

**Korom Erzsébet¹ – Pásztor Attila² –
Gyenes Tamás³ – B. Németh Mária²**

¹ Szegedi Tudományegyetem Oktatásmélet Tanszék

² MTA-SZTE Képességfejlesztés Kutatócsoport

³ Szegedi Tudományegyetem TTIK fizikatanár–matematikatanár MSc szak, hallgató

Kutatási készségek online mérése a 8–11. évfolyamon

A SAILS projekt a kutatási készségek ('inquiry skills') értékelésekor elsősorban a tanórai tevékenységekre koncentrált, és a kutatási készségek formatív értékelésének módjait, lehetőségeit helyezte előtérbe. A hazai munkacsoport a formatív értékelés mellett a kutatási készségek diagnosztikus értékelési lehetőségeit is vizsgálta azokra a tapasztalatokra építve, amelyek a tudás- és képességvizsgálatok, valamint a számítógép-alapú mérések területén az SZTE Oktatásméleti Kutatócsoport és az MTA-SZTE Képességfejlesztés Kutatócsoport munkái révén az elmúlt évtizedekben összegyűltek. Kidolgoztunk egy olyan természettudományos tesztet, amely számítógép-alapon működve méri négy kutatási készség (kutatási kérdés azonosítása, vizsgálatok tervezése, adatok értelmezése, következtetések levonása) fejlettségét. A tanulmányban a tesztfejlesztés tanulságait mutatjuk be egy gimnazisták körében 8–11. évfolyamon zajlott pilot mérés eredményeinek tükrében.

Bevezetés

A kutatási készségek fejlettségének diagnosztizálásra irányuló kutatásunk közvetlen előzményének tekinthető, hogy elkészítettük és bemértük a SAILS projektben részt vevő török kutatócsoport által kidolgozott 20 ítemes, papíralapú teszt online változatát. Az öt készséget (kutatási kérdés, hipotézis és változók azonosítása, adatok közlése és értelmezése) vizsgáló, készségenként négy többszörös választásos feladatot tartalmazó tesztet (Yalaki, Cakmakci, Yahsi, Gümüs, Gürel, Yüksel és Ince, 2014) 6. és 8. évfolyamos tanulók körében próbáltuk ki három iskola négy osztályában (Korom, B. Németh és Pásztor, 2015). Az eredmények jelezték, hogy a 6. évfolyamon a teszt nem mér megbízhatóan, ebben az életkorban a diákok még nem ismernek olyan fogalmakat, mint például kutatási kérdés, hipotézis, változó. A 8. évfolyamosok is elsősorban azokban a feladatokban teljesítettek jól, amelyek szerves részét képezik a magyar természet-tudományos oktatásnak (pl. adatok értelmezése táblázatok, ábrák alapján).

A kutatás folytatásaként – felhasználva a hazai online felmérések tapasztalatait (Molnár és Csapó, 2013; Molnár, 2015) és a természettudományos tudás diagnosztikus értékelésére (Csapó és Szabó, 2012) és online mérésére (Csapó, Korom és Molnár, 2015) kidolgozott tartalmi keretet – saját teszt fejlesztésébe kezdtünk. Mivel hazánkban nincs hagyománya a kutatási készségek vizsgálatának, nem állnak rendelkezésre feladatok, mérési tapasztalatok, célunk egy olyan teszt kidolgozása volt, amely tartalmilag és for-

mailag változatos feladatokkal, számítógép-alapon működve fedi le a természettudományos vizsgálódás, kutatás legfontosabb elemeit. A tanulmányban a tesztfelkészítés tanulságait és a gimnazisták körében zajlott pilot mérés eredményeit összegezzük.

Elméleti háttér

A természettudományos gondolkodás a gondolkodás specifikus típusa, olyan értelmi folyamatok összessége, amelyek természettudományos témákkal, problémákkal való foglalkozás során, illetve a tudományos megismerő tevékenység különböző formái között zajlanak (Kuhn, 2002; Dunbar és Fugelsang, 2005). A természettudományos gondolkodás részét képezi mindaz, amit Piaget (1964) formális műveleti gondolkodásnak vagy Lawson (1995) hipotetikus-deduktív gondolkodási mintázatoknak nevezett. A terület-általános gondolkodási képességek mellett a természettudományos gondolkodásban fontos szerepet játszanak a terület-specifikus tartalmi elemek is (Zimmerman, 2007). Mindkét tudásterület kutatása hosszú múltra tekint vissza. A tartalmi vagy fogalmi tudás vizsgálatában a tanulás, ismeretszerzés eredményén van a hangsúly, azt vizsgálják, hogyan értelmezik a tanulók a tudományos fogalmakat, jelenségeket. A gondolkodási képességek esetében a tudásszerzés folyamata, a kutatás, vizsgálódás, valamint a problémamegoldás és az ehhez szükséges készségek, képességek kerülnek a középpontba.

Számos elméleti keret és mérőeszköz született a természettudományos gondolkodáshoz szükséges gondolkodási képességek feltárására. Piaget (1964) munkáira alapozva kezdetben egyéni interjúkkal, később osztálytermi környezetben alkalmazható módszerekkel diagnosztizálták, hogy a gyerekek a kognitív fejlődés mely szakaszában vannak (lásd pl. Shayer, Adey és Wylam, 1981). A formális gondolkodás vizsgálatára több papír-ceruza teszt is készült. Ilyen például Lawson (1978) tesztje (Classroom Test of Scientific Reasoning, CTSR) és annak továbbfejlesztett változatai. A Tobin és Capie (1981) által kidolgozott Test of the Logical Thinking (TOLT) öt képességet mér: változók kontrollja, arányossági, valószínűségi, korrelatív és kombinatív gondolkodás. Raodranga (1991) Group Assessment of Logical Thinking (GALT) tesztje a konzervációt, arányossági gondolkodást, változók kontrollját, kombinatív, valószínűségi és korrelatív gondolkodást mérve különíti el a tanulókat a gondolkodás konkrét, illetve formális szintje szerint.

A természettudományos gondolkodás specifikus elemei a természettudományos tárgy vizsgálatához kötődő kutatási készségek. Jelentőségüket jelzi, hogy hangsúlyosak a természettudományos nevelés standardjaiban, valamint a nemzetközi vizsgálatokban is. A TIMSS 2011 vizsgálat elméleti kerete a kognitív műveletek (tudás, alkalmazás, gondolkodás) mellett külön felhívja a figyelmet a tudományos vizsgálódás fontosságára (Mullis, Martin, Ruddock, O'Sullivan és Preuschoff, 2009). A 4. és a 8. évfolyamon egyaránt elvárt, hogy a tanulóknak megfelelő ismereteik legyenek a tudomány működéséről, a tudományos kutatás természetéről, és rendelkezzenek a tudomány műveléséhez szükséges olyan alapvető készségekkel, mint a kutatási kérdések és hipotézisek megfogalmazása, vizsgálatok tervezése, adatok ábrázolása, adatok elemzése és értelmezése, következtetések levonása. A kutatási készségek mérése nem elkülönülten, hanem az egyes tartalmi területekbe (biológia, fizika, kémia és földrajz) ágyazottan történik.

A PISA vizsgálatokban (OECD, 2014) a természettudományos műveltség része a szaktudományi ismeretek mellett az episztemológiai (a tudományos vizsgálatokról és azok eredményeként létrejövő tudományos értelmezésekről való tudás) és a procedurális tudás, valamint a tudás alkalmazása különböző kontextusokban. A tudományos vizsgálódás mint procedurális tudás (a PISA kifejezésével kompetencia) van jelen, melynek három összetevője a tudományos kérdések azonosítása, a jelenségek tudományos magyarázata és a tudományos bizonyítékok használata.

A nemzetközi vizsgálatok mellett több átfogó, nemzeti felmérés is foglalkozik a kutatási készségekkel. Az amerikai természettudományos standardokban (NRC, 1996; NGSS, 2013) a természettudományos műveltség elemei között kiemelt szerepet kap a természettudományos vizsgálódás. Iskolai fejlesztéséhez részletes útmutatók, oktatási segédanyagok készültek (lásd pl. NRC, 2000). A kutatási készségek pontosabb meghatározásához, hierarchikus szerveződésük értelmezéséhez Wenning (2005, 2007) dolgozott ki elméleti keretet, és egy arra épülő tesztet (Scientific Inquiry Literacy Test, ScInqLit). A német ESNaS-projektben (Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level), követve a 2004-ben kidolgozott nemzeti standardokat, országos longitudinális vizsgálattal mérték a kutatási készségeket biológiai tartalmakon (Wellnitz, Hartmann és Mayer, 2010).

A természettudományos gondolkodás fejlesztésére irányuló kutatások részeként részletesebben is vizsgálták a gondolkodás néhány elemét. Például a változók kontrollját és annak összefüggését a feladatok, adatok típusával (Zhou, Han, Koenig, Raplinger, Pi, Li, Xiao, Fu és Bao, 2015), a tudomány természetének ismeretét, az elméletek megértését, kísérletek tervezését, eredmények interpretálását és azok összefüggését az intelligenciával, szövegértéssel, problémamegoldással, térbeli képességekkel (Mayer, Sodian, Koerber és Schwippert, 2014). Számos tapasztalat gyűlt össze a kutatásalapú tanulás hatékonyságának értékelési lehetőségeiről is (lásd pl. Gormally, Brickman, Hallar és Armstrong, 2009; Liu, Lee és Linn, 2010; Harlen, 2013).

A természettudományos gondolkodás tartalom-specifikus jellegére számos kutató felhívja a figyelmet, ami kihat a mérésre használt módszerekre is (Osborne, 2013). Kind (2013) olyan skálák alkalmazását javasolja, amelyek a természettudományos gondolkodást három tevékenységgel (hipotézisek megfogalmazása, vizsgálatok tervezése és megvalósítása, valamint a kutatási eredmények értelmezése) írják le, és figyelembe veszik a hozzájuk szükséges tartalmi, procedurális és episztemológiai tudást.

Az eddigi vizsgálatokban elsősorban papír-ceruza teszteket alkalmaztak. A technológiai alapú értékelésben rejülő előnyök kiaknázásával lehetőségünk van arra, hogy a mindennapi pedagógiai gyakorlatban is könnyen alkalmazható diagnosztikus tesztekkel vizsgáljuk a diákok természettudományos gondolkodását, kutatási készségeit, kövessük azok fejlődését. Ehhez kidolgoztuk a természettudományos gondolkodás online diagnosztikus mérésének elméleti hátterét (Adey és Csapó, 2012; Korom, Molnár és Csapó, 2015), és számos feladatot fejlesztettünk a kutatási készségek mérésére (Nagy L.-né, Korom, Pásztor, Veres és B. Németh, 2015).

A kutatási készségek vizsgálata

Az adatfelvétellel elsősorban arra kerestünk választ, hogy mérhető-e a számítógépes felületre készült feladatokkal, feladattípusokkal a kutatási készségek. Célunk az alkalmas módszerek megalapozása, megbízhatóan működő feladatok kidolgozása, feladatsorok szerkesztése volt, de a természettudományos teljesítményekkel összefüggésben is megfogalmaztunk kérdéseket. Választ kerestük arra, hogy megbízhatóan mér-e a teszt, illetve az egyes készségek működésére irányuló résztesztek. Vizsgáltuk, hogy van-e fejlődés a kutatási készségek működésében a középiskolai évek alatt; van-e különbség a természettudományos tantárgyakat integráltan, illetve diszciplínákra (tantárgyakra) tagoltan tanuló diákok között. Ez utóbbi részmintában különböznek-e a természettudományos tantárgyakat normál, illetve emelt óraszámú tanulóknak eredményei.

Minta

A mintát (1. táblázat) egy fővárosi hat évfolyamos alternatív gimnázium 8. és 9., valamint két kisvárosi négy évfolyamos gimnázium 9–11. évfolyamos tanulói alkották (N=227; a fiúk aránya 41,3 százalék).

1. táblázat. A minta jellemzői

Iskolatípus	Évfolyam	Tanulók száma			
		Normál óraszám*	Emelt óraszám	Összesen	
Gimnázium (diszciplináris oktatás)	9.	8	19	27	120
	10.	18	19	37	
	11.	40	16	56	
Alternatív gimnázium (integrált oktatás)	8.	45	-	45	107
	9.	62	-	62	
Teljes minta		173	54	227	227

*Megjegyzés: az alternatív gimnáziumban a normál óraszámhoz képest 30–40%-kal kevesebb.

Az alternatív gimnáziumban természettudomány tantárgyat tanulnak a diákok, a 8. évfolyamon heti 4, a 9. évfolyamon heti 3 órában. Ez az általános, diszciplináris tantervekhez képest kb. 30–40 százalékkal kevesebb. A tanítás nem a központi tankönyvek alapján történik, hanem az iskolában fejlesztett taneszközök, digitális tananyagok, illetve feladatok felhasználásával. A tananyag a szokásosnál kevesebb ismereti elemet tartalmaz, de megtalálhatók benne a fizika, kémia, biológia alapvető fogalmai, modelljei. A hangsúly a gondolkodáson, a tantárgyközi kapcsolatokon és az attitűdformáláson van. Rugalmasabbak az időkeretek, gyakori a projekt-, illetve a differenciált foglalkoztatás, a gyakorlati feladatok megoldása, a kísérletezés. A vizsgált csoportokban a tanítás során előforduló kísérletek döntően nem irányított vagy nyitott kutatások voltak, inkább strukturált feladatok.

A két kisvárosi gimnáziumban biológia-kémia, illetve matematika-fizika tagozat egyaránt megtalálható. A vizsgálatba bevont tanulók egy része a normál óraszámú kerettanterv szerint halad, a tagozatos diákok emelt óraszámú tanulóknak egy vagy több természettudományos tantárgyat.

Mérőeszköz

A 20 feladatból álló, összesen 68 ítemet tartalmazó teszt négy kutatási készséget mér: (1) kutatási kérdés azonosítása, (2) vizsgálatok tervezése, (3) adatok értelmezése, (4) következtetések levonása. Mivel a cél a tudományos gondolkodás, a természettudományos kutatáshoz szükséges tudás vizsgálata, a válaszok megadása nem igényel szaktudományi ismereteket, vagy ha igen, azok megtalálhatók a feladatok szövegében. A teszt alternatív és többszörös választást, szelektálást igénylő zárt feladatokból áll, illetve három feladat a megadott elemek manipulatív kombinálásával kéri vizsgálatok tervezését.

A továbbiakban egy-egy példát mutatunk be a négy kutatási készséget mérő feladatokra. A kutatási kérdés azonosításához a tanulónak azt kell eldönteni, hogy a leírt kísérleti helyzettel mely kutatási kérdések válaszolhatók meg (1. ábra). A helyes megoldáshoz a kísérleti szituációban azonosítani kell a változókat, azok értékeit.

A tanulók műanyag zacskóból és textilből ejtőernyőt készítettek, amire játékerberkét függesztettek. Az ejtőernyőt leejtették. Többféle ejtőernyővel, de egyféle emberkével kísérleteztek.



Választ kaphattak-e a következő kutatási kérdésekre a kísérlet segítségével?
Kattintással válaszolj!

Hogyan függ az esés ideje...

...az emberke tömegétől? igen nem

...az ejtőernyő tömegétől? igen nem

...az ejtőernyő anyagától? igen nem

...az ejtőernyő méretétől? igen nem

Előző Következő

1. ábra. Példa a kutatási kérdés azonosítását mérő feladatra

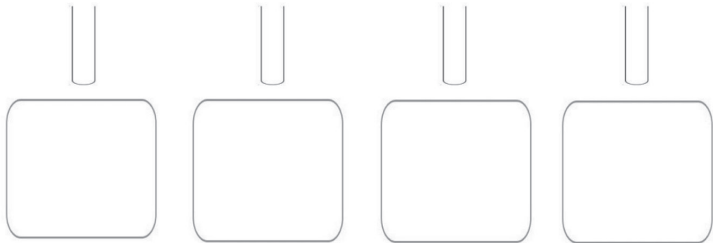
A vizsgálat tervezése részben feladatai a kísérleti elrendezés megtervezését, vizsgálati módszerek kiválasztását, a függő, független és a rögzített változók azonosítását, a kontroll szerepének ismeretét mérik. A 2. ábrán látható feladat interaktív, amelyben nemcsak megtervezni, ténylegesen össze is kell állítani egy vizsgálatosorozatot. Segítette a megoldást, hogy megadtuk, kombinatorikailag négy eset lehetséges.

Egy kísérletben a kutatók az élesztőgombák légzését vizsgálták. Négyféle kémcsőkísérlet állítottak össze desztillált víz, cukor és szárított élesztő felhasználásával. A légzés során keletkező szén-dioxid kimutatása egy jelzőanyag (indikátor) segítségével történt.

Állítsd össze a kísérleti rendszert úgy, hogy bizonyítható legyen: a fejlődő szén-dioxid az élesztőgombák légzésének eredménye!

Mit tegyünk a kémcsövekbe a bizonyításhoz? Egy kémcsőbe több anyagot is beleszálthatsz.
Húzd az anyagok nevét a kémcsövek alatti helyre!

desztillált víz cukor szárított élesztő jelzőanyag



Vizsgálat 1 Vizsgálat 2 Vizsgálat 3 Vizsgálat 4

Előző Következő

2. ábra. Példa a vizsgálat tervezését mérő feladatra

Az adatok elemzését, értelmezését vizsgáló feladatokban a diákoknak táblázatban vagy grafikonon, diagramon közölt adatokat kell elemezniük, adatsorokat összevetniük (3. ábra).

A táblázat egy elejtett (szabadon eső) test által függőlegesen lefelé megtett utat mutat az elejtés pillanatától kezdve. Vizsgáld meg, hogy milyen messzire kerül a test az elejtés helyétől, majd válaszolj a kérdésekre!

Kattints a legördülő menüben a válaszra!

Idő (s)	Hely (cm)
0	0
0,1	5
0,2	20
0,3	45
0,4	80
0,5	125
0,6	180

Melyik az a két időpillanat, amelyek között a legrövidebb utat tette meg a test?

Válassz! ▾

Melyik az a két időpillanat, amelyek között 35 cm utat tett meg a test?

Válassz! ▾

Hogyan változik a test által 0,1 másodpercenként megtett út hossza?

Válassz! ▾

Hogyan mozog a test?

Válassz! ▾

◉ Előző

Következő ◉

Legördülő listában: 0–0,1 s / 0,1–0,2 s / 0,2–0,3 s / 0,3–0,4 s
0,1–0,2 s / 0,2–0,3 s / 0,3–0,4 s / 0,4–0,5 s

Ugyanolyan mértékben nő. / Ugyanolyan mértékben csökken. / Nem ugyanolyan mértékben nő. / Nem ugyanolyan mértékben csökken.

Gyorsul. / Lassul. / Egyenletes a sebessége.

3. ábra. Példa az adatok értelmezését mérő feladatra

A következtetést igénylő feladatok egy-egy kísérletet, annak eredményeit és tapasztalatait írják le. A tanulóknak a rendelkezésre álló információk alapján kell eldönteni, hogy helyesek-e vagy sem a megadott következtetések (4. ábra).

Egy zöldségből vagy gyümölcsből és két fémdarabból elemet állítottak össze a tanulók. Megmérték az elem által létrehozott feszültséget. Többféle összeállítást készítettek. Ezek voltak a tapasztalataik:

- János citromba szűrt egy magnézium- és egy rézlemez, 1,6 V-ot mért.
- Erzszi citromba szűrt egy cink- és egy rézlemez, 0,9 V-ot mért.
- Olivér burgonyába szűrt két vasszöveget, 0 V-ot mért.
- Kati burgonyába szűrt egy cink- és egy rézlemez, 1,1 V-ot mért.

Le lehet-e vonni ezeket a következtetéseket a méréseikből? Kattintással válaszolj!

A feszültség függ a fémek anyagától.	<input type="radio"/> igen	<input type="radio"/> nem
A feszültség nem függ a fémek távolságától.	<input type="radio"/> igen	<input type="radio"/> nem
A feszültség nem függ a fémdarabok nagyságától.	<input type="radio"/> igen	<input type="radio"/> nem
A feszültség függ a zöldségtől vagy a gyümölcstől.	<input type="radio"/> igen	<input type="radio"/> nem

4. ábra. Példa a következtetések levonását mérő feladatra

Adatfelvétel

A mérés 2015 tavaszán zajlott osztálytermi környezetben. A teszt megoldására egy tanóra állt rendelkezésre. Az adatfelvétel számítógépen, az eDia platform (Molnár és Csapó, 2013, Molnár, 2015) alkalmazásával valósult meg. Az adatok elemzése az SPSS 22.0, az Mplus 7.31 és a ConQuest programmal történt.

Eredmények

A teljes teszt reliabilitásmutatója mind a négy vizsgált évfolyamon megfelelő, ami arra utal, hogy összességében a teszt alkalmas a természettudományos kutatási készségek vizsgálatára (2. táblázat). A részterületekre bontásból kiderült, hogy a Cronbach- α értéke három készség esetében megfelelő, a kutatási kérdés azonosításánál viszont alacsony. Ennek egyik oka lehet, hogy ilyen típusú feladatokkal a tanulók valószínűleg még nem találkoztak. Előfordulhat, hogy a „kutatási kérdés” kifejezés volt ismeretlen számukra, és nem értették meg a feladatot, de a későbbiekben azt is meg kell vizsgálnunk, hogy nem feladat- vagy tesztszerkesztési probléma áll-e a háttérben. Mivel e részteszt megbízhatósága nem megfelelő, kihagytuk a tanulói teljesítmények elemzéséből.

2. táblázat. A teszt reliabilitása résztesztenként és évfolyamonként

Résztesztek	Itemszám	Cronbach- α			
		8. évf.	9. évf.	10. évf.	11. évf.
Kutatási kérdés azonosítása	16	0,47	0,55	0,64	0,50
Vizsgálatok tervezése	20	0,85	0,86	0,83	0,85
Adatok értelmezése	16	0,77	0,82	0,79	0,78
Következtetés	16	0,72	0,75	0,72	0,86
Teljes teszt	68	0,90	0,91	0,91	0,90
Teszt (a kutatási kérdés azonosítása nélkül)	52	0,92	0,91	0,90	0,91

A teszt belső összefüggései

A teszt továbbfejlesztéséhez hasznos információkkal szolgál a teszt struktúrájának, belső összefüggésrendszerének elemzése. A korrelációanalízis szerint a résztesztek között a teljes mintában és évfolyamok szerinti bontásban is szignifikáns, közepesen erős a kapcsolat (3. és 4. táblázat). Ez arra utal, hogy a résztesztek által mért készségek kapcsolatban vannak egymással, a mért konstruktum szerves részét képezik.

3. táblázat. A résztesztek közötti korrelációs együtthatók a teljes mintán

Változók	Teszt	Vizsgálatok tervezése	Adatok értelmezése
Vizsgálatok tervezése	0,88		
Adatok értelmezése	0,75	0,52	
Következtetés	0,80	0,55	0,54

4. táblázat. A résztesztek közötti korrelációs együtthatók évfolyamok szerinti bontásban

Változók	Évfolyam											
	Teszt				Vizsgálatok tervezése				Adatok értelmezése			
	8.	9.	10.	11.	8.	9.	10.	11.	8.	9.	10.	11.
Vizsgálatok tervezése	0,87	0,90	0,90	0,82								
Adatok értelmezése	0,79	0,83	0,78	0,75	0,48	0,61	0,49	0,43				
Következtetés	0,83	0,82	0,91	0,81	0,55	0,59	0,72	0,42	0,58	0,58	0,66	0,52

Az elméleti modell további teszteléséhez megerősítő faktoranalízist ('confirmatory factor analyses', CFA) végeztünk. Az elemzés alapját a feladatok adták: hat feladat a vizsgálatok tervezése, négy feladat az adatok értelmezése, és szintén négy feladat a következtetés faktorhoz kapcsolódott. A háromfaktoros modell illeszkedési mutatói megfelelőnek bizonyultak: $\chi^2=103,040$; $df=74$, $CFI=0,941$, $TLI=0,927$, $RMSEA=0,042$. Összehasonlításként az egydimenziós modellre is elvégeztük az elemzést. A modell illeszkedési mutatói jelentősen romlottak: $\chi^2=176,418$; $df=77$, $CFI=0,797$, $TLI=0,760$, $RMSEA=0,076$, a két modell közötti különbség szignifikáns: $\chi^2=72,626$; $df=3$ $p<0,001$. Ezek az eredmények arra utalnak, hogy a három látens faktor empirikusan is igazolható.

Az itemek működésének elemzéséhez a valószínűségi tesztelmélet eszköztárából a dichotóm Rasch-modellt használtuk. Az EAP/PV reliabilitás-értéke megfelelő: 0,882. A személy-item térkép (5. ábra) alapján elmondható, hogy a teszt csak részben fedi le a vizsgált minta tanulóinak képességtartományát. A tanulók többsége a magasabb képességtartományokban, a 0 és 2 logit között helyezkedik el. Mind az alsó, mind a felső tar-

ományban vannak tanulók (az utóbbiban több), akik képességszintjének nem megfelelő a teszt, túl nehéz, illetve túl könnyű. Az itemek nehézségét tekintve megállapítható, hogy az alacsonyabb képesség tartományokban több olyan item is van, amelynél nincsenek tanulók, azok az itemek túl könnyűek voltak. A magasabb képességszinteken fordított a helyzet, nincs olyan item, ami differenciálja a magas képességszintű tanulókat.

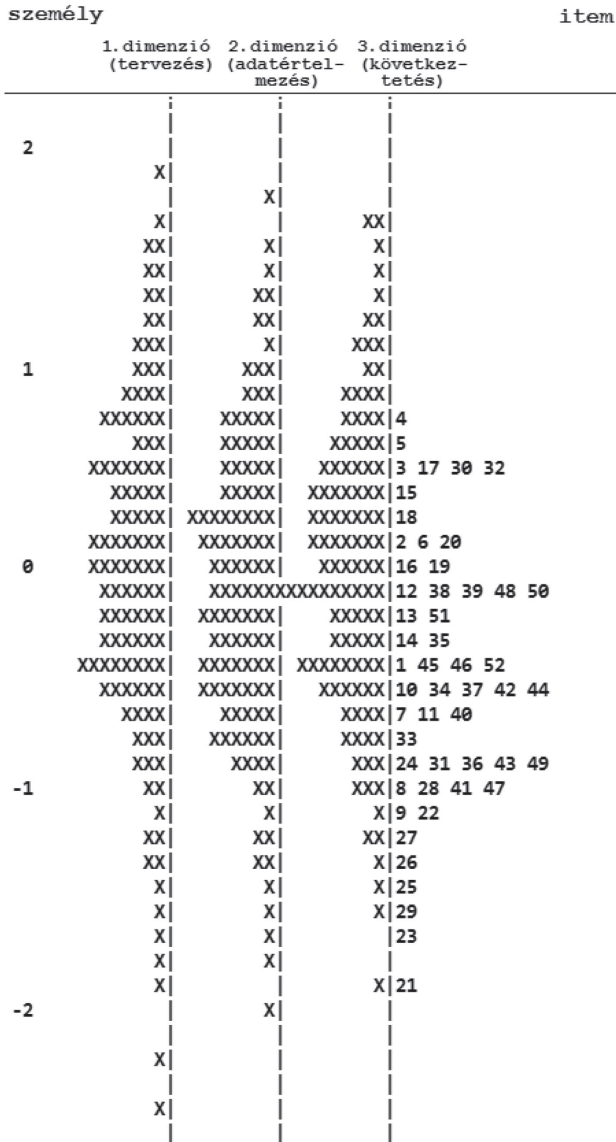
személy	item	
3	XX	
	X	
	XX	
	XX	
	XXX	
	XXXX	
2	XXXXX 4	
	XXXXXXXX 30	
	XXXXXX 5 32	
	XXXXXXXX 17	
	XXXXXXXX 3	
	XXXXXXXX 15	
1	XXXXXX 6 18	
	XXXXXXXX	
	XXXXXXXXXX 2 16 20	
	XXXXXX 48	
	XXXXXXXX 19 38 39 50	
	XXXXXXXX 12 51	
0	XXXXXXXXXX 13 14	
	XXXXXXXX 35	
	XXXXXX 1 46 52	
	XXXXXX 44 45	
	XXXXX 7 10 11 34 37 42	
	XXXX 40	
-1	XX	
	XXX 24 31 33	
	X 36	
	XX 8 28 43 47 49	
	22 41	
	9	
-2	27	
	26	
	25 29	
	23	
	-3	21
		X
X		

Ha az elemzést a résztesztek szerint képzett dimenziókra is elvégezzük, tovább árnyaljuk a képet. A multidimenzionális within-item módszerrel kapott személyi-item térképek (6. ábra) szerint az 1. dimenzióban (tervezés) jelenik meg a legszélesebb képesség tartomány. Mindhárom részteszt (dimenzió) felső képesség tartományában relatíve nagy – a tervezésben 40 (17,6 százalék), az adatértelmezésben 28 (12,3 százalék), a következtetésben 32 (14,1 százalék) – azoknak a tanulóknak a száma, akik képességszintje magasabb, mint ami a teszt bármely itemének megoldásához szükséges. De vannak olyanok is, relatíve kevesebben – a tervezésben 22 (9,7 százalék), az adatértelmezésben 4 (1,8 százalék), a következtetésben 14 (6,2 százalék) –, akik képességszintje alacsonyabb, mint az itemek nehézsége.

A tanulói teljesítmények elemzése

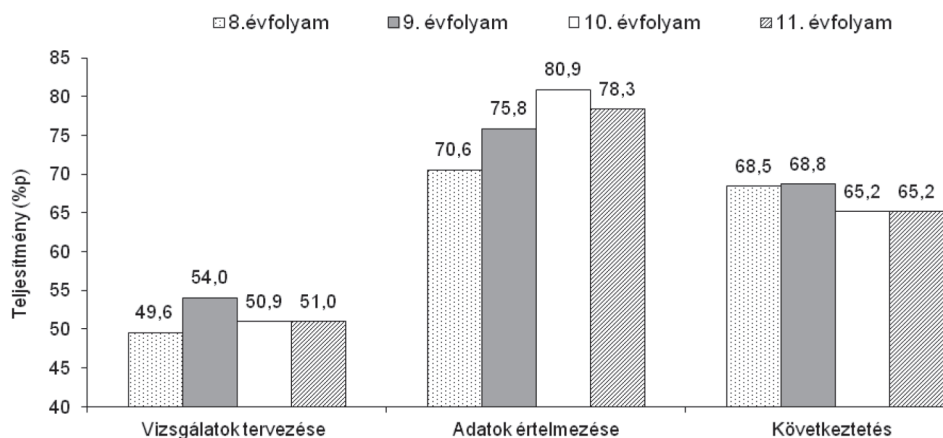
Összehasonlítva az évfolyamokat, a teljes teszten nem, a résztesztek közül csak az adatértelmezésben találtunk szignifikáns különbséget, a 8. évfolyamon gyengébb, a 10. évfolyamon jobb teljesítmények születtek (7. ábra). A nemek között a teljes teszten nincs szignifikáns különbség (fiúk: átlag=61,8 százalékpont, szórás=19,33 százalékpont, lányok: átlag=63,5 százalékpont, szórás=16,7 százalékpont; $F=0,426$, $p=0,51$; $t=0,711$, $p=0,48$). Az adatértelmezésben viszont a lányok voltak jobbak (fiúk: átlag=73,1 százalékpont, szórás=21,9 százalékpont, lányok: átlag=78,3 százalékpont, szórás=16,6 százalékpont; $F=6,662$, $p=0,01$; $t=1,942$, $p=0,05$).

5. ábra. A tanulók és az itemek elhelyezkedése a közös, logitegységben kifejezett képességskálán (egy 'x' átlagosan 1,5 tanulót reprezentál)

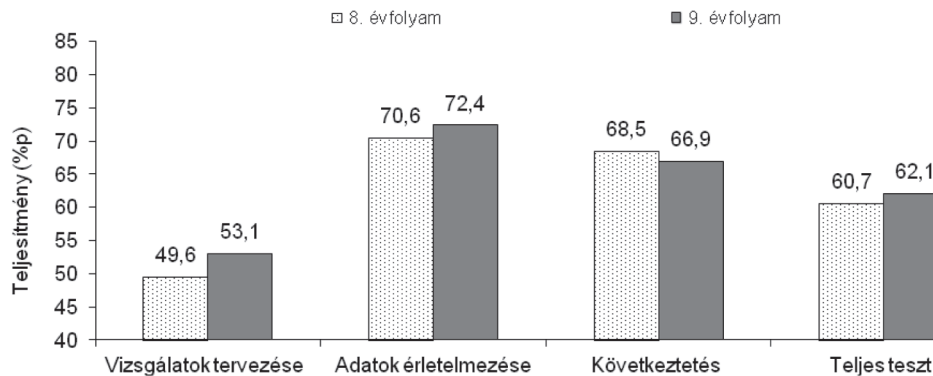


6. ábra. A tanulók és az itemek elhelyezkedése a közös, logitegységben kifejezett képességszállán a három dimenzió szerint (egy 'x' két tanulót reprezentál; itemek sorszáma: tervezés 1–20; adatértelmezés 21–36, következtetés 37–52)

Iskolatípusok szerint vizsgálva a teljesítményeket, az alternatív gimnázium két évfolyamán egyik részteszt esetében sem mutatkozott szignifikáns különbség (8. ábra). Ez azt jelzi, hogy egy tanév alatt lényegesen nem változik a tanulók tudása a vizsgált területen.



7. ábra. A teszten és a részteszten elért teljesítmények évfolyamonként



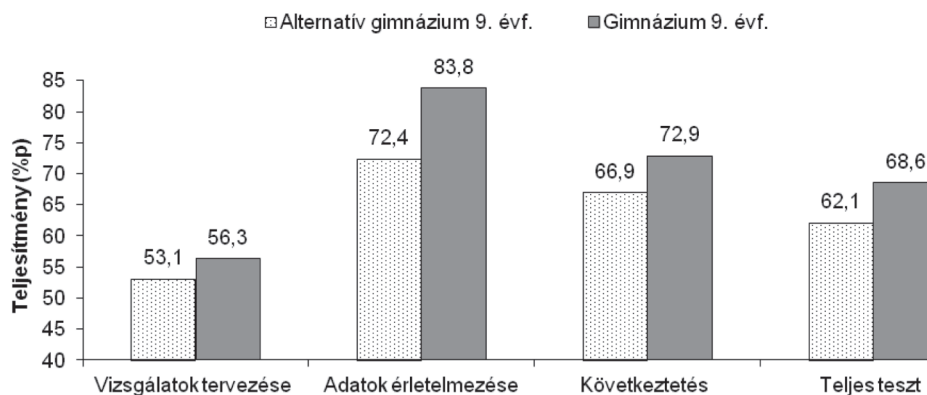
8. ábra. A teszten és a részteszten elért teljesítmény az alternatív gimnáziumban

A négy évfolyamos gimnáziumokban – a várakozásnak megfelelően – a természettudományokat emelt órázámban tanulók teljesítménye szignifikánsan magasabb a teljes teszten ($F=8,599$, $p=0,004$, $t=4,117$, $p=0,00$) és a részteszten egyaránt ($F=1,195$, $p=0,277$, $t=2,674$, $p=0,09$; $F=8,025$, $p=0,005$, $t=3,201$, $p=0,02$; $F=9,900$, $p=0,002$, $t=4,134$, $p=0,00$). Az eredményeket az 5. táblázat mutatja.

5. táblázat. A teszten elért eredmények a négy évfolyamos gimnáziumokban a természettudományos képzés jellege szerint

Részteszt	Normál órászám (N=66)		Emelt órászám (N=54)	
	Átlag (%p)	Szórás (%p)	Átlag (%p)	Szórás (%p)
Vizsgálatok tervezése	50,1	24,8	58,1	20,7
Adatok értelmezése	74,5	20,5	85,2	10,3
Következtetés	63,3	24,1	75,5	15,0
Teljes teszt	60,5	19,2	70,4	11,9

Az iskolatípusok tanulóinak összehasonlítására csak a 9. évfolyamon van lehetőségünk (9. ábra), az eredmények a kis elemszámok miatt tájékoztató jellegűek. A hatosztályos, a természettudományos tantárgyakat integráltan oktató alternatív gimnázium és a négy évfolyamos, diszciplináris természettudományos képzést folytató gimnáziumok között csak az adatértelmezésben van szignifikáns különbség a négy évfolyamos gimnáziumi tanulók javára ($F=9,185$, $p=0,003$ $t=3,223$, $p=0,002$). Mindez arra utal, hogy a kutatási készségek fejlesztésében a lényegesen kevesebb óraszámú megvalósuló, integrált szemléletű oktatás hasonló eredményeket tud elérni.



9. ábra. A hat évfolyamos alternatív és a négy évfolyamos gimnáziumok 9. évfolyamos tanulóinak teljesítményei

Az elemzés szerint az összes lehetséges részmintában a vizsgálatok tervezése részteszt bizonyult a legnehezebbnek, az adatok értelmezése pedig a legkönnyebbnek. Különösen alacsony volt a teljesítmény a 2. ábrán bemutatott feladatban, ahol a tanulók 34 százaléka nem ért el pontot, és csak 5,3 százaléka adott tökéletes megoldást. A nehézséget valószínűleg az okozta, hogy a jó megoldáshoz szisztematikusan át kellett gondolni az egyes vizsgálatokat úgy, hogy minden megadott tényező szerepére külön-külön lehessen következtetni a vizsgálatok eredményei alapján.

Összegzés

Kutatásunk a természettudományos gondolkodás egyik alapvető összetevőjének, a tudományos kutatás készségeinek feltárására irányuló teszt kismintás kipróbálására irányult gimnazista tanulók körében. A vizsgálat eredményei alapján elmondható, hogy a teszt megbízhatóan működik a 8–11. évfolyamon, a mérési adatok illeszkednek a tesztfejlesztés során kialakított elméleti struktúrához; a kutatási kérdések azonosítása részteszt esetében azonban további pontosítás, fejlesztés szükséges.

A teljesítmények csak részben igazolták hipotéziseinket. A 8–11. évfolyamos korosztályban nincs szignifikáns különbség a tanulók eredményeiben. Iskolatípusok szerinti bontásban a négy évfolyamos gimnáziumok tanulói az adatértelmezésben jobbnak bizonyultak, mint az integrált természettudományt tanuló, hat évfolyamos gimnáziumba járó diákok. A természettudományos tantárgyakat emelt szinten tanulók minden kutatási készség esetében szignifikánsan jobban teljesítettek, mint a normál óraszámú képzésben részt vevők.

Pilot kutatásunk eredményei felhívják a figyelmet arra, hogy a középiskolai tanulmányok során nem történik lényeges változás, ami arra utalhat, hogy a hazai természettudományos oktatásban nem kap jelentőségének megfelelő hangsúlyt a természettudományos gondolkodás és a tudományos megismerés módszereinek tanítása, készségeinek fejlesztése.

Adataink jelzik, hogy a mérőeszköz továbbfejlesztéséhez, a kutatási kérdések és a hipotézisalkotás vizsgálatára új feladattípusok kidolgozására van szükség. A kutatás folytatásában az átdolgozott tesztet nagyobb mintán, illetve szélesebb körben próbáljuk ki, fiatalabb korosztályokat és más iskola-típusokat is bevonva.

Köszönetnyilvánítás

A kutatás a SAILS (Strategies for Assessment of Inquiry Learning in Science) FP7-es projekt és az MTA Szakmódszertani Pályázat 2014 támogatásával valósult meg.

A 8–11. évfolyamos korosztályban nincs szignifikáns különbség a tanulók eredményeiben.

Iskolatípusok szerinti bontásban a négy évfolyamos gimnáziumok tanulói az adatértelmezésben jobbnak bizonyultak, mint az integrált természettudományt tanuló, hat évfolyamos gimnáziumba járó diákok.

A természettudományos tantárgyakat emelt szinten tanulók minden kutatási készség esetében szignifikánsan jobban teljesítettek, mint a normál órászámú képzésben részt vevők.

Irodalomjegyzék

- Adey, Philip és Csapó Benő (2012): A természettudományos gondolkodás fejlesztése és értékelése. In: Csapó Benő és Szabó Gábor (szerk.): *Tartalmi keretek a természettudomány diagnosztikus értékeléséhez*. Nemzeti Tankönyvkiadó, Budapest. 17–58.
- Csapó Benő, Korom Erzsébet és Molnár Gyöngyvér (2015, szerk.): *A természettudományi tudás online diagnosztikus értékelésének tartalmi keretei*. Oktatókutatató és Fejlesztő Intézet, Budapest.
- Csapó Benő és Szabó Gábor (2012, szerk.): *Tartalmi keretek a természettudomány diagnosztikus értékeléséhez*. Nemzeti Tankönyvkiadó, Budapest.
- Dunbar, K. és Fugelsang, J. (2005): Scientific Thinking and Reasoning. In: Holyoak, K. J. és Morrison, R. G. (szerk.): *The Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press, Cambridge. 705–725.
- Gormally, C., Brickman, P., Hallar, B. és Armstrong, N. (2009): Effects of inquiry-based learning on students' science literacy skills and confidence. *International Journal for the Scholarship of Teaching and Learning*, 3. 2. sz. 16.
DOI: [10.20429/ijotl.2009.030216](https://doi.org/10.20429/ijotl.2009.030216)
- Harlen, W. (2013): *Assessment and inquiry based science education: Issues in policy and practice*. TWAS-Strada Costiera, Trieste.
- Kind, P. M. (2013): Establishing assessment scales using a novel disciplinary rationale for scientific reasoning. *Journal of Research in Science Teaching*, 50. 5. sz. 530–560. DOI: [10.1002/tea.21086](https://doi.org/10.1002/tea.21086)
- Korom Erzsébet, B. Németh Mária és Pásztor Attila (2015): Kutatási készségek online vizsgálata 6. és 8. évfolyamon. In: Csikos Csaba és Gál Zita (szerk.): *13. Pedagógiai Értékelési Konferencia: PÉK 2015: Program, tartalmi összefoglalók*. SZTE BTK Neveléstudományi Doktori Iskola, Szeged. 174.
- Korom Erzsébet, Molnár Gyöngyvér és Csapó Benő (2015): A természettudományi online diagnosztikus mérések tartalmi kereteinek elméleti háttere. In: Csapó Benő, Korom Erzsébet és Molnár Gyöngyvér (szerk.): *A természettudományi tudás online diagnosztikus értékelésének tartalmi keretei*. Oktatókutatató és Fejlesztő Intézet, Budapest. 17–33.
- Kuhn, D. (2002): What is scientific thinking and how does it develop? In: Goswami, U. (szerk.): *Handbook of childhood cognitive development*. Blackwell, Oxford. 371–393. DOI: [10.1002/9780470996652](https://doi.org/10.1002/9780470996652)

- Lawson, A. E. (1978): The development and validation of a Classroom Test of Formal Reasoning. *Journal of Research in Science Teaching*, **15**. 1. sz. 1–24. DOI: [10.1002/tea.3660150103](https://doi.org/10.1002/tea.3660150103)
- Lawson, A. E. (1995): *Science teaching and the development of thinking*. Wadsworth Publishing Company, California.
- Liu, O. L., Lee, H-S. és Linn, M. C. (2010): Multifaceted assessment of inquiry-based science learning. *Educational Assessment*, **15**. 2. sz. 69–86. DOI: [10.1080/10627197.2010.491067](https://doi.org/10.1080/10627197.2010.491067)
- Mayer, D., Sodian, B., Koerber, S. és Schwippert, K. (2014): Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction*, **29**. 43–55. DOI: [10.1016/j.learninstruc.2013.07.005](https://doi.org/10.1016/j.learninstruc.2013.07.005)
- Molnár Gyöngyvér (2015): *A képességmérés dilemmái: a diagnosztikus mérések (eDia) szerepe és helye a magyar közoktatásban*. Géniusz Műhely Kiadványok. 2. sz. 16–29.
- Molnár Gyöngyvér és Csapó Benő (2013): Az eDia online diagnosztikus mérési rendszer. In: Józsa Krisztián és Fejes József Balázs (szerk.): *II. Pedagógiai Értékelési Konferencia: PÉK 2013: Program, előadás-összefoglalók*. SZTE BTK Neveléstudományi Doktori Iskola, Szeged. 82.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O’Sullivan, C. Y. és Preuschoff, C. (2009): *TIMSS 2011 assessment frameworks*. TIMSS & PIRLS International Study Center, Boston College. Chestnut Hill, MA.
- Nagy Lászlóné, Korom Erzsébet, Pásztor Attila, Veres Gábor és B. Németh Mária (2015): A természettudományos gondolkodás online diagnosztikus értékelése. In: Csapó Benő, Korom Erzsébet és Molnár Gyöngyvér (szerk.): *A természettudományi tudás online diagnosztikus értékelésének tartalmi keretei*. Oktatókutató és Fejlesztő Intézet, Budapest. 35–116.
- National Research Council (NRC) (1996): *National science education standards*. National Research Council. National Academies Press, Washington, DC. DOI: [10.17226/4962](https://doi.org/10.17226/4962)
- National Research Council (NRC) (2000): *Inquiry and the national science education standards: A guide for teaching and learning*. Center for Science, Mathematics, and Engineering Education, National Research Council. The National Academies Press, Washington, DC. DOI: [10.17226/9596](https://doi.org/10.17226/9596)
- NGSS Leads States (2013): *Next generation science standards: For states, by states*. The National Academies Press, Washington, DC. DOI: [10.17226/18290](https://doi.org/10.17226/18290)
- OECD (2014): PISA 2012 results in focus. What 15-year-olds know and what they can do with what they know. OECD, Paris. DOI: [10.1787/9789264208780-en](https://doi.org/10.1787/9789264208780-en)
- Osborne, J. (2013): The 21st century challenge for science education: Assessing scientific reasoning. *Thinking Skills and Creativity*, **10**. sz. 265–279. DOI: [10.1016/j.tsc.2013.07.006](https://doi.org/10.1016/j.tsc.2013.07.006)
- Piaget, J. (1964): Cognitive development in children: development and learning. *Journal of Research in Science Teaching*, **2**. 3. sz. 176–186. DOI: [10.1002/tea.3660020306](https://doi.org/10.1002/tea.3660020306)
- Raodrangka, V. (1991): The construction of a Group Assessment of Logical Thinking (GALT). *Kasetsart Journal: Social Sciences*, **12**. 2. sz. 148–154.
- Shayer, M., Adey, P. és Wylam, H. (1981): Group tests of cognitive development: Ideals and a realization. *Journal of Research in Science Teaching*, **18**. 2. sz. 157–168. DOI: [10.1002/tea.3660180208](https://doi.org/10.1002/tea.3660180208)
- Tobin, K. G. és Capie, W. (1981): The development and validation of a Group Test of Logical Thinking. *Educational and Psychological Measurement*, **41**. 2. sz. 413–423. DOI: [10.1177/001316448104100220](https://doi.org/10.1177/001316448104100220)
- Wellnitz, N., Hartmann, S. és Mayer, J. (2010): Developing a paper-and-pencil-test to assess students’ skills in scientific inquiry. In: Çakmakçı, G. és Taşar, M. F. (szerk.): *Contemporary science education research: Learning and assessment*. Pegem Akademi. Ankara. 289–294.
- Wenning, C. J. (2005): Levels of inquiry: Hierarchies of pedagogical practices and inquiry processes. *Journal of Physics Teacher Education Online*, **2**. 3. sz. 3–11.
- Wenning, C. J. (2007): Assessing inquiry skills as a component of scientific literacy. *Journal of Physics Teacher Education Online*, **4**. 2. sz. 21–24.
- Yalaki, Y., Çakmakçı, G., Yahşi, D., Gümüş, B. S., Gürel, A., Yüksel, G. K. és İnce, İ. (2014): *Development and validation of an assessment instrument for inquiry skills*. SAILS / SMEC Conference 2014, DCU, Dublin, Ireland.
- Zhou, S., Han, J., Koenig, K., Raplinger, A., Pi, Y., Li, D., Xiao, H., Fu, Z. és Bao, L. (2016): Assessment of scientific reasoning: The effects of task context, data, and design on student reasoning in control of variables. *Thinking Skills and Creativity*, **19**. sz. 175–187. DOI: [10.1016/j.tsc.2015.11.004](https://doi.org/10.1016/j.tsc.2015.11.004)
- Zimmerman, C. (2007): The development of scientific thinking skills in elementary and middle school. *Developmental Review*, **27**. 2. sz. 172–223. DOI: [10.1016/j.dr.2006.12.001](https://doi.org/10.1016/j.dr.2006.12.001)