamatkategóriában a Clark, D. J. által kidolgozott, következő deskriptort – ötfokozatú minőségi skálát – alkalmazták a tanulók teljesítményeinek értéklésére.

- *Kiváló teljesítmény:* képes felismerni és folytatni egy összefüggés-sorozatot; következtetéseket levonni táblázatokból, grafikonokból; le tud írni megadott formában egy összefüggést.
- *Magas teljesítmény:* képes felismerni és kiterjeszteni egy összefüggés-sorozatot; következtetéseket levonni táblázatokból, grafikonokból; gyakorolta le tud írni egy összefüggést.
- *Megfelelő teljesítmény:* képes felismerni, kiterjeszteni és megmagyarázni egy összefüggés-sorozatot; következtetéseket levonni egyszerű táblázatokból és grafikonokból.
- *Korlátozott teljesítmény:* képes megadott feltételeknek megfelelően általánosítani egyes esetekből; többnyire fel tud ismerni és folytatni egy egyszerű összefüggés-sorozatot.
- *Nem megfelelő teljesítmény:* nem képes sem egyes példák alapján megadott feltételeknek megfelelően általánosítani; sem felismerni, sem folytatni egy összefüggés-sorozatot.

A tanároknak, akik ilyen minőségi skálákat használnak, még szükségük van megfelelő értéklő feladatokra is, amelyek lehetőséget adnak a listába foglalt magatartások megnyilvánulására.

A normatív értéklés

Egyes tanulók eredményeinek más tanulók teljesítményére vonatkoztatott interpretációja az értéklésről való beszámolás jól járható útjának bizonyult.⁵

Az értéklés ezen típusának alkalmazásakor az egyén teljesítményét vagy egy referenciacsoportéhoz, vagy az egyén valamely osztálytársáéhoz viszonyítják. A normatív értéklés népszerű azon országok nevelési rendszereiben, ahol az iskolázásra fordítható pénzügyi források igen korlátozottak, és emiatt a tanulmányi eredmények értéklése azok kiválasztását szolgálja, akik alkalmasak arra, hogy eljussanak az iskolázás következő fokára. Az ilyen döntő ponton történő tanulmányi eredménymérés alkalmával inkább a jól begyakorolt készségeket értéklék, mint a curriculumban foglalt készségek szélesebb körét. Ha ez kombinálódik a vizsgákon azzal a tendenciával, amely tények és verbalizált általánosítások visszaidezését célozza, akkor az ilyen iskolarendszerekben tanuló diákok kettős hátrány-

⁴ A standardizált tesztek megfelelően kiválasztott, statisztikailag reprezentatív minta alapján dolgozzák ki. Fraise: „A teszt standardizált formájában a mérésnek olyan eszköze, amely az egyén teljesítményének színvonalát azon csoport normájához ... viszonyítva állapítja meg, amelyhez az illető tartozik.” A gyakorlati pedagógusok szempontjából a standardizált tantárgytesztek a leginkább figyelemre méltók. Ilyeneket alkalmaztak a Magyarországon is elvégzett IEA-felmérésben, s ilyeneket dolgozott ki és adott közre például Nagy József (A témazáró tesztek validitása és reliabilitása. Standardizált témazáró tesztek. JATE, Szeged, 1976).

⁵ A normaorientált értéklés a klasszikus pszichometria hagyományos eljárásaként a kiválasztott minta adataihoz, többnyire átlagához viszonyítja az egyes tanulók aktuálisan mért teljesítményét, tulajdonságait, jellemzőit. A tantárgyi feladatlapon, valamint a standardizált tesztek lényegében a normaorientált pedagógiai értéklés eszközei.

től szenvednek. A sikeresek esetében a kapott oktatás elhanyagolja a számukra szükséges problémamegoldó készségek kialakítását, és behatárolja a továbbtanuláshoz szükséges kapacitásukat is. A nem sikeresekre gyakorolt negatív hatás még komolyabb hátrányt jelent, hiszen számukra az iskolai tanítás nem igazán releváns jövődől munkájuk szempontjából, és megfoszthatja őket még a továbbtanulás lehetőségétől is.

A fő gondot nem annyira a normatív értékelés használata, mint inkább a felületes, indirekt mérési feladatok alkalmazása okozza a reális, direkt értékelés helyett. A jobb, a curriculumban megkövetelt készségek egész sorát lefedő vizsgafeladatok alapját jelenthetnék a szelekciónak, noha a tanárok az osztályban végbemenő tanítási-tanulási folyamat fejlesztése szempontjából hasznosabbnak tartják az iskolai keretekben megszervezett formatív értékelést.

MELYEK AZ ÉRTÉKELÉS ELJÁRÁSAI?

A tanulmányi eredménymérés eljárásai az értékelésbe bevontak száma szerint országonként jelentékeny mértékben változnak. Például Kínában 1990-ben a felsőfokú oktatásba történő felvételi vizsgát 2,78 millió jelölt tett, Ausztrália New South Wales tartományában 54 000 hatodik osztályos tanulóit tesztelték olvasásból, anyanyelvből, matematikából, ugyanakkor az értékelési folyamat egyetlen osztályra is kiterjedhet.

A direkt értékelés

Komplex, realiztikus feladatok bevonása növeli az értékelési feladatok érvényességét, mivel teljesebben reprezentálják a curriculum céljait. Az a trend, hogy szubordinált egyes összetevők diszkrét értékelése helyett inkább alkalmaznak komplex, direkt feladatokat, a pontozás folyamán nehézségekre vezet. Mutatkoznak például az írásbeli fogalmazásoknak az értékelésénél, de jelentkeznek projekteknek, portfólióknak (írásbeli tanulói munkák, rajzok, gyűjtemények) és gyakorlati, mindenkor összetett feladatoknak az értékelésekor is.

Sokféle megközelítés alkalmazható esszék, *írásbeli fogalmazások* értékelésénél. Huot, B. áttekinti az írásbeli fogalmazások értékelési eljárásainak főbb trendjeit. Megállapítja, hogy a fogalmazások minőségében elért fejlődés értékelésének alternatív stratégiai közül az indirekt értékelés sok gondot okoz. Így például egyes nyelvi készségek indirekt mérése azzal a pedagógiai problémával jár, hogy a tanárok – magasabb tesztpontszámok elérése érdekében – hajlamosak tanítványaikat grammatikára, a mondatok struktúrájára és használatára oktatni. Ezzel eltorzítják azt az értékelők által feltételezett összefüggést, amely a mért készségek és a fogalmazások minőségében elért fejlődés között fennáll.

Huot az írásbeli teljesítmény direkt értékelése három fő eljárásának a következőket tekinti:

A fő vonások pontozása a fogalmazás egy vagy több, a sajátos retorikai helyzettel összefüggő releváns jellemzőjéhez kötődik. Ha például a feladat azt kívánja a tanulótól, hogy valamely fogalomhoz hozzákapcsolja saját tapasztalatait, akkor a megfelelő vonás az, hogy a tanuló mennyire képes a specifikus részletektől eljutni az illusztrálandó általános fogalomhoz.

Az analitikus pontozás a „jó” írás sokféle minősége alapján értékeli. A jó írásnak ezek a minőségei meghatározhatók. Egy írásbeli fogalmazás minősége azon az alapon ítéltető meg, hogy a jó írásnak hány jellemzőjét tartalmazza.

A holisztikus pontozás az értékelőnek az adott írásról alkotott általános benyomását tükrözi vissza. A legtöbb holisztikus értékelés alkalmával azonban a pontozáshoz kidolgozott irányelvek

részletezik, hogy mely általános jellemzők reprezentálják az írás minőségét a felhasznált, többnyire négy- vagy hatfokú skála egyes pontjain.

Ausztráliában jelentős kezdeményezések történtek a tanulók szóbeli és írásbeli teljesítményei tanári értékelésének a tökéletesítésére. Továbbképzést szerveztek az iskolázás alsó fokán oktató tanárok számára annak érdekében, hogy fejlesszék bennük a kisebb tanulók hangos olvasása értékelésének a készségeit. Gyakoroltatták az egyes gyerekek hangos olvasását rögzítő magnófelvételek készítését és interpretálását. Ausztrália északi területén kifejlesztették az 5 és a 7 évesek írásbeli munkái megítélésének eljárásait. A gyermekmunkákból mintákat gyűjtöttek össze és megküldték azokat értékelésre különböző iskolákba. Majd egy úgynevezett csökkentő bizottság minden írásfajtából (riport, elbeszélés, érvelés, magyarázat) reprezentatív mintákat választott ki, azokat kategorizálta – jó, megfelelő, gyenge –, magyarázatokkal látta el, és megküldte az iskolákba, példaképpen, az érintett pedagógusok értékelőmunkájának a megkönnyítésére. (*Northern Territory Department of Education*)

Annak az USA-beli erőfeszítésnek a jegyében, hogy fejlesszék a magasabb rendű gondolkodás oktatását és értékelését, a *National Assessment of Educational Progress* (NAEP), a nevelés-oktatás fejlődésének értékelésére alakult nemzeti bizottság egy kézikönyvben több gondolkodási szintre adott közre mintafeladatokat. A hierarchia első fokán a tanulóknak az a feladatuk, hogy *osztályozzanak és csoportosítsanak* növényeket és állatokat azok közös vonásai alapján. A következő fokon anyagokat, felszerelést és/vagy készülékeket kapnak a diákok, hogy sajátos matematikai vagy tudományos jelenségeket, illetve összefüggéseket *mutassanak ki, figyeljenek meg, vonjanak le belőlük következtetéseket* és fogalmazzanak meg *hipotéziseket*. A feladatok másik sorozata annak felmérését szolgálja, hogy a tanulók mennyire képesek *adatokban a rendszer felismerésére és az eredmények értelmezésére*. A legkomplexebb feladati szinten az a tennivaló, hogy a tanulók *tervezzenek meg és vezessenek teljes kísérleteket*. (*Educational Testing Service*)

A vizuális információk értékelése

Az a törekvés, hogy az értékelési feladatok tükrözzék vissza a curriculum szándékait, az értékelés egy második eljárására vezetett, amely komplexebb, vizuális vagy grafikus ábrázolásokat bevonó feladatoknak az értékelésére vállalkozik.

Gaulin, C. és Puchalska, E. felfigyeltek arra, hogy alapvető fontosságú a térbeli információk grafikus úton (rajzok, diagramok, térképek, tervek, grafikák, sémák stb.) történő közvetítése. Az adatoknak képi formában való felülvizsgálata ugyanis hasznosul az ember legmagasabb rendű információfeldolgozási kapacitásának (például ráismerés, osztályozás, vizuális mintákra való visszaemlékezés) a fejlődésében.

Spence, I. és Lewandowsky, S. rámutattak: az ember nagyon alkalmas arra, hogy felismerjen és feldolgozzon vizuális benyomásokat. Az agy jelentős területeinek ugyanis éppen a vizuális információk feldolgozása a rendeltetése, és kevesen vitatnák azt az állítást, hogy a látás a domináns érzékelési módozat. Ha az adatokat vizuális úton mutatják be valakinek, akkor annak alapján gyakran vesz észre, fog fel olyan finomságokat is, amelyek észrevétlenek maradtak számára táblázatos bemutatás esetén. *Lowe R. K.* megjegyzi, hogy a kézikönyveknek azért van nagyobb hatásuk a vizuális úton való tanulásra, mivel nagyszámú képet és diagramot tartalmaznak. Bár olyan világban élünk, amelyben dominálnak a vizuális médiák, a kevésbé gyakorlott szem egyeseket közülük könnyebben, másokat nehezebben képes befogadni. Ezért fontos lenne gyakoroltatni a tanulókkal a

vizuális benyomások befogadását. Annál inkább szükséges ez, mert sajátos, lényeges készségek fejlesztésére szintén alkalmas. A vizuális befogadásban tapasztalt, illetve tapasztalatlan egyetemi hallgatók vizsgálata azt tanúsította, hogy például a diagramok feldolgozása bizonyos kifinomultságot és sajátos készségeket követel. Olyanokat, amelyek hasonlóak a szövegfeldolgozáshoz szükségesekhez.

Értékelés teljesítménymutatók és minősítő skálák alkalmazásával

Az értékelésnek ez a harmadik eljárástípusa a megfigyelésből eredő információkra épül. Az elért teljesítmény mérésére teljesítménymutatókat és minősítő skálákat használnak.

A teljesítménymutatók

Fitz-Gibbon, C. T. és társukutatói behatóan foglalkoztak a teljesítménymutatóknak (teljesítményindikátoroknak) nevelési rendszerek monitor vizsgálataiban való szisztematikus alkalmazásával.⁶ *Boulter*nak az az álláspontja, hogy az egyes iskolák nem hasonlíthatók össze a nyers vizsgaeredmények alapján, tekintettel a beiskolázási körzeteknek, a társadalmi-gazdasági tényezőknél, a szülői attitűdöknek, a szelekciós eljárásoknak, valamint a pedagógiai programok támogatottságának az eltéréseire. *Gamble, R.* megkülönböztet alacsony és magas szintű teljesítményindikátorokat. Alacsony szintű mutatóknak számíthatnak például az iskolalátogatás aránya, azoknak a csoportoknak a százalékos aránya, amelyek a vizsgáztatás speciális standardját teljesítették. A magas szintű indikátorok részben az alacsony szintűekre épülhetnek, de szakmai interpretációt követelnek. Ilyen indikátorok irányulhatnak például az oktatás vezetésére vagy a pedagógusok szakmai fejlődésére. Szép példái a teljesítménymutatókat alkalmazó értékelésnek *Brandsma, J.–Ten Brummelhuis, A.–Plomp, T.* három curriculuminnovációs programot összehasonlító elemzése, valamint *Franke–Wikberg, S.* áttekintése arról, miként használják fel a tudásszintmérés útján nyert információkat az iskolázás második és harmadik fokán végbemenő oktatás minőségének az értékelésére.

A minősítő skálák

Az oktatásügyi adminisztráció feladata a vizsgáztatási adatoknak, a teljesítményellenőrzésnek folyamatos kontrollálása. Ők a teljesítményellenőrzés értékelésére a vizsgáztatási/értékelési adatok egész sorozatához férhetnek hozzá. Azonban az iskolázás korai szakaszában, valamint a felnőttek írás-olvasás értékelésének a területén kevés adatuk van, vagy egyáltalában nincs. Ilyen körülmények között minősítő skálákat alkalmaznak, annak ellenére, hogy használhatóságuk bizonyos mértékben behatárolt (a minta nem reprezentatív, hiányzanak az érvényességet ellenőrző adatok).

Kenny, D. E. és *Chekaluk, E.* leírták az iskolában várhatóan tanulási nehézségekkel küzdő gyerekek korai felismerését célzó – a pedagógusok számára készített

⁶ Magyarországon az országos körű monitor típusú ciklikus felmérések rendszerét az Országos Pedagógiai Intézet 1984-ben létesített Értékelési Központja indította meg, és ugyanő folytatta az Országos Közoktatási Intézet Értékelési Központjaként (OKI-ÉK), valamint bonyolítja le ma is. Az OPI-ÉK a felmérések célját 1985-ben a következőképpen fogalmazta meg: a) adjon valóságos képet a tanulók tudásáról, az oktatás hatékonyságáról, b) dolgozzon ki megfelelő értékelési stratégiát, c) a folyamatos fejlesztés programjának megfelelően, a kimenetelzabályozás előkészítésére gondolja újra az oktatás „hozájárál” kapcsolatos elképzeléseket. (*Báthory Zoltán* és mtsai: Monitor típusú felmérések a közoktatás rendszerében. [Tertv tanulmány] Az Értékelési Központ Kiadványai 2. Országos Pedagógiai Intézet, EK, 1985. [Kézirat])

– minősítő skála kialakításának a folyamatát. Mintájukban voltak mind normális ütemben fejlődő kisgyerekek, mind olyanok, akikről feltételezhető volt, hogy az iskolában az elvártnál alacsonyabb szinten fognak teljesíteni. A gyerekeket nyelvi, olvasási és emlékeztetvizsgáló tesztek egész sorozatával tesztelték, és a kapott eredményeket összevetették a tanárok által adott minősítésekkel. A tanulmány azt sejteti, hogy a sikeres olvasáshoz szükséges készségek változni látszanak az iskolázás első három évében: amikor a gyerekek túljutottak az iskolázás korai szakaszán mind a tanárok, mind a tesztek jobb előrejelzőivé válnak az olvasás területén bekövetkező fejlődésnek.

Griffin, P. és társai (Jewell, R.–Forwood, A.–Francis, R.) leírják az olvasási, írási, matematikai alkalmasság indikátorai alapján a felnőttek alapképességeit minősítő skálák kidolgozásának folyamatát. Ez az *International Literacy Year* (a mi szóhasználatunkkal kifejezve: a 16 évesnél idősebb fiatalok és felnőttek írás-olvasás-matematika tudását szorgalmazó „nemzetközi év”) ausztráliai programja által támogatott kutatás kiterjedt az említettek szempontjából különböző alkalmassági szinten lévőkhöz azonosítására és a kompetenciájuk színvonalának felismeréséért segítő leírások kidolgozására.

A kutatás során az egyes skálák számára kidolgozták az indikátorok egy piramisát. A talpán olyan magatartások foglalnak helyet, amelyekkel szinte mindenki rendelkezik, a piramis tetején pedig olyanok, amelyekkel csak kevesen. Fontos az indikátorok közötti implikációs viszony: egymással összefüggő sort alkotnak. A piramis tetején lévő magatartásnak magában kell foglalnia a piramis lejjebb lévő szintjein minden bizonnyal meglévő magatartásokat. A többi szintnek azonban nem kell tartalmaznia a felettük lévőkhöz megjelenő magatartásokat. A piramis lehetővé teszi egy sor kohéziósan összefüggő magatartás felismerését, azonosítását, amely esetleg a minőségi skálák egy sorozatát fogja eredményezni, amelyekből viszont egyéni profilok alakíthatók ki. Ennélfogva az alapképességek fejlődésének fajtáit és szintjeit megfigyelhető viselkedésekkel le lehet majd írni (Griffin, P. és társai).

Standardok alkalmazása

Zieky, M. megjegyzi, hogy a standard értékelés módszere nem veszi figyelembe az értékelendő feladatok komplex voltát. Ámbár a módszer nem tökéletes, mégis jól szolgálja egyfelől az értékelők figyelmének azokra a tényezőkre összpontosítását, amelyeket a tudásszintmérés során tekintetbe kell venni, másfelől sok vizsgálati személy tanulmányi előrehaladásának szintjére vonatkozó értékelést gyűjt össze.

Az eljárás a következőket követeli: megfelelő értékelők kiválasztását, olyan vizsgálati személynek a meghatározását, akinek az ismeretei és készségei a „mastery” és a „non mastery”⁷ (megfelelt – nem felelt meg) tanulók csoportjának a határvonalán helyezkedik el, az értékelők képzését a módszer használatában, valamint a „megfelelés pontszámának” meghatározását.

Berk, R. A. áttekinti a standardok megállapításának, illetve a hibaszámárányok elemzésén és osztályozásán alapuló standardok 38 eljárását. Véleménye szerint a végzett kutatómunka és a témában megjelent számos publikáció ellenére is rendszertelen és ellentmondásos a standardok megállapításának a folyamata.

⁷ A „mastery – non mastery” kategorizálás a „mastery learning”, az „elsajátításhoz vezető tanulás” fogalmával kapcsolatos, amely azt célozza, hogy a tanulók 75–90%-a váljék képessé olyan teljesítményre, amelyre a hagyományos osztályrendszerű tanítás körülményei között a tanulók legskorábban 25%-a képes. (Bövebben l. Csapó Benő: A mastery learning elmélete és gyakorlata = Magyar Pedagógia, 1978. I. sz.; Nagy József: A megtanítás stratégiája. = Köznevelés, 1981; Uő.: A megtanítás stratégiája. Tankönyvkiadó, Bp., 1984.)

Megpróbál tehát kutatás helyett standardizálni, a meglévő kutatási eredményekhez hozzákapcsol dokumentált tapasztalati anyagot, s megkísérli mindezt a gyakorlat szakemberei számára könnyen olvasható formába önteni.

Berk tanulmánya a standardizálással kapcsolatban hat technikai kritériumot állapít meg. A módszer nyújtson a csoportosításhoz megfelelő információt. Állapítson meg átlagpontoszámot, amely dichotomikus (kétpólusú) klasszifikációt (például: kiváló–nem kiváló, megfelelő–nem megfelelő, kompetens–nem kompetens, esélyes–esélytelen) tesz lehetővé. Noha ez művi dichotómia, amelyet a folyamatok teszteredmény-eloszlás fölé helyeznek, mégis fontos több pedagógiai és a szakmaválasztással kapcsolatos döntés szempontjából. – A módszer legyen érzékeny a tesztek teljesítménye iránt. A kritériumorientált tesztekkel mért készségek ugyanis különböző összetettségűek és bonyolultságúak, ezért az értékelési eljárásnak érzékenynek kell lennie a teszt elemei, azaz az itemek nehézségi foka és más lényeges jellemzői, mint a megfelel–nem felel meg közötti megkülönböztetés iránt. – A módszernek továbbá szentítvének kell lennie aziránt, hogy milyen oktatásban, képzésben részesültek a tanulók: figyelembe kell vennie a nekik nyújtott aktuális tanítást. – A módszernek statisztikailag megbízhatónak kell lennie: az ítéletek összegezésének és a tesztben nyújtott teljesítmény leírásának statisztikai eszközei közül az adott mérés céljának megfelelő, korrekt statisztikai eljárásokat kell alkalmaznia. – A módszernek ki kell fejeznie a standard és a pontskálán mért eredmény között mutatkozó különbséget. – A módszernek érvényes ítéleteket kell hoznia, tekintettel arra, hogy standardra alapozott döntésekről van szó.

MELYEK AZ ÉRTÉKELÉS JELENLEGI ÉS A JÖVŐBEN VÁRHATÓ TRENDJEI?

A tanulmányi eredmény változásának, a tudásszintnek a mérése csak egy pillantást jelent, meghatározott időpontban, az adott helyzetre. Ez a pillantás azonban elveszti érvényét, amint újabb tanulásra és a tudás akkumulációjára kerül sor.

Meddig érvényes a mérési információ?

Időszerűtlenné válik-e a mérési eredmény, ha újabb tanulás követi, noha az alkalmasság és a szakértelem fejlődése többet foglal magában a tudás egyes, diszkrét elemeinél. *Glaser, R.* rámutat: a gyakorlottabbak integrálják tudásuk elemeit, közös jellemzőket azonosítanak a megismert folyamatokban, és ellenőrzik saját tanulási stratégiáik sikerességét. Míg eleinte a tanuló/hallgató elszigetelt meghatározásokat tud, és azok megértése felületes, a hozzáértés, a szakértelem növekedésével az információ strukturálódik és integrálódik a korábbi szerkezetekkel, úgy, hogy az elemek, itemek visszakereshetők az emlékezet átfogóbb egységeiből. A legfejlettebb tanulók/hallgatók felismerik az alapelveket, mintákat és összefüggéseket is. Felülvizsgálják, vajon tudásuk releváns-e, tanulási stratégiáik megfelelőek-e.

Ha a tanulás csupán a tudás diszkrét, különálló elemeinek az elsajátítására vezet, akkor gyorsan elavul, az elemek elfelejtődnek, vagy helyüket más, összefüggéstelen itemek foglalják el. És ellenkezőleg: amennyiben az új ismeretek integrálódnak a megelőzően szerettekkel, akkor valószínűbb, hogy az értékelés érvényes marad és összegezi a tanuló fejlődését.

Az értékelés jelenlegi trendjei

Már széles körben ismert az a tekintélyes, nemegyszer kedvezőtlen hatás, amelyet az értékelés gyakorol az oktatásra. Ezzel kapcsolatban világossá vált az is, hogy az értékelés nem ad teljes képet egy-egy fiatal tanulmányi fejlődéséről. Gondot jelent továbbá az a tény is, hogy a szokványos, a százalékos arányt kifejező pontszámok értéke erősen függ a méréseknél alkalmazott tesztkérdések nehézségétől. Mindennek következtében a tudásszintmérés elméletében és gyakorlatában új fejlődési irányok jelentkeztek.

Új stratégiák a ceruza-papír mérési technikán belül és azon túl

Az értékelési technikák kiterjesztése abban a törekvésben gyökerezett, hogy feladataik az addigalatt teljesebben fedjék le, reprezentálják a curriculum követelményeit. Kibontakozott az a trend, amely a feladatoknak olyan struktúráját kívánta létrehozni, amely megfelelő alapot jelenthet a curriculumban megcélzott változatos magatartások meglétének, minőségének az értékelésére. Napjainkban az értékelési stratégiák kiterjesztésére és alternatív eljárásainak a kidolgozására egyaránt találunk példákat. Lássunk közülük néhányat mindkét csoportból!

Hongkongban a középfokot befejező 100 000 tanulóra kiterjedő értékelés a vizsgáztatottak nagy száma ellenére magában foglalt a szóbeli tesztek, a hallott szöveg megértését követelő tesztek mellett projekttevékenységet és gyakorlati feladatokat tartalmazó teszteket is. A technikai tantárgyakban a projekttevékenység mintegy 30 műhelymunkaórát vett igénybe két tanév alatt. A tanulók értékelésre bocsátották projektjük termékét, és beszámolóit készítettek előállításának folyamatáról, beleértve az alternatív tervekkel kapcsolatos kutatást és elemzést, a tervezés menetét és a munkarajzokat. A haladó szint jelöltjei a biológia, fizika és kémia tantárgyakban gyakorlati vizsgát tettek, amely megfigyelést, tervezést, analízist, beszámolót tartalmazott, és manipulatív készségeket követelt.

Érdekes *Szingapúr* példája, ahol mind országos, mind iskolai szintű értékelést szerveznek. Mindkettő tartalmaz folyamatorientált kérdéseket. A tanulók az iskolázás első fokán, 12+ éves korban leteszik az első fokú iskolázást befejező országos vizsgát (Primary School Leaving Examination), és iskolai szinten is részt vesznek értékelésben (tesztelés, gyakorlati munka és projekttevékenység formájában). Ezek eredménye döntő jelentőségű a különböző tanulmányi kurzusok irányában való továbbhaladásnak, a továbblépés megtagadásának, a szelekciónak.

Ausztráliai kutató, *Neuman, M. A.* állapította meg, hogy az értékelési stratégiák átléphetik a tantárgyi határokat. Ő, függetlenül attól, hogy melyik tantárgyról van szó, a hibák következő öt típusát különbözteti meg.

– *Olvadási hibák:* a hibázás akkor minősül ilyennek, ha a tanuló a problémaleírás egyik kulcsszavának vagy valamely szimbólumnak csak olyan szintű elolvasására volt képes, amely nem volt elégséges a problémamegoldás valamely helyes útján való előrehaladáshoz.

– *Megértésbeli hibák:* a tanuló a kérdésben foglalt valamennyi szót el tudta olvasni, de nem volt képes a szavak teljes értelmének a felfogására, ezért képtelen volt előrejutni a problémamegoldás valamelyik megfelelő útján.

– *Transzformációs hibák:* a tanuló megértette, hogy a feladat minek a kiderítését követeli tőle, de nem volt képes meghatározni azt a műveletet vagy a műveleteknek azt a sorozatát, amely szükséges a probléma megoldásához.

– *Műveletkészségek-beli hibák:* a tanulónak sikerült egy, a problémamegoldáshoz szükséges megfelelő műveletet vagy a műveletek egymásutánját meghatározni, de a munkafolyamatban nem volt képes a műveletek pontos kivitelezésére.

– *Dekodolási hibák:* A tanuló helyesen dolgozta ki a problémamegoldást, de nem tudta megoldását írásos formában elfogadhatóan kifejezni.

Az értékelés területén igen fontos a *hollandiai CITO*-nak, az országos mérési intézetnek a kutatói-kutatásszervezői tevékenysége (munkakapcsolatban vannak például a Magyar Közoktatási Intézet Értékelési Központjával is). Három kutatójuk: *Alberts, R. V. J., van Beuzekom, P. J. és de Roo, I.* nagy hívei annak, hogy az értékelés tartalmazzon gyakorlati feladatokat. Ők a gyakorlati munkafeladatok bevonását az értékelésbe az érvek következő három kategóriájával támasztották alá.

Segítséget adnak a tantárgyi ismeretek optimális megtanulásához: a változatos munkafeladatok folyamán a tanulók az elméleti anyaggal gyakorlati természetű kapcsolatba kerülnek, s ez az elméleti ismeretek elsajátításának erős motivációs bázisa lehet.

– *Kísérletet jelentenek a nevelés személyiségformáló céljai teljesüléséhez:* jó alkalmak arra, hogy a diákok megtanulják az együttműködést a társakkal a közös munkában, átéljék a gyakorlati problémamegoldást kísérő sokféle érzelmet, fejlődjön helyzetfelismerési- és kezdeményező-készségük, fantáziájuk.

– *A gyakorlati készségek kialakításának lehetőségei:* módot adnak a gyakorlati kutatáshoz szükséges, kisebb arányban kognitív jellegű, nagyobb részükben praktikus, manipulatív készségek gyakorlására.

Az *Amerikai Egyesült Államok* tesztelési szolgálata számos projektjében a magasabb rendű gondolkodási készségek értékelését célozta meg. Az iskolázás negyedik-nyolcadik fokának diákjait bevonták a magasabb rendű gondolkodási készségeket fejlesztő projektekbe. Az értékelést követően, a tudásszint meghatározására, a tanulók részt vettek gyakorló kutatók által kijelölt tevékenységekben, így az adatgyűjtésben, az adatokra építve előrejelzések megfogalmazásában és az eredményeket ellenőrző elvek meghatározásában.

A különböző értékelések egyesítése a tanulók teljes értékelésében

A különböző értékelések adatainak egyesítésével lehetővé válik átfogó, részletezett kép megrajzolása egy-egy tanuló tudásáról. Ennek a fejlődési trendnek az a törekvés a mozgatója, hogy váljék az értékelési folyamat építőbbé, informatívabb jellegűvé és nyíltá. Az ilyen értékelés a fejlődést, a megszerzett tapasztalatokat pozitívan ítéli meg, ismeri el és dokumentálja. Minden egyes tanulót bevon a formatív értékelés és a fejlődésről való beszámolás folyamatába. Ebbe a törekvésbe belejátszik az a feltételezés is, hogy a gondosan kiválasztott tantárgyakban és kurzusokon elsajátított elméleti ismeretek terén elért sikerek értékelése minden bizonnyal elvonja a figyelmet a tudásról, a megértésről, azokról a készségekről és minőségekről, amelyek elengedhetetlen összetevői a személyiség egészére irányuló nevelési folyamatnak. Ezek magukban foglalják a csoportmunka készségeit, a kezdeményező-készség fejlesztését, a tanulástért vállalt nagyobb felelősséget és a sajátlagos tanulási készségek fejlesztését. Illusztrálja a következőkben egy, a matematikatanításban alkalmazott *ausztráliai* példa (*Joffe, L. S.*), hogy milyen mértékben lehetséges megfelelő feladatok segítségével a megkövetelt általános készségeket – a tantárgyi határokat túllépve – értékelni!

A feladat általános megközelítése a tanulók részéről:

Miként fogalmazzák meg maguknak a problémát?

Honnan indulnak ki?

A feladat mely vonásait tekintik lényegeseknek?

A tanulók módszerei:

Használják-e különböző időpontokban rendszeresen a trial és/vagy error megközelítést?

Megfelelő-e ennek a használata?

Milyen tervezői tevékenység mutatkozik, és egyáltalában mutatkozik-e?

Milyen a terv kidolgozása?

Mi történik, ha a terv nem válik be?

Milyen változtatásokat hajt azon végre a tanuló?

A feladat befejezését követő visszajelzés:

Saját megelégedésükre és/vagy a tanár megelégedésére oldották-e meg a tanulók a problémát?

Teljesült-e a feladat minden követelménye?

Képesek-e a tanulók beszámolni arról, amit csináltak?

A tanulók írásbeli terve jó visszautákróződése-e szóbeli beszámolójuknak?

Felmerült-e a helyesbítés szükségessége?

Van-e szükség javításra?

A hagyományos vizsgák és országos szintű tudásszintmérések mellett az iskolákra épített értékelések egyfelől az információk szélesebb körének az összegyűjtését célozzák, másfelől helyet adnak annak a gyakorlatnak, amely egy hosszabb időszakra terjeszti ki az értékelés folyamatát. A szakértők feltételezik, hogy az iskolai szintű értékelések – *a változatosabb, sokrétűbb értékelési feladatok* alapján – pontosabb képet adnak a tanulók fejlődéséről, és inkább elismerik a direkt értékelést.

Hoge, R. D. és Colardaci, T. áttekintették a tanári értékelések szakirodalmát. Arra a következtetésre jutottak, hogy nagyobb erőfeszítéseket kellene tenni annak érdekében, hogy a tanárok jobban megértsék az értékelés fontosságát az oktatási folyamat szempontjából, hogy több tapasztalatot szerezzenek az értékelés alapelveiről, hogy megbarátkozzanak különböző eszközeivel és módszereivel, beleértve az új megközelítéseket, valamint megfogalmazták azt a követelést, hogy a pedagógusok ez irányú képzési programjainak a megszervezését ki kell terjeszteni.

Szerte a világban egyre nagyobb jelentőséget tulajdonítanak az iskolai szintű értékeléseknek. Néhány példa erre.

Több helyütt, így Botswanában, Hongkongban és Skóciában az iskolai tudásszintmérés beépül az egyes tanulók teljes értékelésébe.

Botswana a tanároktól valamennyi tanulóra vonatkozó folyamatos értékelési feljegyzéseket vár el. Az új, kilencéves oktatás curriculumának bevezetését követően az addiginál több tantárgy folyamatos értékelésének az eredményét vonták be a befejező értékelésbe. Így például a mezőgazdaság tárgyban elért eredmények 40%-os, a háztartástanban elérték 20%-os súllyal szerepeltek a befejező értékelésben. A művészetek tárgyban a tanuló által egy év (vagy még hosszabb idő) alatt öt művészeti ágban készített, portfólióba gyűjtött munkáinak több kritérium szerinti minősítését kombinálják a teszteredményekkel. (Ez utóbbiaknak kisebb a fontosságuk!)

A *hongkongi* vizsgáztató bizottság előkészíti azt a gyakorlatot, amely szerint a haladó tanulók szintjén a kémia tantárgy körébe tartozó készségeket – a gyakorlati vizsga alternatívájaként – a pedagógus értékeli folyamatosan, legalább nyolc alkalommal, egy kétéves perióduson keresztül.

Skóciában a kötelező iskolázás végén, 16+ éves korban megtartott vizsgán az értékelésnek és a bizonyítványnak három tipikus eleme van. A természettudományokban például a tudás-megértés, a problémamegoldás és a gyakorlati készségek. Az egyes elemeket külön-külön értékelik és egyenlő a súlyuk.

Az „item response theory” alkalmazása

Hagyományos értékelés esetén, amely az egyes tanulók által elért pontszámokat százalékarányban fejezi ki, nehézségekkel jár az eredmények interpretálása, mivel az értékek a tesztkérdések könnyű vagy nehéz volta szerint módosulnak. További

problémát okoz, hogy a tudásszint-alakulásnak pontszámok százalékarányával való kifejezése értelemszerűen magában foglalja azonos vagy paralel tesztek használatát. S bár egy oktatási program keretében korábban alkalmazott teszt visszatükrözheti az egy osztályon belül bekövetkező fejlődés rangsorát, nem valószínű, hogy egy vagy két évvel később ugyanazok a tesztkérdések megfelelőek a fejlődés mérésére. Amire szükség van, az egy olyan stratégia, amely lehetővé teszi a fejlődés összegezését és az egyik teszten mért teljesítménynek a másik tartalmú teszten elért teljesítményhez való viszonyítását. Az „item response theory” (IRT) modellek lehetőséget nyitnak az összegzésre, valamint a különböző – eltérő tartalmú, de bemért, egymással azonos nehézségű itemekből felépített – teszteken elért teljesítmények összehasonlítására. Ezek a modellek széleskörűen elterjedtek a tanulmányi fejlődés leírására ott, ahol a tesztkérdéseket dichotomikusan – jó vagy rossz – értékelik (*Wright, B. D.–Stone, M. H.*), illetőleg ahol a tesztkérdésekre adott válaszok értékének megfelelő pontszámokkal minősítenek (*Masters, G. N.*). [További tudományos vizsgálatot kíván az, hogy az egyik országban kifejlesztett mérteszköz alkalmazható-e egy másik országban, ahol eltérő a curriculum tartalma. Egyes kutatók szerint nem alkalmazható (*Miller, M. D.–Linn, R. L.*), mások szerint felhasználható (*Mehrens, W. A.–Phillips, S. E.*)]

Az IRT modell használata széleskörűen elterjedt. Az UNESCO-kiadvány több ausztráliai, kanadai, izraeli, szingapúri, hollandiai és USA-beli példát mutat be. Álljon itt közülük kettő!

Izraelben a tesztelessel és értékeléssel megbízott nemzeti intézet kidolgozott egy komputerizált, adaptív, kategóriákba soroló tesztet az angol mint idegen nyelv számára (*Cohen, Y.–Ben-Simon, A.–Tractinsky, N.*). Mintegy 700 mondatbefejezés és mondat-újrafogalmazást követelő itemet mértek be az item response theory modell használatával. A teszt alkalmazása – ha korszerű technológiával történik az eredmények feldolgozása – potenciálisan a kérdések egyharmadával csökkentheti a papír-ceruza tesztekhez szükséges tesztelési időt.

Hollandiában, a már említett *CITO* a nemzeti értékelési program keretében periodikusan beszámolót készít a nevelésügyi minisztériumnak az alapfokú iskolák tanulójának fejlődéséről. Az első fázisban az intézet a 9 éves és a 12 éves tanulók matematikai tudásának értékelté. A 9 évesek számára 275 kérdést 13 tesztbe, a 12 évesek részére pedig 475 kérdést 27 tesztbe foglaltak össze. A tesztek nem feleletválasztásosak voltak, de a válaszokat dichotomikusan értékelték. A csoporthatás (cluster effect) ellenőrzésére (*Ross, K. N.*) a tesztek mint két iskolázási korosztályban 6 csoportra osztották, aminek eredményeképpen minden egyes item megválaszolását minden iskolából mintegy 4 tanuló, az 1987. év végén legalább 500 tanuló töltötte ki.

Az értékelés jövőben várható trendjei

A tudásszintmérés előbbieken leírt három trendjéhez kapcsolódó továbbfejlődés több dimenzióban megy végbe. Ez magában foglalja a komplex magatartás jobb értékelési stratégiáinak kimunkálását; a tanárok számára speciális képzési programok szervezését az értékelés témakörében; korszerű technológiák alkalmazását az információk gyűjtésének, feldolgozásának és interpretálásának folyamatában, valamint a tanulók, a hallgatók részvételét az értékelésben. Az UNESCO-kiadvány a továbbiakban a felsorolt négy fejlődési trend gyakorlati példáiból ad ízelítőt.

Komplezsebb magatartás-értékelés gyakorlati összefüggésben

Vermeulen és Westerhuis kidolgozták a felső középfokú szakiskolák számára az egészségügyi szakdolgozók által használt interjúk értékelésének egy új eszközét. Az eszköz a szakdolgozó és a szolgáltatást igénybe vevő kliens közötti interjúhelyzetet szimulálja. Ezzel lehetőséget teremt annak értékelésére, hogy a szakiskolai tanuló milyen mértékben rendelkezik azokkal a készségekkel, amelyek szükségesek a szolgáltatást igénybe vevő klienssel készítenő érzékeny és strukturált interjúhoz. Értékelhető továbbá az is, hogy a szakiskolai tanuló – leendő szakdolgozó – az interjú során mennyire képes olyan légkört teremteni, amely segíti a klienst abban, hogy hajlandó legyen problémáiról beszélni. A szimulált helyzetben egy színész vállalja a kliens szerepét. Egyes válaszokat egy kétponos skálán értékelnek (a magatartás megnyilvánult–nem nyilvánult meg), másokat a magatartás minőségének megfelelően egy ötfokú skálán pontoznak.

Tanárok továbbképzése az értékelési technikákban

A tudásszintméréssel foglalkozó kutatók nyomatékosan hangsúlyozzák, hogy ha innovatív oktatási tartalmakat és tökéletesebb oktatási stratégiákat akarunk meghonosítani az iskolákban, akkor szükség van a tanárok továbbképzésére az értékelési technikák terén. A pedagógusoknak el kell sajátítaniuk a pontszámok kombinálásának és csoportosításának az eljárásait annak érdekében, hogy egy-egy komponens ne gyakorolhasson túlzott, aránytalan hatást az értékelésre. Ennek a kívánságnak az indoklására rámutathatunk *Viviani, N.*-nak a pénzzel kapcsolatos analógiáját kiterjesztve: egy USA-dollárnak két angol fonttal való összehasonlítása alapján szerzett 3 pontnak más az értéke, mint a két USA-dollárnak egy angol fonttal való összehasonlításával szerzett 3 pontnak. A feladat megoldása során ugyanis könnyebb megállapítani, hogy egy USA-dollár kevesebb két angol fontnál, mint azt átlátni, hogy két USA-dollár értéke hogyan viszonyul egy angol fontéhoz. Ez utóbbi esetben a helyes válaszhoz ismerni kell a két valuta átváltási arányait.

Korszerű technikák alkalmazása az értékelési folyamatban

A komplex tanulmányi fejlődés közvetlenebb értékelésének távlata kiváltképpen vonzó azok számára, akik arra törekcsenek, hogy az értékelés folyamatában alkalmazott feladatok a curriculumot tükrözzék. *Zieky, M.* hangsúlyozza, hogy az elmélet és a technika ma már képes arra, hogy a tesztelést komputer adaptív tesztek alkalmazásával hatékonyabbá tegye, s ami fontosabb, hogy kevésbé frusztrálja a tesztelteteket. A komputer adaptív tesztek itemjei ugyanis a teszteltet korábbi itemválaszai alapján becsült alkalmasságától függőek. A vizsgált tanulók közül azokat, akik kevésbé alkalmasak, nem frusztrálja az, hogy túl sok olyan itemre kell válaszolniuk, amelyek messze meghaladják kompetenciájukat; a jobb felkészültségű, nagyobb kapacitású tanulókat pedig nem untatja a számukra túlságosan könnyű itemek hosszú sora. Komputer mastery tesztek használata esetén a tesztek hossza a teszteltet alkalmasságának a teljesítmény minimum szintjétől való távolságától függ. Azok, akik messze alatta vagy felette vannak az átlagnak, kompetenciájukat (vagy annak hiányát) relatíve kevés számú itemen tudják tanúsítani.

Tekintélyes erőfeszítések történtek a direktsbb mérések kifejlesztésére korszerű technológia segítségével. A hollandiai *CITO* keretében például *Bosman* interaktív lehetőségekkel bíró videodiszk (képlemez) tesztet fejleszt ki a gyógyszerasszisztens szakma kognitív és gyakorlati aspektusainak az értékelésére. Az

utasítások 5 különböző sorozata olyan helyzeteket hoz létre, amelyek alkalmasak a gyógyszerelőírásokkal és a szociális készségekkel – mint a haragos vagy türelmetlen kliensekkel való kommunikációval – összefüggő általános problémamegoldó készségek értékelésére. A videodiszk tudósítja a tesztelt hallgatót mind a döntő tévedései alapján pontozott akcióin és válaszainak az értékeléséről, mind azoknak a több gyógyszerész ítélete alapján megfogalmazott hatékonyságáról is (*Sanders*, személyes kommunikáció).

Az USA-ban a ProCare interaktív, komputerizált videodiszk rendszert olyan személyeknél használják, akik bizonyítványt kívánnak szerezni a házi egészségügyi gondozás területén. A videodiszk technológiához a ProCare újszerű vonással járul hozzá: egyszerre nyújt képzést és értékelést. A rendszer összeütközésre vezető helyzeteket mutat be, és megkérdezi használójától, hogyan oldaná meg a bemutatott problémát. A vizsgálat vezetői a helytelen válaszokat azonnal javítják, hogy a felhasználók ne őrizzék meg azokat emlékezetükben. A tapasztalatok szerint a videobemutatók sokkal valószerűbb, mint amilyen az írásos bemutatás lenne (*Zieky*, személyes kommunikáció).

A tanulók/hallgatók részvétele az értékelés folyamatában

Amint már említettük a jelenlegi fejlődési irányok tárgyalásánál, néhány oktatási program arra ösztönzi a tanulókat, hogy nagyobb felelősséget vállaljanak tanulmányaikért. *Falchikov, N.* és *Boud, D.* megfigyelték, hogy az életet át tartó tanulás megköveteli, hogy az egyén ne csupán önállóan tudjon dolgozni, hanem legyen képes saját teljesítményének és fejlődésének a megítélésére is. Kimutatták felsőfokú oktatási intézmények hallgatói önértékelésével foglalkozó tanulmányoknak a másodlemzése alapján, hogy ugyanannak a hallgatói teljesítménynek tanári, illetőleg hallgatói megítélésében különbségek vannak. Arra a következtetésre jutottak, hogy „az önértékelés készségnek tekintendő, és mint ilyen fejlesztendő”. Véleményük szerint a jó értékelési gyakorlat magában foglalja az értékelők képzését, legyenek azok tanárok vagy hallgatók. Arra a végkövetkeztetésre jutnak, hogy a tapasztalt tanárok sem megbízható értékelők minden helyzetben.

Az önértékelési készségeket talán már az iskolázás alacsonyabb szintjein fejleszteni kellene. *Withers, G.* és *McCurry, D.* azzal érvelnek, hogy a tanulók részvétele egy oktatási kurzus szerkezetének a megállapításában (tartalmának, az értékelésnek stb.) magasabb rendű, mint az a jelenlegi gyakorlat, amelyben a tanmeneteket és az értékelést felülről határozzák meg. Hangsúlyozzák, hogy a részvétel emeli a tanulók iskolai életének a minőségét, pozitív változásokat eredményez az oktatás tartalmában, növeli az adott kurzus hasznosságát, és segíti a tanulókat, a szülőket és a tanárokat az értékelés módjának és a róla való információknak a közös megértésében.

Batten, J. kutatásai bemutatták, hogy hagyományos középiskolai osztályokba járó társaikhoz képest az értékelésben részt vevők, annak eredményeképpen, elismerik iskolai életük minőségének emelkedését, jobban ellenőrzik saját tanulásukat, pontosabb az énképük, helyesebben ítélik meg saját tudásukat, készségeiket és önnön fejlődésüket.

(*John Izard: Assessing Learning Achievement. Educational studies and documents 60. Published by the United Nations Educational, Scientific and Cultural Organization, Paris, 1992. 47 old.*)

Válogatta és fordította: MAJZIK LÁSZLÓNÉ

VÁLOGATOTT BIBLIOGRÁFIA

- Batten, M.* (1989): Year 12: Student's Expectations and Experiences. Hawthorn, Vic., ACER (ACER Research Monograph No. 33.); *Berk, R. A.* (1986): A Consumer's Guide to Setting Performance Standards on Criterion-referenced Tests. = Review of Educational Research, Vol. 56, pp. 137-72.; *Boulter, P.* (1990): Public Examinations Performance at 16+: 1985/88. In: *C. T. Fitz-Gibbon* (ed.): Performance Indicators. pp. 16-18. Clevedon, Multilingual Matters Ltd. (BERA Dialogues 2.); *Brandma, J.-ten Brummelhuis, A.-Plomp, T.* (1989): Evaluation of National Curriculum and Implementation in Junior Vocational Education. In: Studies in Educational Evaluation, Vol. 15, pp. 359-82; *Clark, D. J.* (1989): Mathematics Curriculum and Teaching Program (MCTP) Alternatives in Mathematics. Canberra, Curriculum Development Centre.; *Cohen, Y.-Ben-Simon, A.-Tractinsky, N.* (1989): Computerised Adaptive Test of English Proficiency. Jerusalem, National Institute for Testing and Evaluation.; *Crooks, T. J.* (1988): The Impact of Classroom Evaluation Practices on Students. = Review of Educational Research, Vol. 58, pp. 438-81.; *de Lange, J.-Verhage, H.* (1990): Data Visualization. Utrecht. Research Group on Mathematics Education (OW&OC), University of Utrecht; *Educational Testing Service* (1987): Learning by Doing: A Manual for Teaching and Assessing Higherorder Thinking in Science and Mathematics. Princeton, Educational Testing Service.; *Falchikov, N.-Boud, D.* (1989): Student Self-assessment in Higher Education: A Meta-analysis. = Review of Educational Research, Vol. 59, pp. 395-430.; *Fitz-Gibbon, C. T.* (ed.) (1990): Performance Indicators. Clevedon, Multilingual Matters Ltd. (BERA Dialogues 2.); *Franke-Wikberg, S.* (1989): Evaluating Education Quality on the Institutional Level. Umea, Department of Education, University of Umea.; *Gamble, R.* (1990): Performance Indicators. In: *C. T. Fitz-Gibbon* (ed.): Performance Indicators. pp. 19-25. Clevedon, Multilingual Matters Ltd. (BERA Dialogues 2.); *Gaulin, C.-Puchalska, E.* (1987): Coded Graphical Representations: A Valuable but Neglected Means of Communicating Spatial Information in Geometry. In: *I. Wirszup and R. Streit* (eds.): Developments in School Mathematics Education around the World: Applications-oriented Curricula and Technology-supported Learning for All Students, pp. 514-39. Reston, Va., National Council of Teachers of Mathematics.; *Glaser, R.* (1990): Towards New Models for Assessment. = International Journal of Educational Research, Vol. 14, pp. 475-83.; *Griffin, P.* (1990): Profiling Literacy Development: Monitoring the Accumulation of Reading Skills. = Australian Journal of Education, Vol. 34, pp. 290-311.; *Hoge, R. D.-Colardaci, T.* (1989): Teacher-based Judgments of Academic Achievement: A Review of the Literature. = Review of Educational Research, Vol. 59, pp. 297-313.; *Huot, B.* (1990): The Literature of Direct Writing Assessment: Major Concerns and Prevailing Trends. = Review of Educational Research, Vol. 60, pp. 237-63.; *Izard, J. F.* (1990): Assessing Achievement with Non-pencil-and-paper Tasks. (Paper presented at the Fifth South East Asian Conference on Mathematics Education, Bandar Seri Begawan, Brunei Darussalem, 14-17 June 1990.); *Joffe, L. S.* (1990): Evaluating Assessment: Examining Alternatives. In: *S. Willis* (ed.): Being Numerate: What Counts? pp. 138-61. Hawthorn, ACER.; *Kenny, D.-Chekaluk, E.* (1990): Early Reading Performance: A Comparison of Teacher-based and Test-based Assessments. (Paper presented at the XXII International Congress of Applied Psychology, Kyoto, Japan.); *Lowe, R. K.* (1989): Scientific Diagrams: How Well Can Students Read Them? What Research Says to the Science and Mathematics Teacher, Number 3. Perth, Key Centre for Teaching and Research in School Science and Mathematics, Curtin University of Technology.; *Masters, G. N.*

(1988): *The Analysis of Partial Credit Scoring*. Applied Measurement in Education, Vol. 1, pp. 279-97.; *Miller, M. D.-Linn, R. L.* (1988): Invariance of Item Characteristic Functions with Variations in Instructional Coverage. = Journal of Educational Measurement, Vol. 25, pp. 205-19.; *Newman, M. A.* (1977): An Analysis of Sixth-grade Pupils' Errors on Written Mathematical Tasks. In: *M. A. Clements and J. Foyster* (eds.): Research in Mathematics Education in Australia 1977. Vol. 1. pp. 239-58, Melbourne, Swinburne College Press.; *Northern Territory Department of Education* (1990): Primary Assessment Program: Children's Writing in the Northern Territory 1989. Darwin, Northern Territory Department of Education.; *Palincsar, A. S.-Winn, J.* (1990): Assessment Models Focused on New Conceptions of Achievement and Reasoning. = International Journal of Educational Research, vol. 14, pp. 411-13.; *Ross, K. N.* (1985): Sampling Errors. In: *T. Husen and T. N. Postlethwaite* (eds.): The International Encyclopaedia of Education, pp. 4381-51. New York, Pergamon. *Ross, K. N.* (1987): Sample Design. International Journal of Educational Research, Vol. 11, pp. 57-75.; *Spence, I.-Lewandowsky, S.* (1990): Graphical Perception. In: *J. Fox and J. S. Long* (eds.): Modern Methods of Data Analysis. Newbury Park, Calif, Sage.; *Suggett, D.* (1985): Guidelines for descriptive assessment. Rev. ed. Melbourne, Victorian Institute of Secondary Education.; *Viviani, N.* (1990): The Review of Tertiary Entrance in Queensland 1990. Brisbane, Department of Education.; *Windham, D.-Levin, H.-Bathory, Z.* (1990): Issues in the Design and Development of Educational Information Systems. In: *K. N. Ross and L. Mählck* (eds.): Planning the Quality of Education: The Collection and Use of Data for Informed Decision-making, pp. 79-95. Paris/Oxford, UNESCO/Pergamon.; *Withers, G.-Batten, M.* (1990): Defining Types of Assessment. In: *B. Low and G. Withers* (eds.): Developments in School and Public Assessment. Hawthorn, ACER. (Australian Education Review, 31.); *Withers, G.-McCurry, D.* (1990): Student Participation in Assessment in a Co-operative Climate. In: *B. Low-G. Withers* (eds.): Developments in School and Public Assessment. Hawthorn, ACER. (Australian Education Review, 31.); *Wright, B. D.-Stone, M. H.* (1979): Best Test Design: Rasch Measurement. Chicago, MESA Press.; *Zieky, M.* (1989): Methods of Setting Standards of Performance on Criterion Referenced Tests. = Studies in Educational Evaluation, Vol. 15, pp. 335-8.