



ÖSSZETETT KÉPESSÉGEK MÉRÉSE: DILEMMÁK ÉS MEGKÖZELÍTÉSEK

Pálvölgyi Lajos

ELTE Eötvös Loránd Tudományegyetem Neveléstudományi Intézet

Az összetett képességek mérése a pedagógia igen fontos területe, amely több nyitott kérdést vet fel. Számos empirikus adatfelvétel, tanulóiteljesítmény-mérés nélkülözi a korrekt méréselméleti megalapozást, gyakran alkalmaznak olyan statisztikai módszereket, amelyek csak intervallumskála esetén lennének megengedhetők, de ennek feltételei nem teljesülnek. A megfelelő mérési szint hiánya olyan teszteknel is felmerülhet, amelyek fontos döntéseket alapoznak meg, mint amilyen például az egyetemi felvétel vagy egyes álláshe-lyek betöltése. A dilemmák két, egymással részben összefüggő alapkérdésre vezethetők vissza: az egyik a „mit mérjünk”, a másik a „hogyan mérjünk” kérdése.

Ez a tanulmány a második kérdéshez kapcsolódó nehézségeket és ezek egyes meg-oldási lehetőségeit tekinti át, figyelemmel a többdimenziós modern tesztelmélet (*multidimensional item response theory*) eredményeire is. Mivel a fejlettség rendezést jelent, egy sokdimenziós képesség fejlettségi állapotát úgy kellene jellemezni egyetlen számmal, hogy az megfeleljen a reprezentációs méréselméletnek, és legalább ordinális skála jöjjön létre. Az összetett képességek többdimenziós jellegéből fakadó méréselméleti problémák megoldását segítheti a javasolt újszerű megközelítés. A tanulmány a megoldást a nemzetközi főáramtól eltérő alternatív irányban keresi, de megtartja annak valószínűségi szemléletét. A dimenzionális megközelítéssel szemben egy *feladatorientált modellt* mutat be. A képességet azokkal a feladatokkal jellemzi, amelyek megoldását az lehetővé teszi.

A modellben központi szerepet játszik a *képességvolumen* koncepciója, amely az adott képességállapotot az ezen képességállapot eléréséhez szükséges szokásos tanulási-gya-korlási időráfordítás mértékével jellemzi, alapul véve valamely meghatározott tanulói re-ferenciacsoport adatait. Az így mért képességállapotok rendezhetők; sőt arányskála szin-ten válik mérhetővé a ráfordításalapú képességvolumen véges sok feladatosztály esetén anélkül, hogy az összetett képesség látens dimenzionális szerkezetének feltárása szüksé-gessé válna. A képességvolumen pedagógiai szempontból fontos és jól hasznosítható fel-világosítást adhat az összetett képesség feladatosztályok halmazával reprezentált terjedel-méről. A javasolt eljárás nem magát az összetett képességet igyekszik megragadni annak teljességében, minőségi sokféleségében, hanem egy pedagógiai szempontból lényeges tu-lajdonságát emeli ki, ezzel lehetőséget adva a magasabb szintű skálákra épülő megalapo-zott statisztikai elemzésekre.

Problémafelvetés

A mérés során egy halmaz elemeihez számokat rendelünk meghatározott szabályok szerint. A mérési szintet az alkalmazott hozzárendelési szabályok határozzák meg. Az erősebb szabályok magasabb mérési szintet valósítanak meg, és több lehetőséget biztosítanak a kapott adatok statisztikai elemzésére. Stevens (1946) nyomán négy alapvető mérési skálát különböztetünk meg. A legegyszerűbb skála a nominális skála, amely csak a legalapvetőbb relációt ismeri a mérendő dolgok között: azonos osztályba tartoznak-e vagy sem. A következő szint az ordinális skála, melynél létezik egy rendezési reláció (sorrendiség). Az igazán hatékony statisztikai elemzések az intervallumskálák alkalmazása esetén lehetségesek, a számok itt már az eltérések mértékét is tükrözik. Végül az arányskála képviseli a legmagasabb szintet. A reprezentációs méréselmélet szerint a számokkal végzett műveletek feltétele az, hogy az adott műveletnek legyen valamilyen empirikusan értelmezhető megfelelője a mérendő dolgok valós világában.

Az elmúlt évtizedekben számos publikáció látott napvilágot, melyek a mérési skálátípusok kibővítését célozták, vagy kapcsolódó kritikát, kiegészítést fogalmaztak meg (pl. Hand, 1996; Michell, 1999; Narens, 2002; Velleman & Wilkinson, 1993). A méréselmélet történetét és a fejlődését kísérő vitákat részletesen bemutatja Kehl (2011). Jelen írás kizárólag a pedagógiai mérésekre, ezen belül az összetett képességek mérésének módszertani alapkérdésére szorítkozik.

A pedagógiai mérések problémáit tárgyalva Nahalka (2015) hangsúlyozza, hogy sok adatfelvétel nélkülözi a korrekt méréselméleti megalapozást. Gyakran alkalmaznak olyan statisztikai módszert, amely csak intervallumskála esetén lenne megengedhető, holott ez nem teljesül. Álláspontja szerint a pedagógiát érdeklő összetett képességeknél az akadályozza a mérést, hogy ezek nem egydimenziós jelenségek, és ezért semmilyen skálán nem tudjuk kifejezni fejlettségi szintjüket; az összetett képességek nem rendezhetők (Nahalka, 2015, pp. 30–31).

Más publikációkban is találkozunk hasonló problémafelvetéssel (Cervone & Caldwell, 2008; Trendler, 2009). Tekintettel arra, hogy a magasabb mérési szintek gyakran nem elérhetők, Heilmann (2015) azt javasolja, hogy a reprezentációs méréselméletet inkább a számszerű vizsgálatokat segítő hasznos elvek és szabályok egyfajta könyvtárának tekintsük, melyek lehetőség szerint alkalmazhatók. Azonban a mérési szint problémáját ezzel nem oldja meg. A reprezentációs méréselmélet és a pszichometria viszonyát vizsgálva Vessonen (2020) úgy véli, hogy ezek kiegészítik egymást; és kívánatos lenne, ha csökkenne köztük a távolság, a kölcsönös elszigeteltség. Norman (2010) amellett érvel, hogy számos paraméteres statisztikai vizsgálat eléggé robusztus ahhoz, hogy ne vezessen alapvetően rossz eredményre akkor, ha az ezekhez szükséges intervallumszintű adatok helyett ordinális szintűeket alkalmazunk. Érdeemes hozzátenni, hogy ez viszont egyes konkrét esetekben nehezen bizonyítható, és gyakran nem is áll fenn, másrészt sokszor ordinális szintű adatokig sem sikerül eljutni. Ha viszont sikerül, akkor ezek meghatározott feltételek fennállása esetén becslült intervallumszintű adatokká nemesíthetők (Granberg-Rademacker, 2010).

Ebben a tanulmányban azt a kérdést vizsgáljuk, hogy az összetett képességek mérésakor biztosíthatók-e valamiképp legalább az ordinális skála feltételei. És ha igen, akkor miként; vagy ha nem, akkor miért nem. Érdeemes megjegyezni, hogy bármennyire is alapvető a mérés, nem az egyetlen lehetséges módja annak, hogy megteremtjük a matematika alkalmazási feltételeit a pedagógiában. A számok egyfajta modellnek tekinthetők, de emellett szóba jöhetnek egészen más modellek is. A pedagógia sajátos struktúráinak feltárása igen változatos módon akár újszerű matematikai modellek útján is elképzelhető (Pálvölgyi, 1981). Kulcsszerepet játszhatnak ebben a gráfokkal leírható struktúrák. Erre mutat példát a szociometria és a hálózat kutatás mellett a témánk szempontjából különösen érdekes tudástér-elmélet (Robusto, Stefanutti, & Anselmi, 2010; Tóth, 2005), illetve a fogalmi struktúrák vizsgálata Galois-gráfokkal (Fatalin, 2008; Takács, 2000).

Dimenziókra bontás

Az összetett képesség olyan látens komponensekkel rendelkezik, amelyek egymástól függetlenül (vagy részben függetlenül) változhatnak. Jelen tanulmányban egy összetett képesség *dimenzióját* olyan homogén komponensként értelmezzük, amely ilyen további relatív független részekre már nem bontható. A homogenitás következtében egy adott dimenzió mérésére irányuló elemi kérdések vagy feladatok (itemek) eredményei nagymértékben korrelálnak egymással és magas Cronbach-alfa értéket mutatnak. A klasszikus tesztelmélet megbízhatósági (*reliabilitás*) mutatói azt jelzik, hogy az itemek ugyanazt a dimenziót mérik.

A probléma megoldásának egyik iránya lehet, hogy a vizsgált összetett képességet tartalmi elemzés alapján homogén dimenziókra bontjuk, hogy azután ezeket külön-külön mérhessük. Ha ez megvan, akkor az összetett képesség az egyes dimenziók mérőszámával jellemezhető. Az eredmény tehát nem egyetlen szám lesz, de nem kell, hogy feltétlenül nagyon sok szám legyen. A személyek az egyes dimenziókban összehasonlíthatóvá válnak, mint ahogyan egy adott személy fejlődése is jellemezhető egy-egy dimenzió mentén. Összesítésre viszont nincs mód, mivel ennek méréselméleti alapja hiányzik. A többdimenziós megközelítést jól illusztrálják az eDia online diagnosztikus rendszer tartalmi keretei¹. A matematika, természettudomány és olvasás-szövegértés tárgyú értékelés ebben a rendszerben a Csapó (2008) által kifejlesztett, a tanulást és tudást meghatározó pszichológiai (gondolkodási), társadalmi-kulturális (alkalmazási) és tantárgyi (diszciplináris) dimenziókra épül.

Egy elméleti modell konstruálása során előfordulhat, hogy az első lépésben kapott komponensek bár jól elkülönülnek, és kijelölik a dimenziókra bontás koncepcionális irányait, de szigorú értelemben még nem mindegyik tekinthető homogén dimenziónak. Az ilyen esetekben a lebontást tovább kell folytatni. A dimenziókra bontás megalapozottságát és egyértelműségét a gyakorlatban sokszor nehéz megfelelően biztosítani. Az adott képes-

¹ Kapcsolódó publikációk: <http://edia.hu/projekt/?q=hu/konyvek>

séget leíró, elméletileg megalapozott modelltől indulunk ki (ha van ilyen), és azt vizsgáljuk, hogy van-e illeszkedés a modell komponenseire mint dimenziókra épített teszt adatai és a modell között (l. pl. Molnár & Csapó, 2019). Az illeszkedés arra utal, hogy a modell valós viszonyokat tükröz. Felmerül a kérdés, hogy teljesen lefedi-e a vizsgált képességet (terjedelmi érvényesség). Ez az alkalmazott elméleti modelltől függ.

Egyes esetekben előfordulhat az is, hogy másként felépített modell alapján is készíthető olyan teszt az adott képességhez kapcsolódóan, amely illeszkedést produkál a saját modelljével. Itt az operacionalizálás általános problémájába ütközünk. Egy elvont fogalommal – bármennyire is jól definiált – miképpen lehet azonos, vagy annak teljesen megfelelő egy nagyon konkrét teszteszköz, pontosabban az abban tükröződő fogalomértelmezés? A teszt a maga konkrétságában de-facto újraértelmezi a fogalmat és annak helyébe lép. Aki mindkettőt ismeri, rábólinthat, hogy igen, a két dolog valóban megfelel egymásnak (tartalmi érvényesség), míg egy másik személy felvetheti, hogy talán mégsem, mert a teszteszköz kicsit szűkebb vagy tágabb, vagy másmilyen ahhoz képest, ahogyan ő személy szerint az adott képességet felfogja, terjedelmét, természetét, jelentését magának megkonstruálta. Különösen így lehet ez akkor, ha a fogalmat nem kellő egyértelműséggel határozták meg, illetve ha értelmezése erősen vitatott. További probléma a dimenziókra bontással az, hogy akár extrém sok dimenzióra is szükség lehet, és a részek összege nem mindig adja ki az egészet. Általánosabb képességeknél ez jellemzően felléphet, és ha fellép, valamiképp kezelni kell (keveset mondunk el egy festményről akkor, ha megmondjuk, hogy melyik festékből hány grammot használt a festő).

A 20. század elején született klasszikus tesztelmélettel szemben a valószínűségi szemléletű modern tesztelmélet (*IRT – item response theory*) már kielégíti a reprezentációs méréselmélet követelményeit. Leggyakrabban alkalmazott módszere a Rasch-modellre épül (Molnár, 2013), ami a képességfejlettség és a feladatnehézség látens változóihoz intervallumszintű skálát képes konstruálni. Azonban az eljárás egyik feltétele, hogy a vizsgált képesség egydimenziós legyen (Molnár, 2006, p. 109).

A modern tesztelmélet rendelkezik többdimenziós modellekkel is, melyek a teszt-konstrukció jelenlegi „csúcstechnológiáját” képviselik (An & Yung, 2014; Chalmers, 2012; Immekus, Synder, & Ralston, 2019; Li, Jiao, & Lissitz, 2012). Bonyolultabb matematikai háttérük miatt alkalmazásuk nehezebb és gyakran nagyobb elemszámú mintát igényelnek. Ezek a modellek az itemek változásait egynél több látens változóval (vonással) képesek magyarázni meghatározott feltételek mellett. A dimenziók azonosításának módszereiről Svetina (2013) áttekintése ad értékelést. Hartig és Hühler (2009) a többdimenziós IRT szerteágazó lehetőségeit három alkalmazási terület bemutatásával illusztrálják: (1) egy egydimenziós konstruktum mérésekor tévesen felmerülő többdimenzionalitás kezelése, (2) a képességdimenziók közötti látens kovarianciastruktúrák modellezése és (3) egyes tesztitemek megoldásához szükséges többféle képesség interakciójának modellezése. A többszintű IRT-modellek az egyéneket (pl. tanulókat) valamilyen csoportba vagy szervezeti egységbe (pl. osztályba) ágyazottan vizsgálják (Sulis & Toland, 2016).

Az IRT-modellek aktuális problémái közül kiemelhető, hogy az általuk kezelhető dimenziók száma túlságosan korlátozott a pedagógiai szempontból releváns képességek komplexitásához képest, továbbá itt is felmerül a kérdés, hogy lehet-e a dimenziók mérőszámait valamiképp integrálni, és a mérés eredményét egy számmal kifejezni az összetett

képesség átfogó jellemzése érdekében. A különböző skálák összeigazításának módszerei (*scale alignment*, pl. Feuerstahler & Wilson, 2019) nem válaszolják meg ezt a kérdést, mivel nem ezt a célt szolgálják. Éppen a dimenziók lényegi különbözősége nehezíti az összegző jellemzés elméleti megalapozását. Ha valamiképp mégis összegzünk, félrevezető lehet, ha az egyik dimenzió magas szintje kompenzálja a másik alacsony szintjét. Ezért az ilyen kompenzációs modellekkel szemben DeMars (2016) részleges kompenzációs modellek alkalmazását javasolja, ám kénytelen hozzátenni, hogy ezek paraméterei nehezen meghatározhatók. Meglehetősen komplex szimulációs példát mutat ezen paraméterek becslésére egy konkrét vizsgálat esetében. A szimuláció általában is a modellek értékelésének egyre gyakoribb módszerévé válik (Feinberg & Rubright, 2016).

Új perspektívát nyithat az ismertett problémák kezelésére az együttes mérés (*conjoint measurement*) elmélete (Bouyssou & Pirlot, 2005; Briggs, 2013). Nahalka (2018) szerint az igazán összetett képességek mérésére jelenlegi formájában még ez sem nyújt megfelelő megoldást. Hasonló kritikai következtetésre jut Trendler (2019) is.

Feladatorientált megközelítés

A következőkben a megoldást a nemzetközi főáramtól eltérő irányban keressük, de megtartva annak valószínűségi szemléletét. A dimenzionális megközelítéssel szemben felvetjük a feladatorientált megközelítés lehetőségét. Első lépésben ehhez egy elméleti modellt mutatunk be a kapcsolódó fogalomdefiníciókkal.

A képesség értelmezése

A *képesség* értelmezésünk szerint meghatározott feladatok elvégzését lehetővé tevő pszichikus és testi tulajdonságok egysége. Az összetett képességek részképességekre bonthatók, melyekre hasonló definíció érvényes. A képességek alkalmazása rendszerint ismeretek mobilizálását is szükségessé teszi. A gördülékeny fogalomhasználat érdekében ebben a modellben a képességet úgy értelmezzük, hogy fogalmilag az ismeretek birtoklása, pontosabban ezt bizonyító demonstrálása (elmondás, leírás, kérdésre válaszadás) is beleérthető legyen. Ez könnyen megtehető úgy, hogy az ismeretek előhívását előíró egyszerű feladatokat is beleértjük a képességhez rendelt feladatok körébe. Az igazán izgalmas feladatok természetesen nem azok, amelyek ismeretek egyszerű felidézésére irányulnak, hanem sokkal inkább azok, amelyek ezek valódi alkalmazását igénylik.

Feladatorientált megközelítésünk abból indul ki, hogy a képességeket azokkal a feladatokkal jellemezhetjük, amelyek megoldására irányulnak. A feladatok száma akár végtelen is lehet, emiatt érdemes inkább feladattípusokról beszélni. A *feladattípus* alatt a lényeges tulajdonságok tekintetében azonos feladatok véges vagy végtelen halmazát értjük. A lényeges tulajdonságok a feladatok megoldását befolyásoló azon jellemzők, paraméterek és körülmények, amelyek nem kapcsolhatók a feladatot megoldó személyhez. Ide tartoznak a feladat azon belső paraméterei, jellemzői, amelyek a feladat jellegét, tartalmát és

komplexitását befolyásolják, a feladatokat egy tanulócsoporthoz számára könnyűvé vagy nehezzé teszik, amelyek a feladatot különböző kontextusban jelenítik meg, vagy más és más ismeretek mobilizálását, szabályok alkalmazását igénylik. Ide tartozik a feladat megoldását befolyásoló összes külső körülmény is, melyek kisebb-nagyobb mértékben segíthetik vagy akadályozhatják a megoldást.

A szűken vett feladat és megoldásának adott körülményei, környezete a személy számára szituatív egységet alkotnak. Ugyanaz a feladat más környezetben lehet kicsit könnyebb vagy nehezebb – például feleltetés esetén a kérdező lehet kedves és biztató, vagy lekezelő és gyanakvó: ez akár lényegesen befolyásolhatja az aktuális eredményt. Feladat helyett tehát inkább *feladathelyzetről* van szó. A következőkben a feladat fogalmát a feladathelyzet értelmében használjuk, vagyis minden olyan jellemzőt és körülményt beleértünk, amely a megoldást befolyásolja, befolyásolhatja, és külső tényezőként jelenik meg a megoldó személy számára.

A feladatmegoldás sikerességét befolyásoló belső tényezőket a megoldó személy mérni kívánt képessége, illetve az ezekbe foglalt vagy ezekhez kapcsolódó releváns ismeretek mellett az attitűd jellegű jellemzőknek nevezhető tényezők nehezen áttekinthető sokasága alkotja, melyek tudatos és nem tudatos formában is szerepet játszhatnak a sikeres vagy sikertelen megoldásban. Ilyenek például a motiváció, az érdekek, az értékek, a beállítódások és az érzelmek. Ezek jelentős része tartós hatást gyakorol. Különösen az érzelmek egy része szituatív jellegű is lehet, és részben a feladathelyzet egyes konkrét jellemzőire adott válaszként jelenik meg, vagy a teszt elvégzése előtti órák, napok személyes történéseit tükrözi.

A képességmérést a fentiek alapján tehát a feladat oldalán az nehezíti, hogy a szűken vett feladat olyan külső körülmények között jelenik meg, amelyek a megoldást befolyásolhatják, vagyis a feladathelyzet eltérő lehet. A személy oldalán pedig az nehezíti, hogy a képességhez (meg tudja-e tenni) olyan attitűd jellegű, érzelmi és más hasonló jellemzők társulnak, amelyek egy része tartósan, másik része szituatív jelleggel befolyásolja a megoldást (meg akarja-e tenni). A feladat oldalán jelentkező probléma a feladathelyzet standardizálásával kezelhető. A második probléma izgalmasabb. Ez a kitöltő személy napi formáján túl arra is utal, hogy a képesség mindig egy még összetettebb struktúrába, a *kompetenciába* ágyazódik, és hiába igyekszünk leszűkíteni a vizsgálat tárgyát, a kompetencia nevezett elemei felbukkannak, beleszólnak. Szigorú értelemben tehát nincs képességmérés, hanem csak kompetenciamérés van. Vagy egyik sincs.

A továbbiak szempontjából feltételezzük, hogy a képességmérés során semleges vagy enyhén pozitív légkört teremtő, a lényeges paraméterek tekintetében standardizált feladathelyzetekben dolgozunk, és ezt megfelelően dokumentáljuk. Amennyire lehet, igyekszünk kizárni a zavaró faktorokat, az extrém negatív vagy extrém pozitív szituatív behatásokat, tényezőket, érzelmi állapotokat, illetve ha ilyet tapasztalunk, elhalasztjuk a mérést. Pozitív esetben így tehát a képesség mellett elvileg már csak a tartósan fennálló attitűd jellegű jellemzők maradnak játékban. Ezek befolyásolni fogják a képességmérés eredményét, amit el kell fogadnunk. Például egy matematika feladat esetében szerepet játszhat, hogy ki milyen gyorsan adja fel, illetve megfelelő önbizalom és korábbi sikerélmények birtokában érdeklődéssel, bizakodással és kitarással dolgozik-e a feladaton, esetleg élvezi-e

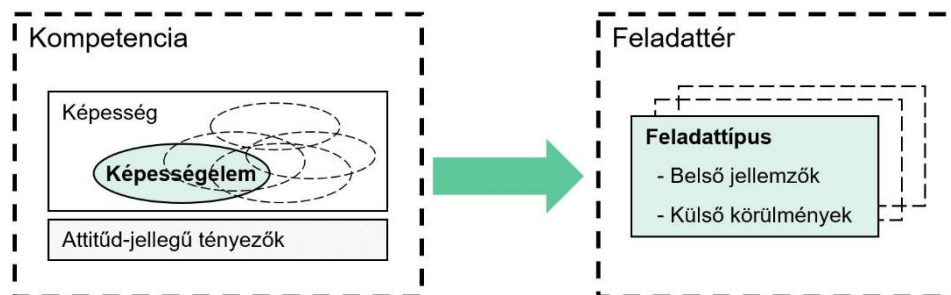
azt. Összefoglalva: ha minden jól megy, akkor a tartósan érvényesülő attitűd jellegű jellemzők által befolyásolt képességet mérjük (vagy becsüljük).

Feladattér

A képesség jellemzésekor a továbbiakban mindig feladattípusokról beszélünk, és nem egyszerűen feladatokról. Az adott képességhez tartozó feladattípusok összessége a *feladattér*. A feladattér hasonló feladattípusait egymás szomszédainak hívjuk. A hasonló kevés tulajdonságban, kismértékben térnek el egymástól. A feladattér fogalma közel áll a tudástér fogalmához (Doignon & Falmagne, 2016), de nem azonos azzal.

A *képességelem* egy meghatározott feladattípus megoldását lehetővé tevő pszichikus és testi tulajdonságok egysége. A képességelem az adott feladattérrel jellemzett képesség része, de általános esetben annak nem struktúrát egyszerűsítő (hierarchikus) lebontása, nem dimenzió (vagy vonás), hanem a feladattípus jellegétől függően maga is többdimenziós összetett struktúra. A képességelem ebben az értelmezésben a képességnek a feladattípusok kialakításától függő metszete, részhalmaza, a feladattípusnak megfelelő, azt tükröző absztrakció. Mivel tág határok között változhat, rendszerint nem (vagy csak nagyon csekély mértékben) jellemzi a képesség feltételezhető belső pszichikus struktúráját. Speciális esetben konstruálhatók olyan feladattípusok is, amelyeknek az adott képesség meghatározott dimenzióinak meghatározott szintjei mint speciális (homogén) képességelemek felelnek meg. Ezt célszerű is megtenni, ha lehetséges, illetve rendelkezésre áll egy ezt megalapozó elméleti modell.

Mindezeket szemlélteti az 1. ábra. Emlékeztetünk itt arra, hogy a képesség fogalmát praktikus okokból jelen modell szempontjából kiterjesztettük úgy, hogy az ismereteket is lefedje, ezért az ismeretek az ábrán nem jelennek meg külön címszóval. Másrészt bár a feladattípusok kialakításuk folytán mindig diszjunkt halmazok, ugyanez nem mondható el a képességelemekről. A képességelemek általános esetben átfedik egymást. Kapcsolatuk és mibenlétük mélyebb vizsgálata nem tárgya ennek a modellnek.



1. ábra

Az összetett képességek feladatorientált megközelítése a képességelemek és a feladattípusok megfeleltetésére épül (saját szerkesztés)

A feladatorientált megközelítés a képesség tartalmát a feladatok nyelvén konkretizálja és terjedelmét lehatárolja. E pontosítás nélkül nehezen tudnánk megmondani, hogy voltaképp mit is szeretnénk mérni. Az egyszerűbb képességeknek kevesebb számú egyszerűbb feladattípus, az összetett képességeknek több és részben bonyolultabb feladattípus felel meg. Az értelmezést azon képességekre korlátozzuk, amelyek véges számú feladattípussal leírhatók. A feladattípusok száma lehet extrém magas is, és a bennük foglalt lehetséges feladatok száma pedig akár végtelen.

Az itt bemutatott megközelítés fókuszában nem a mérni kívánt képesség belső pszichikus struktúrája áll. Az egyes feladattípusokat a feladatok lényeges tulajdonságai alapján határozzuk meg, ügyelve arra, hogy a feladattípusok összessége (a feladattér) lefedje mindazt, amit a képességtől külső teljesítményoldalon elvárunk. Ez egy követelményekből kiinduló megközelítés. Szakmai képességek esetében célszerű olyan referenciacsoport kiválasztása, amelynek tagjai közmegítélés szerint az adott szakmai képességgel megfelelő szinten rendelkeznek, és a tevékenységük alapján célszerű meghatározni, illetve skálázni a feladattípusokat. A skálázás alatt azt értjük, hogy a feladattípusok között legyenek egyszerűbbek és összetettebbek, illetve könnyebbek és nehezebbek is, miközben azok fedjék le, reprezentálják az adott képesség összes részterületét, alkalmazásának tipikus helyzeteit, teljes tartalmi spektrumát. Ha ezt nem teszik, akkor tudatosan vállaljuk és deklaráljuk, hogy a képességnek csak egy részterületéről – vagy egy jól meghatározott szintjéről – van szó.

Egyes feladattípusok esetében lehet ugyan arra törekedni, hogy legyenek olyanok is, amelyek a képesség egyes jól meghatározott komponenseinek feleljenek meg, de a feladattípusok kialakításakor ez modellünkben nem követelmény, mivel sokszor nem is lehetséges. Általános esetben abból indulhatunk ki, hogy minden feladattípus az adott képesség többféle belső komponensét (dimenzióját) igényli (de nem feltétlenül mindig mind-egyiket) és különböző mértékben – azt, hogy melyeket és milyen szinten, nem tudjuk, és jelen megközelítés keretében nem is vizsgáljuk.

Képességjellemzés vektorral

A képesség mérése annak számszerű jellemzése oly módon, hogy a különböző személyek aktuális képességállapotához számokat rendelünk, biztosítva azt, hogy a számok az aktuális képességek között fennálló bizonyos viszonyokat tükrözzenek, így lehetővé téve meghatározott műveleteket a számok között. Például, ha a képességállapotok között értelmezhető a rendezési reláció, akkor az értelmezhető az azokat jellemző számok között is, és ha más reláció nincs, akkor ezzel ordinális mérési skálát kapunk. A cél az, hogy lehetőleg minél magasabb szintű skálát kapjunk, hogy a tesztek eredményeit minél többféle statisztikai elemzésnek vethessük alá. A pontosság tekintetében az észszerű határokon belül egyes esetekben megelégszünk közelítő becslésekkel is, amennyiben ezek még elfogadható szinten tükrözik a valóságosnak gondolt (látens) értékeket. Ilyen kompromisszumok a természettudományokban is előfordulnak.

A képességek feladattípusokkal történő jellemzése kézenfekvő megközelítés. Valójában már a képességek elnevezésében is rendszerint azon feladatokra utalunk, amelyek elvégzéséről szó van. Meghatározott feladatokra gondolunk, és ezt vetítjük vissza az

egyénbe, mert úgy gondoljuk (nem alaptalanul), hogy az adott cselekvésre képes személyben van valami, ami őt erre a cselekvésre képessé teszi. A személyben rejlő előfeltételek összességét tekintjük képességnek anélkül, hogy ezt pontosabban körülhatárolni, illetve mélyebben definiálni tudnánk. A gyakorlatban ez rendszerint senkit nem zavar. A konkrét gyakorlati helyzetekben (például munkahelyen) elsősorban nem az a fontos, hogy mi van az egyénben, hanem sokkal inkább az, hogy „mi jön ki” belőle, vagyis mely feladatokat képes (és hajlandó) ellátni, hogyan működik együtt a többiekkel. A hatékony munkahelyi vezető ismeri a csapattagok munkakör szempontjából releváns képességeit, tulajdonságait anélkül, hogy idevágó méréseket folytatott volna. Ismerete a gyakorlati feladatmegoldások során tett megfigyelésekre, a nyert tapasztalatokra épül.

Egy adott személy aktuális *képességállapotát* az egyes feladattípusokhoz tartozó feladatok megoldási valószínűségeivel jellemezhetjük. A feladattér tetszőleges „ F ” feladattípusához tartozó feladat aktuális megoldási valószínűsége legyen: $p(F)$, ahol: $0 \leq p(F) \leq 1$.

Egy adott személy esetében a $p(F)$ értékét megfelelően közelíti az adott feladattípusból véletlenszerűen kiválasztott feladatok megoldásának relatív gyakorisága nagyszámú kísérlet esetén, feltéve, hogy sem tanulás, sem fáradás, sem felejtés, sem egyéb olyan változás nem történik, amely a tanuló aktuális képességállapotát – és ezzel a megoldások sikerességét nem elhanyagolható mértékben – módosítaná a feladatmegoldások sorozatának időtartama alatt. Ezek alapján egy adott személy képességállapotát az összes feladattípus-hoz tartozó $p(F)$ valós számokból képzett rendezett számsorozat jellemzi. Például egy ilyen: (1, 0,82, 1, 0,78, 0,56, 0, 0,3 ...).

Ez a rendezett véges számsorozat matematikai értelemben (sokdimenziós) vektornak tekinthető. Két képességállapotot azonosnak tekintünk, ha a két vektor egymásnak megfelelő egyes értékei rendre egyenlők egymással (az első az elsővel, a második a másodikkal és így tovább), vagy ezek csak elhanyagolható mértékben különböznek egymástól. Amennyiben összetett képességről van szó, a vektornak sok számot kell tartalmaznia ahhoz, hogy minden feladattípus megoldási valószínűségét jellemezni tudja, ami ezt a kontrukciót gyakorlatilag kezelhetetlenné teszi.

Található-e megfelelő skalár mérőszám?

A bemutatott vektorokkal szemben a skaláris mennyiségek egy mérőszámmal és egy mértékegységgel egyértelműen jellemezhetők. Ha a képességet egyetlen számmal kívánjuk jellemezni, akkor rögtön felmerül, hogy egyetlen számtól talán nem is várható el az, hogy mindazon információt teljeskörűen hordozza, amit egy ilyen sok elemből álló vektor hordozni képes. Ez igaz is, de megváltozik a helyzet, ha azt a kis engedményt tesszük, hogy a valószínűséget nem folytonos skálán mérjük, hanem például a 0, 1, 2 ... 100 százalékokból álló 101 elemű skálán becsüljük. Ily módon egy véges modellt kapunk, ahol a feladattípusok száma is és a lehetséges valószínűségek száma is véges. Ebben a modellben csak véges számú különböző vektor fordulhat elő. Mindegyik vektor kaphat egy sorszámot, amely szám alapján a vektor visszakereshető. Így mégiscsak megkapjuk az óhajtott egyetlen számot. Ez a szám azzal az érdekes tulajdonsággal rendelkezik, hogy elvont értelemben a teljes képességvektort reprezentálja.

Ezzel a megközelítéssel viszont legalább három gond van: (1) nagyon nagy számokról van szó, (2) a szám önmagában közvetlenül nem értelmezhető, (3) ez a szám nem ad lehetőséget statisztikai elemzésekre. Hiszen ez pusztán kódolás, egy kezelhetetlen nominális skála, nem több. Hiányzik a számok között olyan struktúra, reláció, amire méréselméleti szempontból építeni lehet. Úgy tűnik tehát, hogy ez a megközelítés zsákutcába vezet.

A képességállapotot szemléltető dobókocka modell

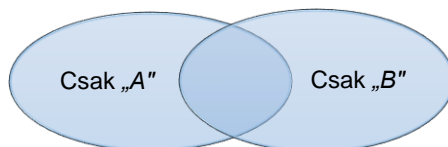
A leírt véges modell lehetőséget nyújt a feladatorientált megközelítés alapgondolatának egyszerű szemléltetésére. Adott személy képességállapotát ennek alapján egy dobókocka modellel illusztráljuk. Képzeljük el, hogy mindenkinek van egy doboza, és ebben annyi dobókockát őriz, ahány feladattípus tartozik az adott képességhez. Minden feladattípushoz tartozik egy-egy szabályos dobókocka. Minden kockának legyen például 100 oldala. (Most ne álljunk meg ott, hogy létezik-e 100 oldalú szabályos test, és ha igen, akkor azt minek nevezik.) Mindegyik oldalra az *igen* és a *nem* szavak valamelyikét írták. A feladat megoldását úgy képzeljük el, hogy ha meg kell oldani egy feladatot, akkor az illető előveszi a dobozból a feladattípusnak megfelelő kockát, és dob egyet. Ha az *igen* van felül, akkor megoldotta a feladatot, ha a *nem*, akkor nem oldotta meg. Ha például 65 *igen* van az adott kockán, az 65% megoldási valószínűséget reprezentál az adott feladattípus vonatkozásában.

A dobókockák összessége ily módon képességének aktuális állapotát tükrözi. A tanulás-gyakorlás során egy vagy több kockán növekszik az igenek száma; felejtés esetén pedig a nemek száma. A dobókocka a képességelemnek felel meg. Míg a dobókockák ebben az egyszerű modellben egymástól függetlenek, a valóságban a képességelemekről ez rendszerint nem mondható el. E szemléletes példával könnyen értelmezhető a team képessége is. Ha egyetlen személy teljesítménye elegendő a csapatból a team előtt álló feladat megoldásához, akkor az események összegének valószínűsége játszik szerepet, ha pedig mindenki teljesítménye szükséges, akkor az események szorzatának valószínűségét kell számolni a megfelelő dobókockák alapján a valószínűségszámítás szabályai szerint.

Miként modellezhető a képesség fejlődése?

Az adott személy képességállapotát ábrázoló nagy elemszámú vektorra visszatérve, problémát jelent az, hogy a vektorral jellemzett személyek képességállapotai nem rendezhetők sorrendbe a kisebb-nagyobb reláció szerint. Azt tudjuk, hogy egy ilyen rendezés esetén a csupa nullát tartalmazó vektor lenne a legkisebb, és a csupa egyest tartalmazó a legnagyobb, de nem mondható semmi a többiről. A problémát az sem oldaná meg, ha a modellt úgy egyszerűsíténénk, hogy nem valószínűségekről beszélünk, hanem csupán a 0 és 1 értékeket engednénk meg aszerint, hogy az adott személy meg tudja-e oldani az adott feladattípust vagy sem.

A probléma lényegét ebben a 0-vagy-1 alapú egyszerűsített modellben a 2. ábra illusztrálja. Azon feladattípusok halmazait mutatja, amely feladattípusokat az „A”, illetve a „B” személy meg tud oldani.



2. ábra

Az eltérő összetételű képességállapotok rendezési problémáját illusztráló Venn-diagram. Az összehasonlítást azon feladattípusok nehezítik, amelyeket csak az egyik, vagy csak a másik képességállapothoz tartoznak (saját szerkesztés).

Általános esetben a két halmazt tekintve (2. ábra) van egy közös rész, illetve egy-egy olyan rész is, ahol csak az egyik, illetve csak a másik személy által megoldható feladattípusok szerepelnek. A gondot ez utóbbi két rész okozza, mivel az itt szereplő elemek összehasonlítása nehezen megoldható. Mi alapján mondható, hogy az egyik többet ér, mint a másik? A legegyszerűbb módszerként az kínálkozik, hogy a feladattípusok számát vesszük alapul, és akinek több van, az lenne előrébb. Ez a megoldás itt nem megalapozott, mivel a feladatok nagyon különbözőek lehetnek. Lehet, hogy az egyik személy jóval kevesebb feladattípust képes megoldani, de az valamilyen szempontból mégis sokkal többet ér. A két eltérő képességszerkezetet úgy kellene összehasonlítani, hogy a rendezési reláció valamilyen releváns szempontból megalapozható legyen. Ameddig ezt a problémát nem oldjuk meg, nem értelmezhető (és így nem is mérhető) a képesség fejlettsége, hiszen a fejlettség rendezést jelent.

Ugyanakkor elképzelhető olyan speciális eset, amikor a „B” halmaz tartalmazza az „A” halmazt. Gyakori lehet ez akkor, ha a két halmaz egyetlen adott személy két különböző időpontban rögzített képességállapotát tükrözi. Ez az eset pedagógiai szempontból nagyon fontos. Éppen ezt várjuk a tanulás-gyakorlás folyamatában: növekszik a megoldott-elsajátított feladattípusok száma, és (aktív szinten tartást feltételezve) nem csökken a már birtokoltak száma felejtés következtében. Az eddigiek alapján, bár két különböző személy képességállapota között általános esetben nem tudunk rendezési relációt definiálni, egy meghatározott személy képességének fejlődése viszont jól értelmezhető és modellezhető az egymást tartalmazó halmazok sorozata által, feltéve, hogy a felejtés kizárható az adott képességterületen. Ezen az úton tovább haladva ordinális skála konstruálható. Ha a tanulás mellett a felejtés is lényeges szerepet játszik, akkor egyazon személy különböző időpontban rögzített képességállapotai között olyan viszonyok alakulnak ki, amelyet a 2. ábra tükröz, vagyis a rendezés leírt módja ekkor már nem lesz lehetséges.

Ideális és reális teszt

Mindezek után nézzük meg, milyen további lehetőségek merülnek fel arra, hogy a képességállapotokhoz számokat rendeljünk úgy, hogy a hozzárendelés legalább ordinális skálát eredményezzen. A fentiek alapján egy *ideális teszt* úgy nézne ki, hogy minden feladattípusból elegendő számú feladatot tartalmazna ahhoz, hogy az adott feladattípushoz tartozó megoldási valószínűsége elfogadható becslést kapjunk. Az ideális tesztnek tehát

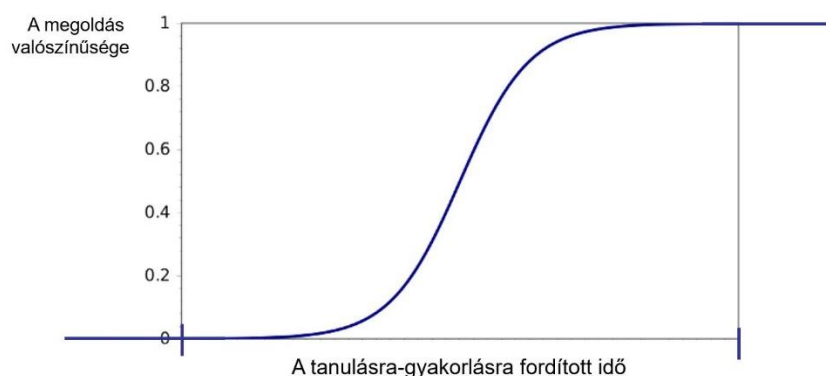
nagyon sok feladattípust, illetve ezekhez kapcsolódóan kezelhetetlenül sok feladatot kellene tartalmaznia. Fontos kiemelni, hogy a képesség jellegétől függően itt egyébként a legkülönbözőbb gyakorlati feladatok is előfordulhatnak, sőt akár dominánssá is válhatnak.

A tesztfeladatok száma csökkenthető, ha a szomszédos hasonló feladattípusokat összevonjuk egy-egy *feladatosztályba*. Ez elfogadható kompromisszum lehet abban a gyakorlati esetben, amikor a hasonló feladattípusok megoldási valószínűsége hasonló nagyságú. Az összevonásokat tovább folytatva a feladattér egyre durvább felosztását kapjuk, ahol már olyan feladattípusok kerülnek egy-egy közös osztályba, amelyek megoldási valószínűsége nagyobb mértékben eltérhet egymástól. Fontos, hogy az összevonások eredményeképpen létrejövő feladatosztályok a feladattípusokat, valamint a bennük foglalt feladatokat lehetőleg kiegyensúlyozottan képviseljék. Ne alakuljanak ki olyan indokolatlan különbségek a feladatosztályok terjedelmét illetően, amelyek torzíthatják az eredményt. Mire gondolunk itt? Egy hasonlattal illusztráljuk. Ha a szavazóközeteket úgy állapítanák meg, hogy mindegyik szavazóközet egy négyzetkilométer terjedelmű lenne, akkor az a városlakók szavazatát erősen leértékelné azokéval szemben, akik vidéken, elszórt kistelepléseken, vagy esetleg éppen tanyán laknak. Szerencsére ez nem így van. A szavazóközetek kialakításakor arra törekednek, hogy mindegyikben megközelítőleg azonos számú szavazópolgár éljen. A mennyiségi értelemben vett arányosság, kiegyensúlyozottság a feladatosztályok kialakításánál is fontos szempont.

A *reális teszt* gyakorlati okokból nem tartalmazhat túlságosan sok feladatot. Ez nem gond akkor, ha a képességnek csupán szűk területét vizsgáljuk. Viszont a reális teszt általános esetben a feladattér nagyvonalú felosztására épül, jelentősen csökkentve összevonásokkal a feladatosztályok számát. A reális tesztre jellemző további súlyos kompromisszum lehet az, hogy minden feladatosztályból csak nagyon kevés, esetleg csak egy-egy feladatot választunk. Az utóbbi esetben az a furcsa dolog történik, hogy egy feladat megoldásából igyekszünk következtetni a feladatosztály valószínűségi látens megoldási valószínűségére. Az egyetlen kiválasztott feladat megoldásának relatív gyakorisága viszont csak kétértékű, 0 vagy 1 lehet egy adott tanulónál. A nulla jól becsüli a feladatosztály látens megoldási valószínűségét, ha az kicsi, és az 1 akkor, ha az nagy, ám a közbülső esetekben egyik érték sem tekinthető jó becslésnek. Két enyhítő körülményre itt azért hivatkozhatunk.

Az egyik enyhítő körülmény egy lehetőség, ami sok esetben megközelítően fennállhat. A témánk szempontjából lényeges szerepet játszik annak a függvénygörbének az alakja, amely egy adott feladattípus megoldási valószínűségének változását írja le a tanulási-gyakorlási idő függvényében. Ha a megoldási valószínűség kezdeti nulla értékéből indulunk ki, előfordulhat, hogy a tanulás-gyakorlás ezt az értéket kezdetben még nem vagy csak nagyon lassan növeli. Tanulja, de még nem tudja. Ezt az előbb stagnáló, majd csak nagyon lassan emelkedő szakaszt egy ponton túl egy meredeken emelkedő szakasz válthatja fel. Itt történik valami, megértés, fogalmi átrendeződés, a megfelelő fogás megtalálása – valami, aminek kívülről jól észrevehető hatása van. A meredekebb szakasz azt jelenti, hogy nekilendül a tanuló, és rövid időn belül a megoldási valószínűség magasabb értéket ér el. Az ezt követő harmadik szakaszban már csak a begyakorlás, tökéletesítés, rögzítés következik, a valószínűség görbéje ismét már csak lassan emelkedve éri el a maximumot. Az így kialakuló ferde S betűre emlékeztető görbét a 3. ábra mutatja. Sok természeti folyamatra jellemző ez a dinamika. Azokban az esetekben, amikor a tanulási folyamat ehhez

hasznos szigmoid függvénnyel írható le (l. pl. Molnár & Csapó, 2003; Leibowitz, Baum, Enden, & Karniel, 2010), és a görbe középső szakasza eléggé meredek, egy adott személyhez tartozó érték az idő nagyobb részében valamelyik végponthoz közeli tartományban lehet (figyelemmel a tanulást megelőző és az azt követő időszakokra is). Ilyen esetekben viszonylag kisebb tehát annak a valószínűsége, hogy a 0 vagy 1 mért érték ne lenne tekinthető a látens valószínűség elég jó (elfogadható) becslésének.



3. ábra

Reális teszt esetén a becslés pontosságát növeli, ha az adott feladattípus megoldási valószínűsége viszonylag meredek S alakú függvényt közelítő módon növekszik a tanulásra-gyakorlásra fordított idő függvényében (saját szerkesztés)

A másik enyhítő körülmény statisztikai jellegű. A 0 és 1 mért értékek a látens valószínűség elfogadható becslésének tekinthetők akkor, amikor a látens valószínűség nagyon kicsi vagy nagyon nagy, és nem tekinthetők ilyenek, amikor nem ez a helyzet. Az utóbbi esetben a vektor mért értékei tehát túl kicsik vagy túl nagyok. Ha viszont az itt jelentkező hibák (eltérések) véletlenszerűek, és a hiba várható értéke nulla (nincs szisztematikus hiba), továbbá ha sok eltérés van, a nagy számok törvénye alapján fellép az a tendencia, hogy a mért értékekből számolt átlag már jobban közelíti a látens valószínűségek átlagát, mert az ellentétes irányú véletlenszerű eltérések a számítás során mintegy kioltják egymást.

De valójában mit is mutat az értékek átlaga? Ha a mért 0 és 1 értékek feladatosztályok látens megoldási valószínűségeinek becslései, vagyis maguk is valószínűségek, akkor ezekből valóban számolható átlag. Ez az átlag nem jellemzi az összetett képességet kielégítő módon, de nem is független teljesen attól. Valójában azt becsüli, hogy mekkora valószínűséggel old meg az adott személy egy véletlenszerűen kiválasztott feladatot akkor, ha minden feladatosztály egyforma eséllyel vesz részt a kiválasztásban. A feladatosztályok, illetve a bennük foglalt feladattípusok között itt nincs semmilyen különbségtétel. Az így kalkulált átlag viszonylag jól orientál a két szélső érték (0 vagy 1) közelében, és rendsze-

rint még akkor is, ha két személy képességállapota, vagy ugyanazon személy két képességállapota között extrém nagy különbséget jelez. Más esetekben nem megfelelő, mert figyelmen kívül hagyja a feladatosztályok (vagy feladattípusok) minőségi különbségeit.

Feladattípusok súlyozása

Az egyes feladattípusok minőségi különbségeit gyakran a feladattípusok súlyozásával igyekeznek kifejezni. Itt két kérdés merül fel: (1) mi lehet a súlyozás elfogadható elméleti alapja, és (2) az előző válasz alapján konkrétan miként határozzuk meg a súlyt.

Kézenfekvőnek látszik a feladattípusok nehézségét alapul venni, arra építve, hogy a tanulók milyen valószínűséggel képesek megoldani az adott feladatot. A feltevés az, hogy a nehezebb feladatot kevesebben oldják meg. Ez a logikus gondolat a már említett Rasch-modell egyik alappillére (Molnár, 2006, p. 106). Összetett képességeknél viszont felmerülhet egy probléma: ez alapján előnybe kerülhet az, akinek képességszerkezete erősen eltér a többiekétől, vagy fordítva, hátrányba az, aki a többségi csoportba kerül. A feladat nehézségének fenti mérése túlértékelheti a többségtől eltérő képességszerkezetet.

A nehézség szerinti súlyozás megoldható szakértői vélemények alapján is, de ez nagyobb ráfordítást igényel, és kevésbé egzakt. Adott kontextusban maga a nehézség is lehet többdimenziós tulajdonság. Elvileg felmerülhetnek még további súlyozási szempontok is. Ilyen például az adott feladattípus fontossága a további tanulmányok szempontjából; vagy éppen hasznossága a gyakorlatban, szakmai megítélés alapján. Bármennyire is érdekes lenne ez utóbbi szempont, a hasznosság sokszor nehezen megítélhető, mivel erősen függhet az adott helyzettől is, hogy mikor mi mennyire hasznos. További súlyozási szempont lehet az adott feladattípus szokásos gyakorisága a mindennapi gyakorlati helyzetekben. Ennek a megközelítésnek praktikus jelentősége lehet: néha célszerűbb a gyakran felmerülő feladatok megoldását lehetőleg mielőbb megtanulni, ez munkahelyen munkaszervezési előnyökkel is járhat. A gyakoriság és a hasznosság akár kombinálható is egymással, így gyakorlatorientált súlyozást kaphatunk.

Ha sikerült megtalálni a súlyozási módszert, akkor visszatérhetünk a 2. ábra kapcsán kifejtett rendezési problémára. Feltéve, hogy a feladattípusok kialakítása mennyiségi értelemben kiegyensúlyozottan történt, vagyis nincsenek feltűnő aránytalanságok ezek terjedelmét illetően (l. a szavazókörzetek példáját), akkor a súlyozás megoldhatja a rendezési problémát is. Az lesz előrébb a sorban, akinél a vektor összes valószínűségének súlyozott átlaga nagyobb értéket mutat. Az elméleti modellhez hasonlóan a tesztek esetében is a feladatmegoldások súlyozott átlaga határozza meg a sorrendet – azt, hogy minek a sorrendjét, érdemes ismét meggondolni. Egy testben egy meghatározott súlyozási eljárás esetén a súlyozott átlag annak valószínűségét becsüli, hogy az adott személy képes megoldani egy véletlenszerűen kiválasztott feladatot a testbe bevont feladattípusok (feladatosztályok) közül, feltéve, hogy a kiválasztás valószínűsége a súlyok arányát tükrözi. Az ilyen típusú súlyozott átlagok az összetett képességek fejlettségét elfogadható módon jellemzik a két végpont (a 0 és az 1) közelében, illetve rendszerint extrém nagy eltérések esetén is, azonban egyéb esetekben sokkal kevésbé. Az egyik probléma az, hogy nem mindig könnyű kellően releváns és elméletileg megalapozott súlyozási szempontot találni, másrészt a súlyozás tényleges megvalósítása gyakran erősen szubjektív. Az így konstruált

skála jó esetben olyasmit mér, ami bár nem független az összetett képesség fejlettségétől (ha ilyen létezik), de azt nem jellemzi megfelelően. Ha jellemezné is, azt akkor is csak a választott súlyozási szempont perspektívájából tenné.

Ráfordításalapú mérés

Beszélhetünk-e fejlettségről, lehetséges-e fejlettség egy összetett képesség esetében? A gyakorlati tapasztalat alapján hajlamosak vagyunk erre igenlő választ adni, hiszen hogyan fejleszthetnénk a képességeket, ha fejlettségük nem is létezik. Kitérő válasz lehet, hogy nem a képességet fejlesztjük, csak egy-egy lehatárolt dimenzióját. Egy pedagógus nem biztos, hogy elfogadna egy ilyen választ. Ő vélhetően az összetett képességet is, a maga teljességében is, fejleszteni kívánja – és valószínű, meg is teszi. Szigorúbb értelemben azonban csak akkor beszélhetünk fejlettségről egy összetett képesség esetében, ha rendezni tudjuk a különböző lehetséges képességállapotokat. A leírt súlyozási technikák alapvető hiányosságai miatt a rendezés kérdésére még nem adtak kielégítő választ.

A továbbiakban egy újszerű eljárást ismertetünk, ami a képességállapotokhoz olyan mértéket rendel, amely a képességet szigorúan szűk pedagógiai szempontból jellemzi, és megoldja a rendezés problémáját. A feladatorientált megközelítésre épülő eljárás nem magát az összetett képességet, hanem annak csupán egyik lényeges tulajdonságát méri. Ennél többet egyetlen számtól vélhetően nem is várhatunk. Ha egy fizikustól azt kérjük, hogy mérjen meg egy szilárd testet, megkérdezi, hogy a test melyik tulajdonságára gondolunk. A tömegére? A hőmérsékletére? A sűrűségére? A tulajdonságok sora még hosszan folytatható. Az adott test eme tulajdonságai jellemezhetők egy-egy számmal és a hozzá tartozó mértékegységgel, de ezek egyike sem jellemzi a testet a maga teljességében. Hasonlóan nem várhatjuk azt, hogy egy összetett képesség egyetlen skalár mennyiséggel vagy pontszámmal kielégítően (teljes körűen) leírható legyen.

A javasolt eljárás során egy feladattér minden feladattípusához hozzárendeljük azt a munkaórában mért időráfordítást, amely a feladatmegoldás elsajátításához átlagosan szükséges, ideértve a tanulás minden szokásos formáját és módját, valamint a feladatmegoldás gyakorlását is más hasonló tanulók nagyobb létszámú csoportjának adatai alapján. A feladatmegoldás elsajátítása azt jelenti, hogy a tanuló az adott feladattípus feladatait (közel) 100% valószínűséggel képes megoldani. Problémának tűnhet, hogy ez az időráfordítás erősen függ – egyebek közt – a személytől, a körülményektől, a tanulás módszereitől és az addig megtett individuális tanulási úttól (vagyis a feladattér mely más feladattípusait tanulta meg korábban). Hasonló körülmények között, hasonló előismeretekkel és előzményekkel rendelkező, hasonló tanulók hasonló módszereket alkalmazó csoportjában kisebb eltérések adódhatnak. Most azonban ez nem is annyira fontos, mivel ezt a mennyiséget nem egy adott tanuló várható időráfordításának előrejelzésére szeretnénk használni, hanem arra, hogy az egyes feladattípusokhoz egy pedagógiai szempontból értelmezhető mértéket rendeljünk. Ez pedig ilyen. Az, hogy egy tanulócsoporthoz tagjai adott szintről indulva átlagosan mennyi ideig tanulnak meg valamit, pedagógiai szempontból nagyon is releváns

és értelmezhető, empirikusan vizsgálható. Igen fontos adat ez a tanulás tervezése, irányítása és támogatása szempontjából. Ráadásul ezek az időmennyiségek *aránykálát* alkotnak, tág teret nyitva a legkülönbözőbb statisztikai feldolgozások számára.

Érdeemes megjegyezni, hogy a ráfordításalapú megközelítés nagyon hasonló logikát követ ahhoz, mint ahogyan a projektmenedzsmentben az elvégzett munkát mérik (PMI, 2019). A konstrukció ott a megtermelt érték (*earned value*) fogalmára épül, melyhez különféle számítási eljárások kapcsolódnak. Az analógia alapján a javasolt mérték megtanult értéknek (*learned value*) lenne nevezhető, ha ez nem ütközne a fogalom már meglévő, más értelmű használatával.

Ahhoz, hogy az eljárás működjön, választ kell adni még egy kérdésre. Feladatorientált elméleti modellünkben egy nagy elemszámú vektor szerepel. A vektort alkotó számok az egyes feladattípusok (vagy feladatosztályok) megoldási valószínűségei. Minden 0 és 1 közé eső szám itt valamilyen közbülső tanulási állapotra utal. Milyen munkaóra mennyiséget rendeljünk ezekhez a közbülső állapotokhoz, ha már ismert a teljes megtanuláshoz szükséges időmennyiség?

Többféle megoldás adódik. Az egyik a lineáris interpoláció: azt mondjuk, hogy nem ismerjük pontosan az időráfordítás és a megoldási valószínűség közötti kapcsolatot leíró függvényt (amely lehet S alakú, de más módon is), de akármilyen is ez, egyenes aránysággal közelítjük. A gyakorlatban ez azt jelenti, hogy ha az adott feladattípus teljes megtanulásához szükséges idő „*t*”, és az adott tanuló ezt a feladattípust adott időpontban aktuálisan „*p*” megoldási valószínűséggel képes megoldani, akkor a közbülső állapothoz rendelt érték a két szám szorzata lesz. A másik lehetőség az, hogy adatokat gyűjtünk a „*p*” valószínűség néhány kitüntetett közbülső értékéhez, és így tudni fogjuk azt, hogy átlagosan mennyi idő kell például a 0,25, a 0,5 és a 0,75 valószínűségi szint eléréséhez. Az ezek közötti köztes értékeket pedig arányosítva számoljuk (pontosított lineáris interpoláció). A harmadik megoldás az, hogy tesztsorozattal az időráfordítás és a megoldási valószínűség közötti kapcsolatról sok adatot gyűjtünk, és meghatározzuk az ezekre legjobban illeszkedő függvényt, ami lehet akár olyan is, amely a 3. ábrán szerepel. Végül a függvény elméleti megfontolások alapján is definiálható. Ha megvan a függvény, akkor ez alapján bármilyen „*p*” értékhez meghatározható a hozzá tartozó időráfordítás. Hasonló feladattípusok esetén azonos vagy nagyon hasonló függvények várhatóak.

Mindezek alapján egy adott képességállapothoz tartozó vektor összes feladattípusához megállapítható a hozzátartozó időráfordítás egy adott referenciacsoporthoz alapul véve. Az új fogalomként szereplő „*V*” *képességvolumen* az adott képességállapotot (annak „terjedelmét”) a képességállapot eléréséhez szükséges összes tanulási-gyakorlási időráfordítás mértékével jellemzi. Az alábbi (1) formula az általános esetet mutatja, a (2) pedig azt, amikor a „*g*” függvényt lineáris interpolációval becsüljük.

$$(1) V = \sum g_i(p_i) * t_i \quad (2) V = \sum p_i * t_i$$

ahol g_i = az *i*-dik feladattípus valószínűség-időráfordítás függvénye,
 p_i = az *i*-dik feladattípus aktuális megoldási valószínűsége, és
 t_i = az *i*-dik feladattípus (közel) 100%-os megoldásának megtanuláshoz szükséges szokásos időráfordítás a kiválasztott referenciacsoporthoz.

Amennyiben egy teszt feladatait a tanulási-gyakorlási időráfordítással arányosan súlyozzák, a teszt százalékos pontszáma azt becsüli, hogy a tesztben reprezentált feladattípusok tekintetében a tanuló hol tart a tanulási folyamatban egy adott referenciacsoport jellemzőit (pl. előismeretet és tanulási hatékonyságot) alapul véve. Az eljárás a már elsajátított területeket és a még elsajátítandó területeket is ezek szokásos tanulási-gyakorlási időráfordításával méri. Aki hátrébb van, az előtt több munka áll. Más kérdés az, hogy az egyéni különbségek, a tanulási módszerek és körülmények függvényében a még hátralévő tanulási idő egyes konkrét esetekben akár jelentősen is eltérhet attól, amit a referenciacsoport adatai alapján határoztak meg. Az eljárás alkalmas arra, hogy egy adott tanuló képességállapotát és annak változását a referenciacsoport adataihoz viszonyítva jellemezze. Alkalmas lehet arra is, hogy a referenciacsoporthoz hasonló tanulócsoportokra vonatkozóan előrejelzéseket alapozzon meg. Azzal, hogy az eljárás a figyelmet a tanulási-gyakorlási folyamatok időbeliségére irányítja, ezek átgondoltabb tervezését és megvalósítását is segítheti, figyelembe véve például a tanulástámogatás és a differenciálás szempontjait is.

A javasolt eljárás nem az összetett képesség látens dimenzionális szerkezetének feltárására épül. A fentiek megerősítik azt a feltételezést, hogy egy összetett képesség állapota nem jellemezhető egyetlen számmal (skalármennyiséggel) úgy, hogy az megfeleljen a reprezentációs méréselméletnek, és ordinális skála, jobb esetben intervallumskála jöjjön létre, ha az eljárástól azt várjuk, hogy az összetett képesség (tartalmi és minőségi) sokféleségét is kellőképpen figyelembe vegye. Amennyiben mégis megkíséreljük az egyetlen számra épülő jellemzést, akkor vélhetően olyan mérési modellhez juthatunk, mint amelyet a képességvolumen mérésének javasolt megoldása képvisel (és amely ezt a sokféleséget nem tükrözi).

A dimenzionális és a feladatorientált megközelítés fentiekben leírt különbségeit az 1. táblázat foglalja össze. A feladatorientált megközelítés a dimenzionális szemlélet általánosításának is tekinthető, mivel a feladatosztályok segítségével dimenziók is definiálhatók, ha ez megfelelő elméleti modell alapján megtehető.

A felsőoktatásban és más területeken is használatos kreditpontok a javasolt eljáráshoz hasonló alap gondolatra épülnek, de (rendszerint) nélkülözik az időráfordítások részletes és pontos mérését. Az eljáráshoz szükség van a feladattípusok szokásos tanulási-gyakorlási idejére a kiválasztott referenciacsoportban. Erre nézve adatok, illetve becslések gyűjthetők például megkérdezés, megbeszélés, megfigyelés és dokumentumelemzés útján, továbbá digitális környezetben történő tanulás esetén az automatizált adatkezelés és adatelemzés új lehetőségeit is hasznosítva (vö. *big data*, *data mining*, *learning analytics* – l. pl. Aldowah, Al-Samarraie, & Fauzy, 2019; Papamitsiou & Economides, 2014), figyelemmel a személyes adatok kezelésének szabályaira is. Az adatbázis (legalább) időszakos frissítése javasolt.

Különösen a szakképzés, a felsőoktatás és a felnőttoktatás területén az időráfordítás mellett egyéb jelentős költségtartalmú ráfordítások is felmerülhetnek. A bemutatott időráfordításalapú megközelítés átalakítható *költségalapúvá* is. Ebben az esetben azt kell meghatározni, hogy az egyes időráfordításoknak milyen költségek feleltethetők meg. Például alapul vehetjük azt az elmaradt bevételt, amire az adott személy hasonló jellegű és hasonló terhelést jelentő munkával szert tett volna, ha tanulás helyett inkább dolgozik. Vagy alapul vehető az a bevétel, amit az adott idő alatt saját szakmájában, munkakörében keresett

volna. A teljes költségráfordítás meghatározásakor ehhez még hozzá kell adni a tanuló további felmerülő közvetlen költségeit (pl. tandíj, könyvek, eszközök, utazás). Ez a megközelítés a képességfejlesztés oktatás-gazdaságtani elemzéséhez járulhat hozzá, adott esetben figyelembe véve az oktatói oldal költségeit is.

1. táblázat. Az összetett képesség kétféle megközelítése (saját szerkesztés)

Összehasonlítási szempont	Dimenzionális megközelítés	Feladatorientált megközelítés
A „ K ” összetett képesség értelmezése	$K = (D_1, D_2, \dots)$ D_x : komponens dimenziók	$K = (F_1, F_2, \dots)$ F_x : komponens feladatosztályok
A komponensek értelmezése	A diszjunkt és homogén D_x dimenziók meghatározásával a képesség szerkezetének elméleti modellben értelmezhető leírására törekszik.	Az F_x feladatosztályok a képesség külső szemléletű, gyakorlati jellegű (praktikus) leírását adják elméleti modell igénye nélkül.
Mérési szint	A D_x dimenziók ordinális szinten mérhetők. A modern testelmélet alapján az egyes dimenziók meghatározott feltételek mellett intervallumskálán is mérhetővé tehetők. A képességállapotok ezt megalapozó elméleti modell hiányában nem rendezhetők.	A dimenzionális szerkezet feltárása nélkül arányskála szinten mérhető a ráfordításalapú <i>képességvolumen</i> véges sok feladatosztály esetén adott referenciacsoportra vonatkozóan. Az így mért képességállapotok rendezhetők.
Magyarázóerő	A dimenzionális struktúra elméletileg megalapozott felvilágosítást nyújthat a képesség feltételezett szerkezetéről, segítve annak mélyebb megértését.	A feladatstruktúra gyakorlati (praktikus) felvilágosítást nyújthat a képesség definiált terjedelméről, segítve annak hatékonyabb elsajátítását.

A képességelem fontos jellemzője hasznossága mellett az is, hogy mennyire könnyű vagy nehéz a megtanulása, vagyis személyes reprodukálása. A ráfordításalapú súlyozás az adott képességelem szokásos társadalmi költségét fejezi ki, annak becslésének tekinthető. A költségesebb képességelemből rendszerint kevesebb áll rendelkezésre. Ami nehezebben megtanulható, az rendszerint ritkább is, és általában jobban becsülendő (kivételek előfordulhatnak). A képességelemek a tanulás-gyakorlás személyes műhelyében létrehozott sajátos termékek. A ráfordítás a termék költsége. A képességelemek kulturális értékükön és saját célú használatukon túl piaci termékek is válhatnak. Értéküket ebben az esetben alapvetően a kereslet és a kínálat határozza meg. Konkrét szakmai képességek esetében, megfelelő adatok birtokában a költségek és a hasznonelemek alapján megtérülési számítások is végezhetők, melyek az esetek többségében remélhetőleg azt támasztják majd alá, hogy a tanulás az egyik legjobb befektetés.

Irodalom

- Aldowah, H., Al-Samarraie, H., & Fauzy, W. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–49. doi: [10.1016/j.tele.2019.01.007](https://doi.org/10.1016/j.tele.2019.01.007)
- An, X., & Yung, Y. (2014). Item Response Theory. Paper SAS364-2014. In *SAS Global Forum, Proceedings 14*, Cary: SAS Institute Inc. Retrieved from <https://support.sas.com/resources/papers/proceedings14/SAS364-2014.pdf>
- Bouyssou, D., & Pirlot, M. (2005). Conjoint measurement tools for MCDM. In J. Figueira, S. Greco, & M. Ehrgott (Eds.), *Multiple criteria decision analysis: State of the art surveys* (pp. 73–132). Springer.
- Briggs, D. C. (2013). Measuring growth with vertical scales. *Journal of Educational Measurement*, 50(2), 204–226. doi: [10.1111/jedm.12011](https://doi.org/10.1111/jedm.12011)
- Cervone, D., & Caldwell, T. L. (2008). From measurement theory to psychological theory, in reverse. *Measurement: Interdisciplinary Research and Perspectives*, 6(1–2), 84–88. doi: [10.1080/15366360802035539](https://doi.org/10.1080/15366360802035539)
- Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, 48(6). doi: [10.18637/jss.v048.i06](https://doi.org/10.18637/jss.v048.i06)
- Csapó, B. (2003). *A képességek fejlődése és iskolai fejlesztése*. Budapest: Akadémiai Kiadó.
- Csapó, B. (2008). A tanulás dimenziói és a tudás szerveződése. *Educatio*, 17(2), 207–217. Retrieved from <http://publicatio.bibl.u-szeged.hu/6099/>
- DeMars, C. E. (2016). Partially compensatory multidimensional item response theory models: Two alternate model forms. *Educational and Psychological Measurement*, 76(2), 231–257. doi: [10.1177/0013164415589595](https://doi.org/10.1177/0013164415589595)
- Doignon, J., & Falmagne, J. (2016). Knowledge spaces and learning spaces. In W. Batchelder, H. Colonius, E. Dzhafarov, & J. Myung (Eds.), *New handbook of mathematical psychology* (Cambridge Handbooks in Psychology) (pp. 274–321). Cambridge: Cambridge University Press. doi: [10.1017/9781139245913.006](https://doi.org/10.1017/9781139245913.006)
- Fatalin, L. (2008). *Hierarchikus fogalmi struktúrák vizsgálata gráfokkal*. Doktori (PhD) értekezés. Debreceni Egyetem. Retrieved from https://dea.lib.unideb.hu/dea/bitstream/handle/2437/85019/ertekezes_magyar.pdf
- Feinberg, R. A., & Rubright, J. D. (2016). Conducting simulation studies in psychometrics. *Educational Measurement: Issues & Practice*, 35(2), 36–49. doi: [10.1111/emip.12111](https://doi.org/10.1111/emip.12111)
- Feuerstahler, L., & Wilson, M. (2019). Scale alignment in between-item multidimensional Rasch models. *Journal of Educational Measurement*, 56(2), 280–301. doi: [10.1111/jedm.12209](https://doi.org/10.1111/jedm.12209)
- Granberg-Rademacker, J. S. (2010). An algorithm for converting ordinal scale measurement data to interval/ratio scale. *Educational and Psychological Measurement*, 70(1), 74–90. doi: [10.1177/0013164409344532](https://doi.org/10.1177/0013164409344532)
- Hand, D. J. (1996). Statistics and the theory of measurement. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159(3), 445–492. doi: [10.2307/2983326](https://doi.org/10.2307/2983326)
- Hartig, J., & Höhler, J. (2009). Multidimensional IRT models for the assessment of competencies. *Studies In Educational Evaluation*, 35(2/3), 57–63. doi: [10.1016/j.stueduc.2009.10.002](https://doi.org/10.1016/j.stueduc.2009.10.002)
- Heilmann, C. (2015). A new interpretation of the representational theory of measurement. *Philosophy of Science*, 82(5), 787–797. doi: [10.1086/683280](https://doi.org/10.1086/683280)
- Immekus, J. C., Snyder, K. E., & Ralston, P. A. (2019). Multidimensional item response theory for factor structure assessment in educational psychology research. *Frontiers in Education*, 4. doi: [10.3389/educ.2019.00045](https://doi.org/10.3389/educ.2019.00045)

- Kehl, D. (2011). Skálák és statisztikák: A méréselméletéről és történetéről. *Statisztikai Szemle*, 89(10–11), 1057–1080. Retrieved from http://www.ksh.hu/statszemle_archive/2011/2011_10-11/2011_10-11_1057.pdf
- Leibowitz, N., Baum, B., Enden, G., & Karniel, A. (2010). The exponential learning equation as a function of successful trials results in sigmoid performance. *Journal of Mathematical Psychology*, 54(3), 338–340. doi: 10.1016/j.jmp.2010.01.006
- Li, Y., Jiao, H., & Lissitz, R. W. (2012). Applying multidimensional item response theory models in validating test dimensionality: An example of k–12 large-scale science assessment. *Journal of Applied Testing Technology*, 13(2), 1–27. Retrieved from <http://www.jattjournal.com/index.php/atp/article/view/48367>
- Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept (Ideas in context)*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511490040
- Molnár, G., & Csapó, B. (2019). Making the psychological dimension of learning visible: Using technology-based assessment to monitor students' cognitive development. *Frontiers in Psychology*, 10:1368. doi: 10.3389/fpsyg.2019.01368
- Molnár, Gy. (2006). A Rasch-modell alkalmazása a társadalomtudományi kutatásokban. *Iskolakultúra*, 16(12), 99–113. Retrieved from <http://real.mtak.hu/id/eprint/57255>
- Molnár, Gy. (2013). *A Rasch-modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában*. Budapest: Gondolat.
- Molnár, Gy., & Csapó, B. (2003). A képességek fejlődésének logisztikus modellje. *Iskolakultúra*, 13(2), 57–69. Retrieved from <http://publicatio.bibl.u-szeged.hu/11277/>
- Nahalka, I. (2015). Tanulói teljesítménymérések alkalmazhatósága a neveléstudományban. In K. Széll (Ed.), *Mit mér a műszer? A tanulói teljesítménymérések alkalmazhatóságáról* (pp. 23–36). Budapest: Oktatókutató és Fejlesztő Intézet. Retrieved from <https://ofi.oh.gov.hu/kiadvany/mit-mer-muszer>
- Nahalka, I. (2018). *Ellentmondások a pedagógiai mérés és értékelés elméleteiben*. Habilitációs értekezés, Eszterházy Károly Egyetem. doi: 10.15773/EKE.HABIL.2018.008
- Narens, L. (2002). A meaningful justification for the representational theory of measurement. *Journal of Mathematical Psychology*, 46(6), 746–768. doi: 10.1006/jmps.2002.1428
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15, 625–632. doi: 10.1007/s10459-010-9222-y
- Pálvölgyi, L. (1981). *A modellezés lehetőségeiről a pedagógiában*. Budapest: Akadémiai Kiadó.
- Papamitsiou, Z., & Economides, A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49–64. Retrieved from <https://www.jstor.org/stable/jeductechsoci.17.4.49>
- PMI (2019). *The standard for earned value management*. Newtown Square, USA: Project Management Institute.
- Robusto, E., Stefanutti, L., & Anselmi, P. (2010). The Gain-Loss model: A probabilistic skill multimap model for assessing learning processes. *Journal of Educational Measurement*, 47(3), 373–394. doi: 10.1111/j.1745-3984.2010.00119.x
- Stevens, S. S. (1946). On the theory of scales of measurement. *American Association for the Advancement of Science*, 103(2684), 677–680. doi: 10.1126/science.103.2684.677
- Sulis, I., & Toland, M. D. (2016). Introduction to multilevel item response theory analysis: Descriptive and explanatory models. *The Journal of Early Adolescence*, 37(1), 85–128. doi: 10.1177/0272431616642328
- Svetina, D. (2013). Assessing dimensionality of noncompensatory multidimensional item response theory with complex structures. *Educational and Psychological Measurement*, 73(2), 312–338. doi: 10.1177/0013164412461353
- Takács, V. (2000). A Galois-gráfok pedagógiai alkalmazása. *Iskolakultúra-könyvek 6*. Szeged: *Iskolakultúra*. Retrieved from <http://misc.bibl.u-szeged.hu/45472/>

Összetett képességek mérése: dilemmák és megközelítések

- Tóth, Z. (2005). A tudásszerkezet és a tudás szerveződésének vizsgálata a tudástér-elmélet alapján. *Magyar Pedagógia*, 105(1), 59–82. Retrieved from http://www.magyarpedagogia.hu/?pid=10_item&cid=217
- Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory and Psychology*, 19(5), 579–599. doi: [10.1177/0959354309341926](https://doi.org/10.1177/0959354309341926)
- Trendler, G. (2019). Conjoint measurement undone. *Theory & Psychology*, 29(1), 100–128. doi: [10.1177/0959354318788729](https://doi.org/10.1177/0959354318788729)
- Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 47(1), 65–72. doi: [10.1080/00031305.1993.10475938](https://doi.org/10.1080/00031305.1993.10475938)
- Vessonen, E. (2020). The complementarity of psychometrics and the representational theory of measurement. *The British Journal for the Philosophy of Science*, 71(2), 415–442. doi: [10.1093/bjps/axy032](https://doi.org/10.1093/bjps/axy032)

Pálvölgyi Lajos

ABSTRACT

MEASURING COMPLEX SKILLS: DILEMMAS AND APPROACHES

Lajos Pálvölgyi

Measuring complex skills is a very important area of pedagogy raising several open questions. Many empirical data collections and student performance measurements lack a proper measurement theoretical foundation, and statistical methods are often used which would only be permissible for interval scales, but the conditions for this are not met. The dilemmas can be traced back to two partly interrelated basic questions: one is „what do we measure” and the other is „how do we measure”. This study reviews the difficulties associated with the second question and some of their possible solutions, also taking into account the results of multidimensional item response theory. Since development means ordering, the developmental state of a multidimensional skill should be characterized by a single number so that the procedure corresponds to the representational theory of measurement, and at least an ordinal scale is created. The proposed novel approach can help solve some measurement theory problems arising from the multidimensional nature of complex skills. The study seeks a solution in an alternative direction from the international mainstream, but retains its probabilistic view. It presents a task-oriented model as opposed to the usual dimensional approach. Skill is characterized by the tasks it allows to be solved. Central to the model is the concept of skill volume, which characterizes a given skill state with the amount of usual learning/practice time required to achieve it, based on the data of a given student reference group. The skill states measured in this way can be sorted; moreover, the effort-based skill volume becomes measurable at ratio scale level for a finite number of task classes, without the need to explore the latent dimensional structure of the complex skill. The skill volume can provide pedagogically important and well-usable information on the scope of a complex skill represented by a set of task classes. The procedure does not try to capture the complex skill itself in its completeness and qualitative diversity, but highlights one of its pedagogically relevant attributes, thus providing an opportunity for well-founded statistical analysis based on higher level scales.

Magyar Pedagógia, 120(2). 149–170. (2020)

DOI: 10.17670/MPed.2020.2.149

Levelezési cím / Address for correspondence: Pálvölgyi Lajos, ELTE Eötvös Loránd Tudományegyetem Neveléstudományi Intézet, H–1075 Budapest, Kazinczy u. 23–27.