

ELTÉRŐ ITEMMŰKÖDÉS VIZSGÁLATA AZ ORSZÁGOS KOMPETENCIAMÉRÉSBEN HALMOZOTTAN HÁTRÁNYOS HELYZETŰ DIÁKOK KÖRÉBEN TÖBBDIMENZIÓS IRT MODELLEK SEGÍTSÉGÉVEL

Kispál Sára¹, Gergely Bence¹

Levelező szerző: Kispál Sára (sarakispal@gmail.com)

Kivonat

Kutatásunk célja, a halmozottan hátrányos helyzetű diákok Országos kompetenciamérésén nyújtott teljesítményének feltérképezése, melynek alapját Kispál, Gergely, T. Kárisz és Takács (2020, kézirat) kutatásának eredményei adták.

A halmozottan hátrányos helyzetű diákok minden évben részt vesznek az Országos kompetenciamérésen, ám teljesítményük elmarad az átlagtól. Fő célkitűzéseink közé tartozik megvizsgálni, hogy itemek szintjén megjelenik-e különbség a halmozottan hátrányos helyzetű diákok feladatmegoldásában, vannak-e olyan feladatok melyek kifejezetten nehezebbek vagy könnyebbek a számukra, illetve tudják-e kompenzálni az egyik kompetenciatérületen való hiányosságukat a másikkal?

Két statisztikai modell vizsgálata során (továbbiakban M2, M4 modell) megállapíthattuk, hogy mindkét esetben fennáll differenciált itemműködés, ugyanakkor az M4-es modell mutatója láthatóan, hogy a teszt információ-görbéje a szövegértés dimenzióról sokkal több információt szolgáltat, mint a matematika dimenzióról. Ugyanakkor jelentősen elmarad az információ-tartalom maximuma a nem halmozottan hátrányos helyzetű csoporttól. Item-jellemzők és tartalmi kódolás vizsgálata esetén csak a tényismeret és egyszerű műveletek területén figyelhető meg szignifikánsan nagyobb esély a differenciált itemműködés megjelenésére.

Összegezve elmondható, hogy a halmozottan hátrányos helyzetű diákok a matematika feladatok megoldásánál sokkal inkább a szövegértési kompetenciájukat veszik igénybe. Abban az esetben, ha a feladat nem kíván több matematika tudást, mint szövegértést, nem okoz differenciált itemműködést, míg ellenkező esetben igen. A szerzők a tanulmány alapjául szolgáló kutatást az Országos kompetenciamérés kutatócsoport 20642B800 témaszámú, a Károli Gáspár Református Egyetem Bölcsészeti és Társadalomtudományi Kara által finanszírozott pályázat keretében végezték. Dolgozatunk mellékleteként angol nyelven is közöljük tanulmányunkat.

Kulcsszavak: Országos kompetenciamérés ■ halmozottan hátrányos helyzetű diákok ■ többdimenziós IRT modellek ■ Differenciált itemműködés

¹ Károli Gáspár Református Egyetem, Pszichológiai Intézet, Budapest, 1034, Bécsi út 324.

Abstract

The aim of our research is to map the performance of disadvantaged students in the National Assessment of Basic Competences, which is based on the research results of Kispál, Gergely, T. Káráz and Takács (2020, manuscript). Disadvantaged students participate in the National Assessment of Basic Competences every year, but their performance lags behind the average. Our main objectives are to examine whether there is a difference at the levels of items in the problem solving of students with disadvantages, whether there are tasks that are specifically more difficult or easier for them, and whether they can compensate for their shortcomings in one area of competence with another.

Examining two statistical models (henceforth M2, M4 models) we found that in both cases there is a differential item functioning; however, the M4 model shows that the test information curve provides much more information about the reading comprehension dimension than the mathematics dimension. At the same time, the maximum information content lags significantly behind the non-disadvantaged group. When examining item characteristics and content coding, a significantly higher chance of the emergence of differential item functioning can be observed only in the field of factual knowledge and simple operations.

In summary, disadvantaged students use their reading comprehension competencies much more when solving math problems. In the case where the task does not require more mathematical knowledge than reading comprehension competencies, it does not cause differential item functioning, while otherwise it does. The authors of the study carried out the research on the basis of the National Assessment of Basic Competences Research Group with the topic number 20642B800, funded by the Faculty of Humanities and Social Sciences, Károli Gáspár University of the Reformed Church in Hungary. As an appendix to our study, we also publish our dissertation in English.

Keywords: disadvantaged students ■ National competence measurement ■ multidimensional IRT models ■ differential item functioning

BEVEZETÉS

A halmozottan hátrányos helyzetű tanulók minden évben részt vesznek az Országos Kompetenciamérésen, viszont teljesítményük elmarad a társaikéhoz képest. Ennek okait több kutatásban igyekeztek feltárni (ld. Mendi, 1999; Buda, 2017; Fejes, 2013 stb.), melyek során a teljesítményt leginkább befolyásoló tényezőként az értelmiségi szerepminta hiányát sorolták fel (Mendi, 1999), hiányzik az életükből a mintaként szolgáló családszerkezet, családfelépítés (Kasik és mtsai., 2014) és a stabil anyagi háttér (Kertesi és Kézdi, 2012).

Jelen vizsgálat fő célkitűzése újraértékelni azt a megítélést, hogy a halmozottan hátrányos helyzetűek rosszabbul teljesítenek az Országos kompetenciamérésen. Lehetséges-e valamilyen eltérés már az itemek szintjén? Vannak-e esetleg olyan feladatok, melyekben a halmozottan hátrányos helyzetű tanulók jobban teljesítenek, mint társaik? Tudják-e az egyik kompetenciaterületen lévő hiányosságaikat a másik kompetenciaterülethez tartozó tudásukkal kompenzálni?

KOMPETENCIA FOGALMA, AZ ORSZÁGOS KOMPETENCIAMÉRÉS ÉS A HALMOZOTTAN HÁTRÁNYOS HELYZETŰ DIÁKOK MAGYARORSZÁGON

A kompetencia fogalom definiálását több szempontból közelíthetjük meg. Hétköznapi értelemben a kompetencia alkalmasságot, ügyességet jelent. Nemzetközi oktatási mérések nyomán vált el egymástól a műveltség (literacy) és a kompetencia (competence) fogalma. Az oktatásban használt kompetencia fogalmak egyik új értelmezése a PISA tesztekhez köthető. Utóbbihoz kapcsolódóan a kompetenciamérés alapját olyan képességek mérésének kell alkotnia, melyek relevánsak a diákok számára a mindennapi életük során (Trier és Peschar, 1995). Többféle kompetenciaértelmezés létezik, akár aszerint is, hogy melyik tudományág használja. Az egyik kompetencia fogalom szerint a kompetencia kontextus-specifikus kognitív diszpozíció, mely szükséges egy adott helyzet kezeléséhez, megoldásához (Weinert, 2001; Klieme és Leutner, 2006). Ez a meghatározás a kompetencia képességalapú megközelítéséhez kapcsolódik. Dweck és Leggett (1988) modellje szerint a kompetencia azon alapul, hogy az adott személy a kihívást jelentő helyzetben milyen szinten tud teljesíteni. Ez a kihívást jelentő helyzet a kudarc utáni teljesítményt jelenti. Kétféle csoportot különböztettek meg e szempont alapján: a telhetetleneket, akik szorongtak, nem tudtak megküzdeni a helyzettel és az elsajátítás orientáltakat, akik intrinzik motivációval reagálnak a helyzetre. Ez a kompetenciafogalom tehát a motiváció oldaláról magyarázza a konstruktumot.

Nemzetközi tudásszintmérő vizsgálatok és az Országos kompetenciamérés

Az 1970-es évek elejétől kezdve fokozatosan voltak jelen Magyarországon nemzetközi vizsgálatok, melyek a diákok teljesítményét: a tudásszintet, illetve a mindennapi élethez szükséges ismereteket, a tudás hasznosítását mérték (TIMSS, PIRLS, PISA). A felmérések fő célja az volt, hogy nemzetközi szinten kapjanak visszajelzést egy-egy ország oktatási helyzetéről és rangsorokat hozzanak létre a diákok eredményei alapján (Csapó, Fehér, Kinyó és Tóth, 2014).

Az Oktatási Hivatal szervezésében, 2001-től kezdve minden év tavaszán megrendezik az Országos kompetenciamérést. „A kompetenciamérés elsődleges célja az oktatási kormányzat tájékoztatása az alap- és középfokú oktatás eredményességéről a tanulási eredmények, valamint az esélyegyenlőség nézőpontjából. Emellett a kompetenciamérés fejlesztői feladatuknak tekintik, hogy az iskolákat és fenntartókat ellássák a helyi szintű értékeléshez és célszerű beavatkozások tervezéséhez szükséges visszajelzéssel, tovább elemezhető adatokkal (fenntartói, intézményi és telephelyi jelentések).” (Országos Kompetenciamérés Technikai Leírás, 2012, p. 8)

Az Országos kompetenciamérésen hatodik, nyolcadik és tizedik évfolyamos diákok vesznek részt minden évben. A teszt időpontját az Oktatási Hivatal egységesen jelöli ki. A mérés két részből áll: szövegértés és matematika feladatsorokat kell kitöltenie a diákoknak, mindkettő teszt sor kitöltésére kétszer negyvenöt perc áll a rendelkezésükre. A feladatsorok kialakításakor a matematikai eszköztudás és a szövegértés fogalmát vették alapul, melynek értelmében a matematika feladatsoroknál az életben való alkalmazáson,

míg a szövegértés esetén a dekódoláson, megértésen és reflektáláson van a hangsúly (Szabó, Szepesi, Takácsné és Vadász, 2017).

A tesztfeladatok minden évben egy számos kritériumot – pl. az életszerűség szempontját - tartalmazó előszűrést követően kerülnek a végleges feladatsorba. Fontos a feladattípusok változatossága is, hiszen egyéb esetben a feladatok könnyen begyakorolhatóak lesznek és a mérés pontatlan lesz (Weinert, 2001). A feladatokban kerülnek a márka- nevek, országnevek és politikai tartalmak használatát, nem megfelelőek a félrevezető feladatok és az olyan feladatok sem, melyek túl hosszúak vagy túl rövidek. A feladatoknak minden esetben besorolhatónak kell lenniük (Országos Kompetenciamérés Technikai Leírás, 2012).

Kutatásunkban a matematikai feladatokban megjelenő itemműködési eltéréseket vizsgáljuk. A kompetenciamérésben a matematika feladatokat a matematikai eszköztudás fogalma alapján alakítják ki, mely magába foglalja „az egyénnek azt a képességét, amelynek segítségével megéri és elemzi a matematika szerepét a valós világban; a matematikai eszköztár készség szintű használatát; az elsajátított matematikai tudás valós élethelyzetekben való alkalmazásának igényét és az erre való képességet; a matematikai eszközök használatát a társadalmi kommunikációban és együttműködésben az egyén életkorának megfelelő szinten.” (Balázi és mtsai, 2014, p. 33). A tesztben különböző tudás-területekről származó feladatok találhatóak: mennyiségek, számok és műveletek; hozzárendelések, összefüggések; alakzatok és tájékozódás; statisztikai jellemzők és valószínűségi számítás. A feladatok ezen felül háromféle gondolkodási művelet alapján csoportosíthatóak: tényismeret, alkalmazás és integráció, illetve a komplex megoldások és értékelés (Balázi és mtsai, 2014). Ez a feladatfelépítés nagyban egyezik a PISA mérés felépítésével (Ehmke, van den Ham, Sälzer, Heine és Prenzel, 2020).

Halmazottan hátrányos helyzet Magyarországon

A halmazottan hátrányos helyzet definícióját az 1997. évi XXXI. törvény a gyermekek védelméről és a gyámügyi igazgatásról szabja meg. Ennek értelmében „67/A § (1) Hátrányos helyzetű az a rendszeres gyermekvédelmi kedvezményre jogosult gyermek és nagykorúvá vált gyermek, aki esetében az alábbi körülmények közül egy fennáll: a) a szülő vagy a családba fogadó gyám alacsony iskolai végzettsége (...) b) a szülő vagy a családbafogadó gyám alacsony foglalkoztatottsága (...) c) a gyermek elégtelen lakókörnyezete, illetve lakáskörülményei (...) ahol korlátozottan biztosítottak az egészséges fejlődéshez szükséges feltételek. (2) Halmazottan hátrányos helyzetű a) az a rendszeres gyermekvédelmi kedvezményre jogosult gyermek és nagykorúvá vált felnőtt, aki esetében az (1) bekezdés a) - c) pontjában meghatározott körülmények közül legalább kettő fennáll, b) a nevelésbe vett gyermek, c) az utódgondozói ellátásban részesülő és tanuló vagy hallgatói jogviszonyban álló fiatal felnőtt.”

A halmazottan hátrányos helyzetű diákok teljesítményét kutató vizsgálatok (pl. Kertesi és Kézdi, 2012; Mendi, 1999; Horváth, 2015; Radó, 1997 stb.) több befolyásoló tényezőt állapítottak meg. Ezek közül a legfontosabbak a szocioökonomiai státusz hatása, az értelmiségi szerepminta hiánya, a motivációvesztettség, az iskolai szegregáció és területi eltérések az országon belül, illetve az esetleges kognitív működésbeli különbségek.

A szocioökonómiai státusz nagy mértékben befolyásolja a diákok iskolában nyújtott teljesítményét. A magasabb SES-el rendelkező diákok sokkal jobban teljesítenek standardizált tesztekben, mint az alacsonyabb SES-es rendelkező társaik (Beese és Liang, 2010). Kertesi és Kézdi (2012) kutatásukban a roma és nem roma fiatalok teszteredményeit vetették össze. Megállapították, hogy a lemaradás háttérben leginkább jövedelmi, iskolázottsági és lakóhelyi különbségek állnak, illetve az anyagi háttér befolyása aránytalanul nagy hatással van a diákok teljesítményére. Fontos az is, hogy a mélyszegénységben élők esetében a jövedelemnövekedés sokkal nagyobb mértékű javulást eredményez, mint a társadalom közép- vagy felső osztályában.

A halmozottan hátrányos helyzetű diákok rosszabb teljesítményét befolyásolja, hogy ezek a diákok sok esetben nem kapnak értelmiségi szerepmintát a családjukon belül (Mendi, 1999 idézi Horváth, 2015). A diákok sokszor mozaikcsaládban vagy egyszülős családban nőnek fel, ez negatívan befolyásolja a szociális-problémamegoldó képességeiket, mely legfőképpen serdülőkorban jelenik meg (Kasik és mtsai., 2014). A halmozottan hátrányos helyzetű diákok sokszor találkoznak negatív előítéletekkel az iskolán belül, akár a pedagógusoktól is, mely szintén csökkenti a teljesítményüket (Radó, 1997 idézi Horváth, 2015).

A motivációvesztettség is jellemezheti a halmozottan hátrányos helyzetű diákokat. Ennek magyarázatául sokszor a tananyag felépítését, akadémiusságát és elméletiességét említik, illetve azt, hogy az iskolában látott minták nem felelnek meg a tanuló addig ismert „valóságképének”. Mivel a halmozottan hátrányos helyzetű tanulók sokszor másfajta kultúrában nevelkedtek, így idegen számukra az iskolában tapasztalt dolgok nagy része, ezért ignorálják azokat (Buda, 2017).

A szegregáció szintén egy olyan tényező, mely negatív befolyással van a halmozottan hátrányos helyzetű diákok teljesítményére. Mivel Magyarországon a szegregáció rendkívül jellemző a közoktatásban (Fejes, 2013), a magyar diákok teljesítményét nagy mértékben befolyásolja. Jelen esetben e fogalom alatt nem a kisebbségek elkülönítését értjük, hanem az egy osztályba vagy iskolába járó gyermekek családi háttérének homogenitását (Fejes, 2013). Sok szülő körzeten kívüli iskolába írta gyermekét, hogy jobb oktatásban részesüljön, így a szegregáció egyre jellemzőbb lesz (Keller és Mártonfi, 2006). Az alacsonyabb szocioökonómiai státusszal rendelkező diákok teljesítményére kedvezően hat, ha olyan intézménybe járnak, melyben a többség családi háttere jónak tekinthető, ám ennek megvalósulása Magyarországon még nem következett be (Fejes, 2006; Cs. Czachesz és Radó, 2003).

A szegregációt elősegítette a tény, hogy a piacgazdaság kialakulásával hazánkban egyre nagyobbak lettek a területi különbségek (város, falu) (Faluvégi, 2004). Az ország Budapest központú lett, az elmaradottabb térségekben a diákok teljesítménye is elmarad a fővároshoz képest (Faluvégi, 2004 idézi Keller és Mártonfi, 2006). Az ország elmaradottabb térségeiben a munkanélküliség aránya sokkal magasabb, így sokkal több a halmozottan hátrányos helyzetű diák (Rajnai, 2012).

A kognitív működésbeli eltérések sok diák esetében okoznak lemaradást társaikhoz képest, így felmerülhet a kérdés, hogy állhat-e ilyenfajta eltérés a halmozottan hátrányos helyzetű diákok teljesítményének háttérben? Az absztrakt konceptualizálás és a vizuális stimulánsok kategorizálásában sok esetben elmaradnak a társaiktól a halmozottan

hátrányos helyzetű diákok (Deutsch, 1964). Siller (1957) szerint, az absztrakcióban való elmaradást az okozza, hogy a magasabb szocioökonómiai státusszal rendelkező diákok sokkal nagyobb arányban alkalmazzák az absztrakciót, mint az alacsonyabb SES-el rendelkező társaik. A hátrányos helyzetű diákok kevésbé flexibilisek feladatok megoldásakor, inkább konkrétan gondolkodnak társaikhoz képest (McCandles, 1952).

A DIFFERENCIÁLT ITEMMŰKÖDÉS ÉS A TUDÁSSZINTMÉRŐ VIZSGÁLATOK

Ahhoz, hogy egy tudásszintmérő vizsgálat eredményeit tekintve össze lehessen hasonlítani különböző csoportokat, meg kell vizsgálnunk, hogy a tesztben szereplő egyes itemek ugyanolyan nehézségűek-e a csoport tagjainak. Ennek megállapítását általában differenciált itemműködést vizsgáló eljárásokkal szokták feltárni (Holland és Wainer, 2012). A DIF akkor jelenik meg, amikor a tesztkitöltők, akik egy adott látens tulajdonság azonos szintjével rendelkeznek, ám más alpopulációból származnak, szignifikánsan eltérő valószínűséggel oldják meg helyesen az adott itemet, azaz a két csoport item-valószínűség görbéje eltér egymástól vagy a nehézség vagy a diszkriminancia index (vagy ezek interakciójának) tekintetében (Huang, Wilson és Wang, 2014; Chalmers, Counsell és Flora, 2016).

Nemzetközi felmérések esetén (pl.: PISA, TIMSS) több olyan kutatást folytattak, mellyel a differenciált itemműködést vizsgálták. Shanmugam (2020) kutatásában a TIMSS teszt matematika kérdéseit tekintette át. Célja az volt, hogy feltárja: függ-e az itemek működése a diák nemétől? Eredménye differenciált itemműködést állapított meg a lányok esetében olyan feladatoknál, melyek magasabb rendű gondolkodási képességeket igényeltek, míg a fiúk számára a törteket és a negatív számokat tartalmazó feladatok bizonyultak nehezebbnek.

Da Costa és Araújo (2012) a bevándorló és nem bevándorló diákok PISA szövegértés tesztjének eredményeit vizsgálta. Feltételezésük az volt, hogy a bevándorló diákok számára nehézséget okoz olyan feladat megoldása, mely a jelenlegi hazájukkal kapcsolatos kulturális vonatkozásokkal rendelkezik. Eredményeikben megállapították, hogy vannak olyan feladatok, melyeket a bevándorló diákok könnyebben oldanak meg, mint nem bevándorló társaik. Jellemzően azoknál a feladatoknál volt előnyük, melyek célja az oktatás, a tudás átadása és olyan szituációkhoz köthetőek melyeket az iskolában sajátíthattak el.

TÖBBDIMENZIÓS IRT MODELLEK ÉS DIFFERENCIÁLT ITEMMŰKÖDÉS

Mivel a kompetencia, ahogyan a definíciója is mutatja, összetett jelenség, a mérőeszközök sok esetben a modern valószínűségi tesztelmélet modelljeit alkalmazzák

(későbbiekben IRT (item response theory). Ennek lényege, hogy minden item esetén megállapítható egy látens dimenzió különböző szintjére, hogy az egyén milyen valószínűséggel válaszol helyesen (Lord, 1980). Mivel a kompetencia mérésénél a hagyományos egydimenziós IRT modellek használata kapcsán sok probléma merült fel, így többdimenziós IRT modelleket érdemes alkalmazni. Ennek alkalmasságát Walker és munkatársai (2008) szemléltették. Kutatásukban a szöveges matematika itemek esetén esetlegesen fellelhető DIF-et vizsgálták, abban az esetben, ha a csoportosító változó a diák szövegértés szintje. Megállapították, hogy a diák matematika feladatmegoldását befolyásolja, hogy milyen szinten olvas. Mivel a matematika feladatokban megjelenik a szövegértés mérése is, így nehezen lehet szétválasztani a két dimenziót, ám Walker és munkatársai (2008) egy félig-feltáró IRT modellel változtatával sikeresen elkülönítették ezeket a dimenziókat.

Kutatásunk célja, hogy az Országos kompetenciamérés matematika itemei között már feltárt DIF-el rendelkező itemeket (Kispál, Gergely, T. Kárász és Takács, 2020, kézirat) megvizsgáljuk, hogy van-e olyan item vagy tartalmi jellemző, amely megléténél nagyobb valószínűséggel fordul elő differenciált itemműködés. Vizsgálatunk során egy többdimenziós IRT modell (Gergely, T. Kárász és Takács, 2020, kézirat) esetén is megvizsgáljuk ezen eltéréseket, ugyanis úgy gondoljuk, hogy vannak olyan itemek melyek esetén a hagyományos eljárások nem tárják fel a DIF-et, mely legfőképpen abból adódik, hogy a tanuló az egyik képesség hiányában a másik képességgel kompenzál.

KUTATÁSI ELŐZMÉNYEK

A jelen tanulmány a Károli Gáspár Református Egyetem Kompetenciamérés kutatócsoportjának munkájához, kutatásaihoz kapcsolódóan jött létre, melynek keretében az Országos kompetenciamérés kutatócsoport 20642B800 témaszámú kutatásához kapcsolódhattunk. A kutatás részben a Károli Gáspár Református Egyetem Bölcsész- és Társadalomtudományi Kara által finanszírozott pályázat keretében valósul meg. Dolgozatunk mellékleteként angol nyelven is közöljük tanulmányunkat. A dolgozat előzményeihez szervesen hozzátartoznak az alábbi közlemények: a szülői háttér családi szintű aspektusairól Nyitrai és munkatársai két tanulmányban is elérhetővé tették eddigi eredményeiket (Nyitrai és mtársai, 2019a, 2019b), illetve angol nyelven szintén elérhető a tanulmány (Nyitrai és mtársai, 2019c). A szülői bevonódás iskolai szintű aspektusairól Koltói és munkatársai számolnak be két magyar és egy angol nyelven írt tanulmányukban (Koltói és mtársai, 2019a, 2019b, 2019c). Fontos kiemelni, hogy a szülői bevonódás mellett az iskolai és családi háttér, szociokulturális státusz is jelentős különbségekkel társul az iskolai teljesítmény tekintetében. Az Országos kompetenciamérés kapcsán e különbségekről írnak Harsányi és munkatársai szintén két magyar és egy angol nyelvű anyagukban (Harsányi és mtársai, 2019a, 2019b, 2019c). Továbbá mindenképpen meg kell említenünk azt is, hogy az Országos kompetenciamérés hivatalos eredményeihez módszertanilag más megközelítések tartoznak. Esetünkben nem is a közvetlen eredményekre, hanem a háttérjelenségek összefüggéseire szerettünk volna rávilágítani, így e

tekintetben nem kellett az alapvetően lényegesen bonyolultabb matematikai-statisztikai módszertanra támaszkodnunk, melyről magyar és angol nyelven egyaránt elérhető T. Kárász Judit összefoglaló munkáiban (T. Kárász, 2019a, 2019b).

MÓDSZER

Minta és felhasznált mérőeszköz

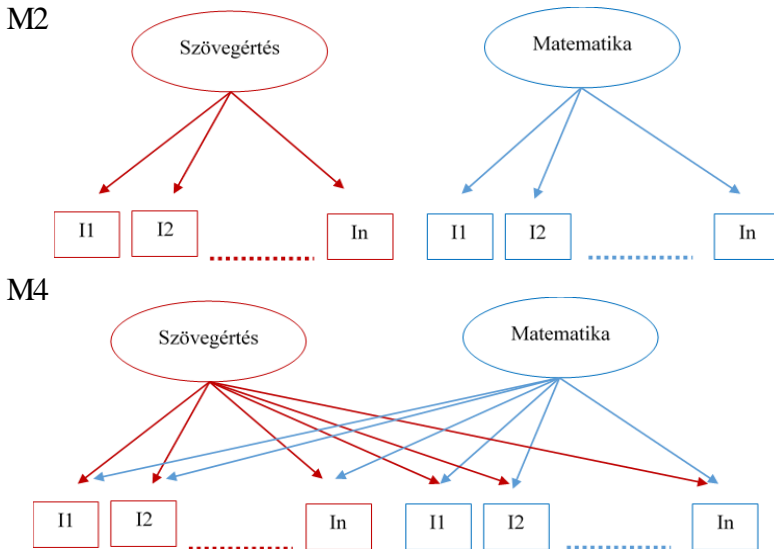
Elemzéseinket a 2017/2018-as Országos kompetenciamérésben részt vevő 10. osztályos tanulók itemszintű adatbázisán végeztük el. Azokat az itemeket töröltük, melyek egy feladat részeredményeire vonatkoztak, csak a végeredményt jelző változókat használtuk, illetve egy item esetében rendkívül alacsony válaszadási tendenciát tapasztaltunk, így azt is kivontuk az elemzésből (OP00103 - 15%-os válaszadás).

A használt eljárások feltétele, hogy ne szerepeljenek hiányzó értékek az adatbázisban, ezért azokat a tanulókat, akik az egyik teszt részt nem töltötték ki, kivontuk a vizsgálatból.

Az adatállomány nagysága így 76644 fő, melyből 2928 tanuló halmozottan hátrányos helyzetű, míg 73716 nem. A tesztsor összesen 118 itemet tartalmazott, melyből 57 matematika, 61 pedig szövegértés kompetenciát mér. A differenciált itemműködés vizsgálata során csak az 57 matematika itemet vizsgáltuk. Az adatbázisból a diákok nem beazonosíthatóak.

A vizsgált modellek

Gergely, T. Kárász és Takács (2020, kézirat) eredményeire építve a teljes tesztsorra két különböző modellt illesztettünk, az egyértelműség kedvéért a nevüket az eredeti publikációban közöltként megőriztük. A jelenleg használt mérési modell az Országos kompetenciamérésben az M2 modell. A szövegértés és a matematika itemek ebben az esetben két külön látens dimenzióra illeszkednek (Gergely, T. Kárász és Takács, 2020, kézirat). A másik használt modell az M4 modell, mely egy olyan többdimenziós IRT modell, amely figyelembe veszi az itemeken belüli multidimenzionalitást, azaz a matematika és a szövegértés itemek egyaránt súlyozódnak mindkét látens dimenzióra. Ebben az esetben a matematika feladatokban is fellelhető a szövegértés nehézség, illetve a szöveg-



1. ábra: M2 és M4 modell látens és manifeszt változóinak kapcsolat-ábrája, Gergely, T. Kárász és Takács (2020, kézirat) alapján

értés feladatokban is a matematikában használatos gondolkodási műveletek (Gergely, T. Kárász és Takács, 2020, kézirat).

Statisztikai elemzések

Az elemzéseket az R (R Core Team, 2013) MIRT (Chalmers és mtsai., 2016) programcsomagjában végeztük el. Mindkét modell esetén 2 paraméteres logisztikus modellt használtunk, a két paraméter minden item esetén a nehézség és a diszkriminancia index voltak. A modell illesztéséhez EM algoritmust használtunk, amelynél a toleranciaszintet 0,001-re állítottuk a futási idő csökkentésének érdekében.

Ahhoz, hogy a FHH kódal rendelkezőket és nem rendelkezőket össze tudjuk hasonlítani többcsoportos MIRT eljárást alkalmaztunk (multigroup MIRT), mely után

többdimenziós IRT modelleken használható DIF-et futtatunk (Chalmers, Counsell és Flora, 2016). A DIF futtatása során szekvenciális-elimináció algoritmust használtunk.

Jelen kutatásunk fő kérdése, hogy megállapítsuk azt, hogy mely változók jelzik előre, vagy vannak kapcsolatban azzal, hogy mely esetekben jelenik meg differenciált itemműködés. Ennek megállapítására egy új adatstruktúrában rögzítettük az itemek tesztfüzetben felűntetett paramétereit és két független kódoló segítségével a feladatok tartalmát csoportokba osztottuk (a kódolók konzisztenciáját inter-rater korrelációval ellenőriztük). Ezután keresztábrás elemzéseket végeztünk el.

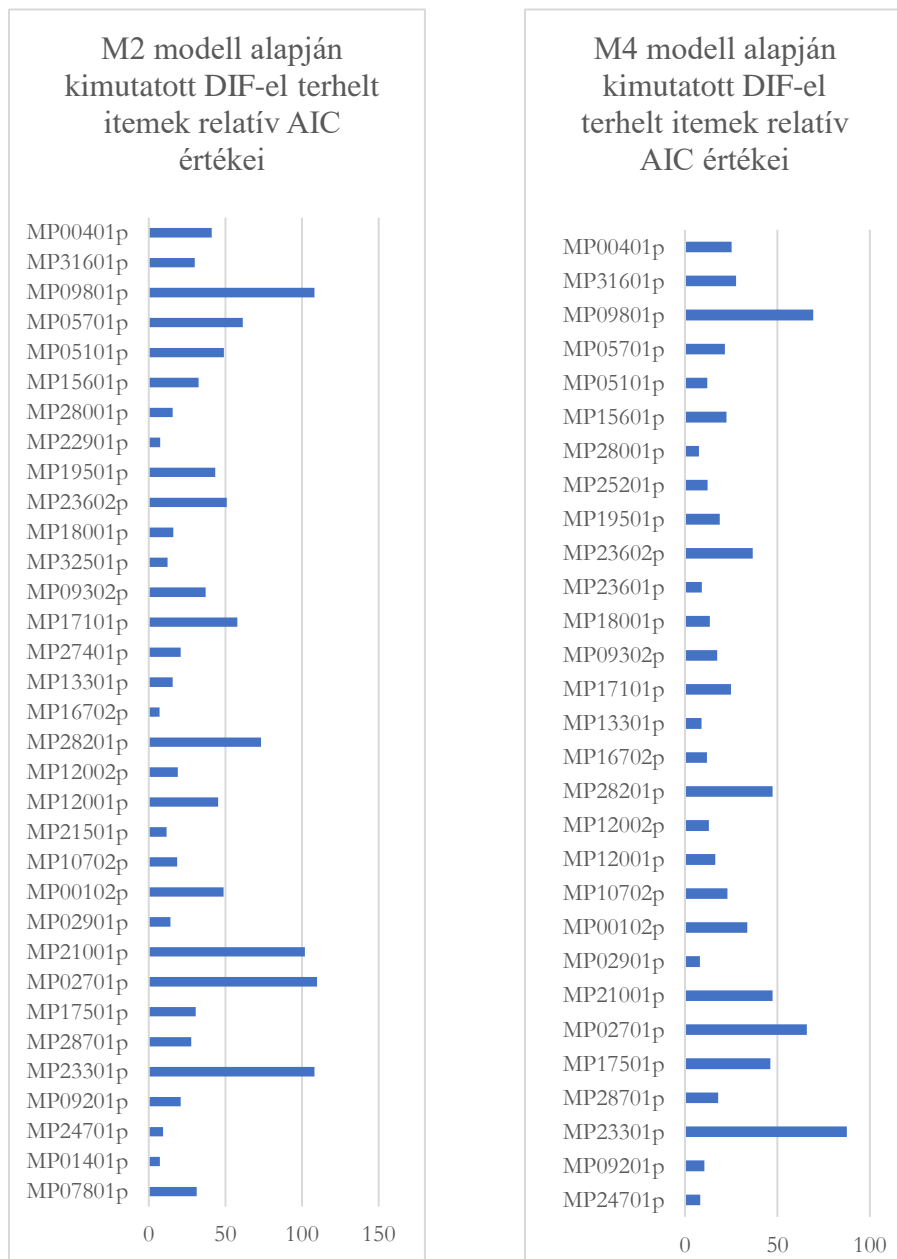
Tartalmi terület	Kognitív művelet	Nehézségi szint	Táblázat	Feladattípus	A feladat szövegeinek tartalma
Mennyiségek, számok, műveletek	Tényszerű, egyszerű műveletek	1-től 7-ig	Van	Zárt végű/feleletválasztó	Csak városban elérhető tevékenységek
Hozzárendelések, összefüggések	Alkalmazás integráció				Kulturális/politikai
Alakzatok tájékozódás	Komplex megoldások és értékelés		Nincs	Nyílt végű	Tudományos szöveg
Statisztika, valószínűség					Hétköznapi tevékenységek

2. ábra: A feladatokhoz hozzárendelt itemjelzők az Országos kompetenciamérés tantervi felelősi által, illetve a feladatok tartalmának kódolása a szerzők által

EREDMÉNYEK

A modellek illesztése 0,001 toleranciaszint mellett végbement, így a teljes modell illeszkedése megfelelő volt. Mivel a két csoport mintaelemszáma között nagy az eltérés, így az alkalmazott hátrányos helyzetűek részmintáján illesztett modell mutat kedvezőbb AIC és BIC értékeket. Az AIC az Aikake-féle információs kritérium, míg a BIC pedig a Bayes-féle információs kritérium értéket jelöli. Hagyományosan a modell illeszkedésének jószágát szokták vele mérni, minél kisebb az értéke, annál nagyobb a modell illeszkedése a mintára (ld. Vrieze, 2012).

A DIF vizsgálata esetén szekvenciális-elimináció algoritmust használtunk, mely mellett mindkét esetben BH korrekciót alkalmaztunk, így korigálva a fals pozitív eredményeket (Benjamini és Hochberg, 1995). Az M2 modell esetén 33, míg az M4 modell alkalmazása esetén 29 itemnél tartunk fel differenciált itemműködést.



3. ábra: A két modell alapján feltárt DIF-el terhelt itemek AIC értékeinek különbségének abszolút értéke a két csoport esetén.

A Khi-négyzet próba elvégzése előtt két független kódoló tartalomelemzést hajtott végre az itemeken (inter-rater reliabilitás = 0,869), ha a két kódolás eltért egymástól egy harmadik személy bevonásával döntöttünk a két választás között. A Khi-négyzet próba

eredményei alapján a vizsgált változók közül a kognitív művelettípusok és az M4 modell alapján feltárt DIF között van közepes mértékű együttjárás (Khi négyzet = 9,611, $p = 0,008$, Cramér - $V = 0,411$). A páronkénti összehasonlítás alapján a Tényszeret egyszerű műveletek típusba tartozó itemek között magasabb a valószínűsége a DIF-el rendelkező itemek előfordulásának.

	Tartalom (OH alapján)	Kog. művelet (OH alapján)	Nehézség szint (OH alapján)	Van/nincs táblázat	Nyílt/Zárt item	Tartalmi kód
M2dif	3,320 ($p=0,344$)	2,589 ($p=0,274$)	0,499 ($p=0,779$)	0,438 ($p=0,508$)	0,489 ($p=0,484$)	1,661 ($p=0,645$)
M4dif	1,628 ($p=0,653$)	9,611 ($p=0,008$)	1,450 ($p=0,484$)	0,160 ($p=0,688$)	0,144 ($p=0,7040$)	1,196 ($p=0,753$)

1. táblázat: Az itemek jellemzőinek és tartalmának összehasonlítása a DIF előfordulását rögzítő M2dif és M4dif változókkal (Khi-négyzet próbát alkalmazva)

DISZKUSSZIÓ

Jelen tanulmány fő kérdése, hogy az Országos kompetenciamérésen részt vevő halmozottan hátrányos helyzetű diákok teljesítményének melyek a fő befolyásoló tényezői és megjelenhet-e a teljesítménybeli különbség itemek szintjén is? Célunk volt, hogy feltárjuk a differenciált itemműködéssel rendelkező itemeket (DIF) és esetleges közös itemjellemzőket állapítsunk meg. Ennek alapját Kispál, Gergely, T. Kárász és Takács (2020, kézirat) kutatásának eredménye jelentette.

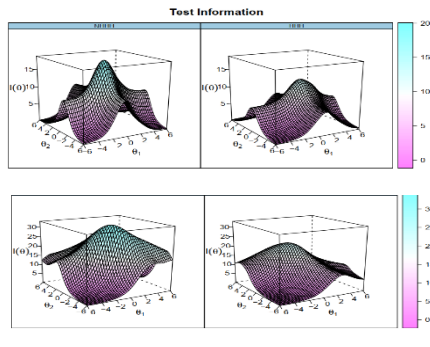
Mivel Magyarországon tanulnak a köznevelésben halmozottan hátrányos helyzetű gyermekek (az általános- és a középfokú oktatásban összesen 58023 fő (Oktatási Hivatal, 2021), fontosnak találtuk azoknak a tényezőknek a feltárását, melyek számukra nehézségeket okoznak egy tudásszintmérő vizsgálat során. A hátrányos helyzetű diákok teljesítményére ható tényezőket kutató eddigi vizsgálatok (pl.: Fejes, 2013; Mendi, 1999; Keller és Mártonfi, 2006) mindegyike különféle szempontból közelítette meg a kérdéskört, azonban egyetlen esetben sem volt arra példa, hogy itemszinten vizsgálják meg az eltéréseket a HHHH-s és nem HHHH-s diákok teljesítménye között, tehát az feladatmegoldásbeli eltérésekre helyezték a hangsúlyt.

A differenciált itemműködés vizsgálatát sok esetben végezték el nemzetközi tudásszintmérő vizsgálatok esetén (pl.: Yıldırım, 2019; Da Costa és Araujo, 2020 stb.), viszont eddig tudomásunk szerint, az Országos kompetenciamérés itemeinek esetében nem történtek ilyen vizsgálatok. A fent említett kutatások sok esetben tártak fel DIF-es itemeket a TIMSS és a PISA tesztekben.

A statisztikai elemzéseket az Országos kompetenciamérés 2017/2018-as 10. évfolyamos diákjainak itemszintű adatbázisán végeztük el. Két modellt alkalmaztunk, melyeket Gergely, T. Kárász és Takács (2020) kézírata alapján választottunk, ezen modellek az M2 (a jelenleg is a kompetenciamérésben használt mérési modell) és az M4 (többdimenziós IRT modell, mely figyelembe veszi az itemeken belüli multidimenzionalitást) modellek voltak.

Többcsoportos MIRT eljárást használtunk a statisztikai elemzések során, mely után többdimenziós IRT modelleken használható DIF-et alkalmaztunk (Chalmers, Counsell és Flora, 2016) és szekvenciális elimináló algoritmust. Ahhoz, hogy Khi-négyzet próbát hajthassunk végre, két független kódoló segítségével tartalmilag kódoltuk és az itemek paramétereit és egyéb információit, illetve, hogy tapasztalható-e DIF egy külön adattáblában rögzítettük.

Az M2 és M4 modell esetében közel egyforma mértékben jelentek meg DIF-el terhelt itemek, ugyanakkor csak az M4 modell segítségével tudjuk feltámi, hogy mi az eltérések oka. Míg az M4 modell esetén a nem HHH-s csoportnál mindkét látens változóról hozzávetőlegesen egyenletesen biztosít információt, addig a HHH-s csoport esetében enyhe eltolódást látunk a szövegértési kompetenciát mérő dimenzió irányában. Ez a különbség M2 modell esetén nem látható. Az M4 megmutatja, hogy a két csoport esetén a feladat más képességeket mér.



4. ábra: Tesztinformáció görbe M2 és M4modell esetén

Az item jellemzők és a tartalmi kódolás vizsgálatakor csak a Tényismeret és egyszerű műveletek kognitív művelési területen mutatható ki szignifikánsan magasabb esély arra, hogy az adott item DIF-el terhelt legyen. Ennek magyarázata lehet, hogy sok esetben a halmozottan hátrányos helyzetű diákok egyszülős családban nőnek fel, nem rendelkeznek értelmiségi szerepmintával, így nem tudják ezeket a képességeket elsajátítani (Mendi, 1999; Horváth, 2015). Az iskolai szegregáció is állhat a lemaradás hátterében (Horváth, 2015).

Összegezve elmondható, hogy a halmozottan hátrányos helyzetű diákok a matematika feladatok megoldásakor bizonyos esetekben a szövegértés tudásukra támaszkodnak. Érdeemes kiemelni, hogy a feladatok közül a Tényismeret és egyszerű műveletek kognitív művelési területen van a legnagyobb esély DIF megjelenésére halmozottan hátrányos helyzetű diákok esetén, így a pedagógusok nagyobb hangsúlyt fektethetnek ezen

feladatok gyakorlására és megértetésére a halmozottan hátrányos helyzetű tanulók esetében. A későbbiekben a vizsgálati eredményeink pontosításához célszerű lehet kvalitatív vizsgálatok végrehajtása is, hogy részletesebben feltárhatóak legyenek a feladatok, melyek nehézséget okoznak a HHH-s diákoknak.

BIBLIOGRÁFIA

1997. évi XXXI. tv. a gyermekek védelméről és a gyámsági igazgatásról. Letöltve: 2021. 04. 27. <https://netjogtar.hu/jogszabaly?docid=99700031v>
- Balázs I., Balkányi P., Ostorics L., Palincsár I., Rábainé Szabó A., Szepesi L., Szipőcsné Krolopp J. & Vadász Cs. (2014). Az Országos Kompetenciamérés Tartalmi Keretei: Szövegértés, matematika, háttérkérdőívek. Budapest: Oktatási Hivatal
- Beese, J. & Liang, X. (2010). Do resources matter? PISA science achievement comparisons between students in the United States, Canada and Finland. *Improving Schools*, 13(3), 266-279.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 57, 289-300.
- Buda, M. (2017). Hátrányos helyzet és iskola. Merre keressük a kiutat a motívatlanság csapdájából?. In Polonyi, T., & Abán, K. (eds). *Digitális tanulás és tanítás*. (pp. 9 - 26). Debreceni Egyetem.
- Belinszki, B., Szepesi, L., Takácsné Kárász, J., & Vadász, Cs. (2019). Országos kompetenciamérés: Országos jelentés. Budapest: Oktatási Hivatal.
- Cs. Czachesz, E. & Radó, P. (2003). Oktatási egyenlőtlenségek és speciális igények. In Falás, G. & Lannert, J. (szerk). *Jelentés a magyar képzésről 2003*. (pp. 351 - 376). Budapest: Országos Közoktatási Intézet.
- Chalmers, P., Pihlkin, J., Robitzsch, A., Zoltak, M., Kim, K., Falk, C. F. & Oguzhan, O. (2016). Package: mirt: Multidimensional Item Response Theory - Version 1.29.
- Chalmers, R. P., Counsell, A., & Flora, D. B. (2016). It might not make a big DIF: Improved differential test functioning statistics that account for sampling variability. *Educational and Psychological Measurement*, 76(1), 114-140.
- Csapó B., Fejes J. B., Kinyó L., & Tóth E. (2014). Az iskolai teljesítmények alakulása Magyarországon nemzetközi összehasonlításban. In Kólosi Tamás & Tóth István György (szerk). *Társadalmi riport 2014 - Tanulmánykötet* (pp. 110 - 136). Budapest: TÁRKI.
- Da Costa, P. D., & Araújo, L. (2012). Differential item functioning (DIF): What functions differently for Immigrant students in PISA 2009 reading items. *JRS Scientific and Policy Reports*. Luxembourg: European Commission.
- Deutsch, M. (1964). Early social environment: Its influence on school adaptation. In Schreiber, D. (ed). *The school dropout* (pp. 89 - 100). Washington DC: National Educational Association
- Ehmke, T., van den Ham, A. K., Sälzer, C., Heine, J., & Prenzel, M. (2020). Measuring mathematics competence in international and national large scale assessments: Linking PISA and the national educational panel study in Germany. *Studies in Educational Evaluation*, 65, 100847.
- Faluvégi, A. (2004). A társadalmi-gazdasági jellemzők területi alakulása és várható hatásai az átmenet időszakában. Műhelytanulmányok Discussion Papers, 2004/5. Budapest: MTA Közgazdaságtudományi Kutatóközpont.
- Fejes, J. B. (2006). Miért (nem) fontosak a hátrányos helyzetű tanulók?. *Új Pedagógiai Szemle*, 56(7-8), 17-26.
- Fejes, J. B. (2013). Miért van szükség deszeregációra?. In Fejes József, Balázs & Szűcs Norbert (szerk). *A szegedi és békéscsabai deszeregációt támogató Hallgatói Mentorprogram. Öt év tapasztalatai*. (pp. 15 - 35). Szeged: Belvedere-Medionale.
- Gergely, B., Takácsné Kárász, J., Takács, Sz. (2020). Többdimenziós IRT modellek alkalmazása az Országos Kompetenciamérésben (IDK, kézirat)
- Harsányi, Sz, G. Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019). Az iskolai teljesítménykülönbség és a szocioökonómiai státusz összefüggései – az országos kompetenciamérés eredményeinek vizsgálata a szülők munkájának rendszeressége, az észlelt társadalmi helyzet és a lakókörnyezet vonatkozásában, *Psychologia Hungarica Carolinensis*, 16, 148-221.
- Harsányi, Sz, G. Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019). Születni tudni kell: Az Országos kompetenciamérés eredményeinek vizsgálata a szülők munkájának rendszeressége, észlelt társadalmi helyzet és a lakókörnyezet vonatkozásában, *Psychologia Hungarica Carolinensis*, VII/2, 64-85.
- Harsányi, Sz, G. Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Smohai, M., Simon, G., Takács, N., Takács, Sz. (2019). The relationship of school achievement with parents' employment status, perceived

- social status, and living environment as reflected in findings of the 2017 National Assessment of Basic Competencies (NABC), *Psychologia Hungarica Caroliensis*, VII/2, 167-189.
- Holland, P. W., & Wainer, H. (2012). *Differential item functioning*. Routledge.
- Horváth, Á. (2015). Hátrányos helyzetű diákok tanulási motivációi. *Studia Mundi - Economica*, 2(3), 97-112.
- Huang, X., Wilson, M., & Wang, L. (2014). Exploring plausible causes of differential item functioning in the PISA science assessment: language, curriculum or culture. *Educational Psychology*, 36(2), 378-390.
- Kasik, L., Gutí, K., & Gáspár, Cs. (2014). Hátrányos és nem hátrányos helyzetű diákok szociálisprobléma-megoldó gondolkodása. *Magyar Pedagógia*, 114(1), 49-63.
- Keller, J., & Mártonfi, G. (2006). Oktatási egyenlőtlenségek és speciális igények. In Halász G. & Lannert J. (szerk). *Jelentés a magyar köznevelésről 2006.* (pp. 377 - 411). Budapest: Országos Közoktatási Intézet.
- Kerteszi, G., & Kézdi, G. (2012). A roma és nem roma tanulók teszteredményei közti különbségekről és e különbségek okairól. *Budapesti Munkagazdaságtani Füzetek*, 5, 1-84. Budapest: Magyar Tudományos Akadémia Közgazdaság- és Regionális Tudományi Kutatóközpontjának Közgazdaságtudományi Intézetében működő Munkapiaci Kutatások, Budapesti Corvinus Egyetem Emberi Erőforrások Tanszéke.
- Kispál, S., Gergely, B., Takácsné Kárász, J. & Takács, Sz. (2020). Eltérő itemműködés vizsgálata az Országos kompetenciamérésben halmozottan hátrányos helyzetű diákok körében többdimenziós IRT modellek segítségével (IDK, kézirat).
- Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms bei der DFG. *Zeitschrift für Pädagogik*, 52, 876 - 890.
- Koltói, L., Harsányi, Sz. G., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019a). A szülői tanulmányokba való bevonódásának összefüggése az iskolai teljesítménnyel, *Psychologia Hungarica Caroliensis*, VII/2, 86-103.
- Koltói, L., Harsányi, Sz. G., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019b). Az iskolai szülői bevonódás iskolai szintű vizsgálata megyei és regionális szinten az országos kompetenciamérés 2017-es és 2018-as adatai alapján, *Psychologia Hungarica Caroliensis*, 16, 222-258.
- Koltói, L., Harsányi, Sz. G., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019c). The relationship between school achievement and paternal involvement in children's school activities as judged by headmasters in the 2017 National Assessment of Basic Competencies (NABC), *Psychologia Hungarica Caroliensis*, VII/2, 190-212.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale NJ: Lawrence Erlbaum Associates.
- McCandless, B. (1952). Environment and intelligence. *American Journal of Mental Deficiency*, 56, 674 - 691.
- Mendi, R. (1999). Felsőoktatásban tanuló roma fiatalok pályaszocializációs és személyiségvizsgálata. *Bölcsészleírás szakdolgozat*, ELTE Pszichológiai Intézet. Letöltve: 2021. 04. 27. <http://mekoszk.hu/02000/02034/html/index.htm>
- Nyitrai, E., Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Mátay, G., Nagybányai-Nagy, O., Pusker, M., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019a). Szülői bevonódás és az iskolai teljesítmény kapcsolata az országos kompetenciamérés 2017-es és 2018-as adatainak tükrében, *Psychologia Hungarica Caroliensis*, 16, 7-51.
- Nyitrai, E., Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019b). Iskolai teljesítmény és szülői bevonódottság, *Psychologia Hungarica Caroliensis*, VII/2, 7-28.
- Nyitrai, E., Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019c). Relations between Parental Involvement and School Performance in the Light of Data from National Assessment of Basic Competencies (NABC) 2017, *Psychologia Hungarica Caroliensis*, VII/2, 115-130.
- Oktatási Hivatal: Letölthető köznevelési adatok, statisztikai kimutatások. Letöltve: 2021. 04. 27. https://data.oktatas.hu/kozerdeku_index
- Országos Kompetenciamérés Technikai Leírás (2012). Letöltve: 2020. 08. 17. https://www.oktatas.hu/pub_bin/dload/kozoktat/meresek/unios_tanulmanyok/OKM_Technikaleiras.pdf
- R Core Team (2013). R: A language and environment for statistical computing.
- Radó, P. (1997). *Jelentés a magyarországi cigány tanulók oktatásáról, szakértői tanulmány a Nemzeti Etnikai Kisebbségi Hivatal számára*. Budapest.
- Rajnai, J. (2012). Lehet másképp?! – Avagy pedagógiai alternatívák a hátrányos helyzetű roma fiatalok nevelésében. *Új Pedagógiai Szemle*, 62(11-12), 54–75.
- Siller, J. (1957). Socioeconomic status and conceptual thinking. *Journal of Abnormal and Social Psychology*, 55, 365 - 371.
- Shanmugam, S. K. S. (2020). Gender-Related Differential Item Functioning of Mathematics Computation Items among Non-native Speakers of English. *The Mathematics Enthusiast*, 17(1), 108-139.
- Szabó, L.D., Szepesi, L., Takácsné, K. J., & Vadász, Cs. (2018). *Országos Kompetenciamérés 2017 Országos jelentés*. Budapest: Oktatási Hivatal

HÁTRÁNYOS HELYZETŰ DIÁKOK TÖBBDIMENZIÓS IRT MODELL VIZSGÁLATA

- T. Kárisz, Judit (2019a): Hibabecslési eljárások véletlen jelenségek paramétereinek becslésére, *Psychologia Hungarica Caroliensis*, VII/2, 104-114.
- T. Kárisz, Judit (2019b): Estimation methods on standard error of different statistical parameters, *Psychologia Hungarica Caroliensis*, VII/2, 213-220.
- Trier, U. P., & Peschar, J. (1995). Cross-Curricular Competencies: Rationale and Strategy for Developing a New Indicator. In OECD (ed.), *Measuring what Students Learn* (pp. 99 - 109). Paris: OECD
- Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods*, 17(2), 228.
- Walker, C. M., & Beretvas, S. N. (2003). Comparing multidimensional and unidimensional proficiency classifications: Multidimensional IRT as a diagnostic aid. *Journal of Educational Measurement*, 40, 255 - 275.
- Weinert, F. E. (2001). Concept of competence: a conceptual clarification. In Rychen, D. S., & Salganik, L. H. (eds), *Defining and selecting key competencies* (pp. 5 - 65). Seattle: Hogrefe & Huber Publishers.
- Yıldırım, O. (2019). Detecting Gender Differences in PISA 2012 Mathematics Test with Differential Item Functioning. *International Education Studies*, 12(8), 59 - 71.

DIFFERENTIAL ITEM FUNCTIONING IN THE NATIONAL ASSESSMENT OF BASIC COMPETENCIES AMONG DISADVANTAGED CHILDREN USING MULTI-DIMENSIONAL IRT MODELS

Sára Kispál, Bence Gergely, Judit Takácsné Kárász, Szabolcs Takács

Correspondence author: Sára Kispál (sarakispal@gmail.com)

Abstract

The aim of our research is to map the performance of disadvantaged students in the National Assessment of Basic Competencies, which is based on the research results of Kispál, Gergely, T. Kárász and Takács (2020, manuscript). Disadvantaged students participate in the National Assessment of Basic Competencies every year, but their performance lags behind the average. Our main objectives are to examine whether there is a difference at the levels of items in the problem solving of students with disadvantages, whether there are tasks that are specifically more difficult or easier for them, and whether they can compensate for their shortcomings in one area of competence with another.

Examining two statistical models (hereinafter M2, M4 models) we found that in both cases there is a differential item functioning; however, the M4 model shows that the test information curve provides much more information about the reading comprehension dimension than the mathematics dimension. At the same time, the maximum information content lags significantly behind the non-disadvantaged group. When examining item characteristics and content coding, a significantly higher chance of the emergence of differential item functioning can be observed only in the field of factual knowledge and simple operations.

In summary, disadvantaged students use their reading comprehension competencies much more when solving math problems. In the case where the task does not require more mathematical knowledge than reading comprehension competencies, it does not cause differential item functioning, while otherwise it does. The authors carried out the research on which the study is based in the framework of the National Assessment of Basic Competencies Research Group 20642B800 tender, funded by the Faculty of Arts and Social Sciences of the Károli Gáspár University of the Reformed Church in Hungary.

Keywords: disadvantaged students ■ National Assessment of Basic Competencies ■ multidimensional IRT models ■ differential item functioning

INTRODUCTION

Disadvantaged students participate in the National Assessment of Basic Competencies every year, although their performance lags behind their peers'. The reasons behind this were investigated in several researches (see Mendi, 1999; Buda, 2017; Fejes, 2013 etc.) in which the lack of intellectual role pattern was listed as the most influential factor that affects their performance (Mendi, 1999). The absence of stable financial background

(Kertesi & Kézdi, 2012) and the stable family structure (Kasik et al., 2014) were also listed as one of the influential factors.

The main objective of the present study is to re-evaluate the view that students with cumulative disadvantages perform worse on the National Assessment of Basic Competencies. Is there any item-level differentiation? Are there tasks in which these children perform better than their peers? Can they compensate their shortcomings in one area of competence with their knowledge in another area?

THE CONCEPT OF COMPETENCE, THE NATIONAL ASSESSMENT OF BASIC COMPETENCIES AND STUDENTS WITH MULTIPLE DISADVANTAGES IN HUNGARY

The concept of competence can be approached from several perspectives. In everyday terms, competence means suitability, but the creation of the concept of competence used in education can be linked to the PISA test. According to this, competence measurement should be based on the measurement of skills that are relevant to students in their daily lives (Trier & Peschar, 1995). According to the currently official concept of competence, competence is a context-specific cognitive disposition that is necessary for the management and solution of a given situation (Weinert, 2001; Klieme & Leutner, 2006).

International knowledge assessments and the National Assessment of Basic Competencies

From the beginning of the 1970s, international knowledge assessments (IEA, TIMSS, PIRLS, PISA) were present in Hungary. The main aim of the surveys was to get feedback on the educational situation in a country at an international level and to create ranking based on students' results (Csapó, Fejes, Kinyó & Tóth, 2014).

Since 2001, the National Assessment of Basic Competencies has been held every spring organized by the Educational Authority. "The primary goal of competency measurement is to inform the education government about the effectiveness of primary and secondary education in terms of learning outcomes and equal opportunities. In addition, the developers of competence measurement consider it their task to provide schools and their maintainers with the feedback needed for local evaluation and planning of appropriate interventions, with data that can be further analysed (maintainer-, institutional- and site reports)." (National Assessment of Basic Competencies Technical Description, 2012, p. 8)

Sixth, eighth and tenth grade students take part in the National Assessment of Basic Competencies every year. The date of the test is set uniformly by the Education Office. There are two parts to the measurement: reading comprehension and mathematical tests for students to complete, both of which have two times forty-five minutes for complete with a little break in between. The planning of the task sets was based on the concepts of mathematical tool knowledge and text comprehension, according to which

the emphasis in mathematics task sets is on real-life application, while on text comprehension the emphasis is on decoding, comprehension and reflection (Szabó, Szepesi, Takácsné & Vadász, 2017).

The test tasks undergo a pre-screening each year, so that only those tasks that meet the pre-selection procedure are included in the final task list (one task must meet several criteria: eg. they must be lifelike etc.). Variety of task types is also important, as otherwise tasks will be easy to practice and measurement will be inaccurate (Weinert, 2001). Tasks avoid the use of brand names, country names and political content, they also avoid misleading tasks, too long or too short questions. Tasks must be classifiable in all cases (National Assessment of Basic Competencies Technical Description, 2012).

In our research, we examine the differences in the operation of items in mathematical problems. In the National Assessment of Basic Competencies, mathematics tasks are developed based on the concept of mathematical tool knowledge, which includes “the ability of the individual to understand and analyze the role of mathematics in the real world; skill-level use of the math toolbox; the need to be able to apply the acquired mathematical knowledge in real life situations and the ability to do so; the use of mathematical tools in social communication and collaboration at an age-appropriate level.” (Balázi et al., 2014, p. 33). The test includes tasks from different areas of knowledge: quantities, numbers, and operations; assignments, relationships; shapes and orientation; statistical characteristics and probability calculations. In addition, the tasks can be grouped based on three types of thinking operations: factual knowledge, application and integration, and complex solutions and evaluation (Balázi et al., 2014). This structure of the tasks is very similar to the structure of the PISA measurement (Ehmke, van den Ham, Sälzer, Heine & Prenzel, 2020).

Cumulatively disadvantaged children in Hungary

The definition of cumulative disadvantage is set out in Act XXXI of 1997. Act on the Protection of Children and Guardianship Administration. According to it, “Section 67/A § (1) A disadvantaged child is a child entitled to regular child protection allowance and an adult who has one of the following circumstances: a) low education of the parent or foster parent (...) b) low employment of the parent or foster carer (...) c) insufficient living environment or housing conditions of the child (...) where the conditions necessary for healthy development are limited. (2) A cumulatively disadvantaged is a) a child entitled to a regular child protection allowance and an adult who has at least two of the conditions specified in subsection (1) a)-c), b) a child in foster care, c) a young adult in care and in a pupil or student relationship.”

Studies investigating the performance of multiple disadvantaged students have identified several influencing factors. The most important of these are the impact of socioeconomic status, the lack of an intellectual role model, loss of motivation, school segregation and territorial differences within the country, and possible differences in cognitive functioning.

Socioeconomic status considerably influences students’ school performance. Students with higher SES perform better in standardized tests than their peers with lower

SES (Beese & Liang, 2020). In their research, Kertesi and Kézdi (2012) compared the test results of roma and non-roma youth. It was found that differences in income, education and residence are the main reasons for the lag, and the influence of financial background has a disproportionately large impact on student performance. It is also important that for those living in deep poverty, income growth results in a much greater improvement than in the middle or upper class of society.

The poorer performance of cumulatively disadvantaged students is influenced by the fact that in many cases these students do not receive an intellectual role pattern within their family (Mendi, 1999 quotes Horváth, 2015). Students often grow up in a mosaic family, which negatively affects their social problemsolving abilities, which appears mainly in adolescence (Kásik et al., 2014).

Cumulatively disadvantaged students often encounter negative prejudices in the school, even from teachers, which also reduces their performance (Radó, 1997 quoted by Horváth, 2015).

Loss of motivation is common among cumulatively disadvantaged students. This is often explained by the structure, academic and theoretical nature of the curriculum, and the fact that the patterns seen in school do not correspond to the student's previously known "picture of reality". Because cumulatively disadvantaged children were often brought up in a different culture, most of the things they experienced at school are not familiar to them, so they ignore them (Buda, 2017).

Segregation is also a factor that has a negative impact on the performance of cumulatively disadvantaged students. As segregation is extremely common in public education in Hungary (Fejes, 2013), it greatly influences the performance of Hungarian students. In the present case, we do not mean the segregation of minorities, but the homogeneity of the family background of children attending a class or school (Fejes, 2013). Many parents enrol their children in schools outside the district where they live, to receive a better education, so segregation becomes more prevalent (Keller & Mártonfi, 2006). The performance of students with lower socio-economic status is positively affected by attending an institution in which the family background of the majority can be considered good, but this has not yet realised in Hungary (Fejes, 2006; Cs. Czachesz & Radó, 2003).

Segregation was facilitated by the fact that with the development of the market economy, regional differences in Hungary became larger and larger (Faluvégi, 2004). The country has become Budapest-centric, and the performance of students in less developed regions is also lower than in the capital (this is typically in the region of Eastern Hungary) (Faluvégi, 2004 quotes Keller & Mártonfi, 2006). In the less developed regions of the country, the unemployment rate is much higher, so there are many more cumulatively disadvantaged students (Rajnai, 2012).

Cognitive dysfunctions cause many students to lag behind their peers, so the question may arise as to whether such a discrepancy may underlie the performance of cumulatively disadvantaged students. In many cases, cumulatively disadvantaged children lag behind their peers in categorizing abstract conceptualization and visual stimulants (Deutsch, 1964). According to Siller (1957) the lag in abstraction is caused by the fact that students with higher socioeconomic status use abstraction at a much higher rate

than their peers with lower SES. Cumulatively disadvantaged students are less flexible in solving tasks, preferring to think more specifically compared to their peers (McCandles, 1952).

DIFFERENTIAL ITEM FUNCTIONING AND KNOWLEDGE ASSESSMENTS

In order to compare different groups in terms of the results of a knowledge assessment we need to examine whether each item in the test is of equal difficulty to the members of the group. This finding is usually explored using differential item function methods (Holland & Wainer, 2012). DIF appears when test fillers who have the same level of a given latent property but come from a different subpopulation have a significantly different probability of solving a given item correctly, i. e. the item probability curve of the two groups is different in the difficulty or in terms of the discrimination index (or their interaction) (Huang, Wilson & Wang, 2014; Chalmers, Counsel & Flora, 2016).

In the case of international knowledge assessments (eg. PISA, TIMSS), several researches were carried out in which differential item functioning was examined. In his research, Shanmugam (2020) examined the mathematics issues of the TIMSS test. Its purpose was to explore: if the functioning of the items depend on the gender of the student He found differential item functioning for girls in tasks that required higher thinking skills, while for boys, tasks containing fractions and negative numbers proved more difficult.

Da Costa and Araújo (2012) examined the results of the PISA comprehension test for immigrant and non-immigrant students. Their assumption was that immigrant students have difficulties solving a task that has cultural implications for their current homeland. Their results found that there are tasks that immigrant students solve more easily than their non-immigrant counterparts. Typically, they had an advantage in tasks that were designed to educate, transfer knowledge, and relate to situations that they could master at school.

MULTIDIMENSIONAL IRT MODELS AND DIFFERENTIAL ITEM FUNCTIONING

As the definition of competence is a complex phenomenon, the models of modern probability test theory (later IRT (item response theory)) are often used in the examination of measuring instruments. The essence of this is that for each item, a different level of a latent dimension can be used to determine the probability that the individual will respond correctly (Lord, 1980). Because many problems have arisen with the use of the traditional one-dimensional IRT models for measuring competence, it is worthwhile to use multidimensional IRT models. The sustainability of this was illustrated by Walker et

al. (2008). In their research they examined the DIF that might be found in mathematics items with text, in the case where the grouping variable is the student's level of comprehension. It was found that the level at which a student reads a math problem is affected by his or her reading. Because the measurement of comprehension also appears in math problems, it is difficult to separate the two dimensions, but Walker et al. (2008) successfully separated these dimensions with a version of a semi-exploratory IRT model.

The aim of our research is to examine the items with a DIF (Kispál, Gergely, T. Kárász & Takács, 2020, manuscript) that have already been explored among the mathematics items of the National Assessment of Basic Competencies, whether there is an item or a content characteristic that is more likely to occur DIF. In our study we also examine these differences in the case of a multidimensional IRT model (Gergely, T. Kárász & Takács, 2020, manuscript), as we believe that there are items for which traditional methods do not reveal DIF. This could be mainly due to the student compensating for the absence of one ability with the other ability.

RESEARCH BACKGROUND

The present study was created in connection with the work and research of the Competence Measurement Research Group of Károli Gáspár University of the Reformed Church in Hungary. The research is partly implemented within the framework of a tender funded by the Faculty of Arts and Social Sciences of Károli Gáspár University of the Reformed Church in Hungary (tender number: 20642B800). The antecedents of the dissertation include the following publications: on the family-level aspects of parental background, Nyitrai et al. made their results available in two studies (Nyitrai et al., 2019a; 2019b), and the study is also available in English (Nyitrai et al., 2019c). School-level aspects of parental involvement are reported by Koltói et al. in two studies written in Hungarian and one in English (Koltói et al., 2019a; 2019b; 2019c). It is important to emphasize that in addition to parental involvement, school and family background and socio-cultural status are also associated with significant differences in school performance. In connection with the National Assessment of Basic Competencies, Harsányi et al. also write about these differences in their two Hungarian and one English language materials (Harsányi et al., 2019a; 2019b; 2019c). Furthermore, it should be mentioned that the official results of the National Assessment of Basic Competencies include methodologically different approaches. In our case, we did not want to shed light on the direct results, but on the connections between the background phenomena, so we did not have to rely on the fundamentally more complicated mathematical-statistical methodology, which is available to readers in both Hungarian and English in Judit T. Kárász's summary work (T. Kárász, 2019a; 2019b).

METHOD

Sample and measuring instrument used

Our analyses were performed on the item-level database of 10th grade students participating in the 2017/2018 National Assessment of Basic Competencies. Items that related to the partial results of a task were deleted, only the variables indicating the final results were used, and an extremely low response trend was observed for one item, so it was also excluded from the analysis (OP00103 - 15% response).

The condition of the procedures used was that no missing values were included in the database, so students who did not complete one of the test sections were excluded from the study.

The size of the data set is thus 76 644 people, of which 2 928 students are cumulatively disadvantaged, while 73 716 are not. The test series contained a total of 118 items, of which 57 measured math and 61 reading comprehension competencies. During the examination of differential item functioning only the 57 mathematics items were examined. Students cannot be identified from the database.

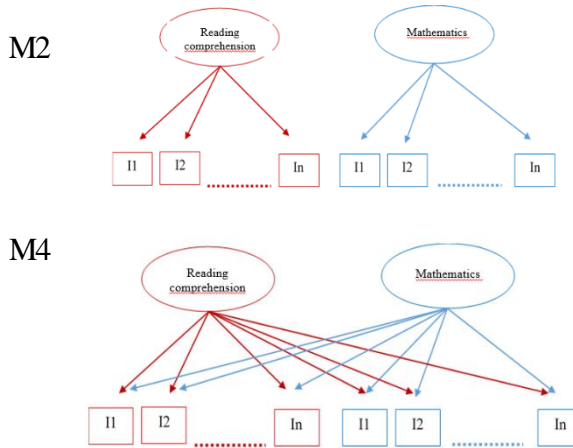
Used models

Figure 1: Relationship diagram of latent and manifest variables of M2 and M4 models, based on Gergely, T. Kárász és Takács (2020, manuscript)

Building on the results of Gergely, T. Kárász and Takács (2020, manuscript), we fitted two different models to the entire test series, and for the sake of clarity, we retained their names as published in the original publication. The currently used measurement model in the National Assessment of Basic Competencies is the M2 model. The reading comprehension and mathematics items in this case fit into two separate latent dimensions (Gergely, T. Kárász & Takács, 2020, manuscript). The other model used is the M4 model, which is a multidimensional IRT model that considers multidimensionality

within items, i.e. both math and reading comprehension items are weighted for both dimensions. In this case, the difficulty of reading comprehension can also be found in mathematics tasks, and the thinking operation used in mathematics in reading comprehension tasks can also be found (Gergely, T. Kárász & Takács, 2020, manuscript).

Statistical analyses

Analyses were performed in the MIRT (Chalmers et al., 2016) software package of R (R Core Team, 2013). For both models, a 2-parameter logistic model was used, with the two parameters for each item being the difficulty and the discretion index. To fit the model, we used an EM algorithm with the tolerance level set to 0.001 to reduce the running time.

To compare those with and without a CDC (cumulatively disadvantaged children) code, we used a multigroup MIRT procedure, followed by a DIF for use in multidimensional IRT models (Chalmers, Counsel & Flora, 2016). A sequential elimination algorithm was used to run DIF.

The main question of our present research is to determine which variables predict or are related to which cases differential item functioning occurs. To establish this, we recorded the parameters of the items indicated in the test booklet in a new data structure and divided the content of the tasks into groups with the help of two independent encoders (the consistency of the encoders was checked by inter-rater correlation). Cross-tabulation analyses were then performed.

Content area	Cognitive operation	Level of difficulty	Spread sheet	Type of exercise	Content of the exercise's text
Quantities, numbers, operations	Factual knowledge, simple operations	1 to 7	There is	Close-ended, multiple choice	Activities only available in the city
Assignments, relationships	Application Integration				Cultural/politics
Shapes, orientation	Complex solutions and evaluation		There is not	Open-ended	Scientific text
Statistics, probability		Daily activities			

Figure 2 Item characteristics assigned to the tasks by the subject managers of the National Assessment of Basic Competencies, and coding of the content of the tasks by the authors

RESULTS

The fit of the models took place at a tolerance level of 0.001, so the fit of the model was correct. Due to the large difference between the sample sizes of the two groups, the model fitted to the sub-sample of the cumulatively disadvantaged shows more favourable AIC and BIC values. AIC is the Akikake information criterion and BIC is the Bayes

information criterion. Traditionally, the goodness of fit of a model has been measured with it, the smaller its value, the greater the fit of the model to the sample (see Vrieze, 2012).

In the case of DIF, we used a sequential-elimination algorithm, in which case BH correction false positive results results (Benjamini & Hochberg, 1995). Differential item functioning was revealed for 33 items in the case of the M2 model and 29 items in the case of the M4 model.

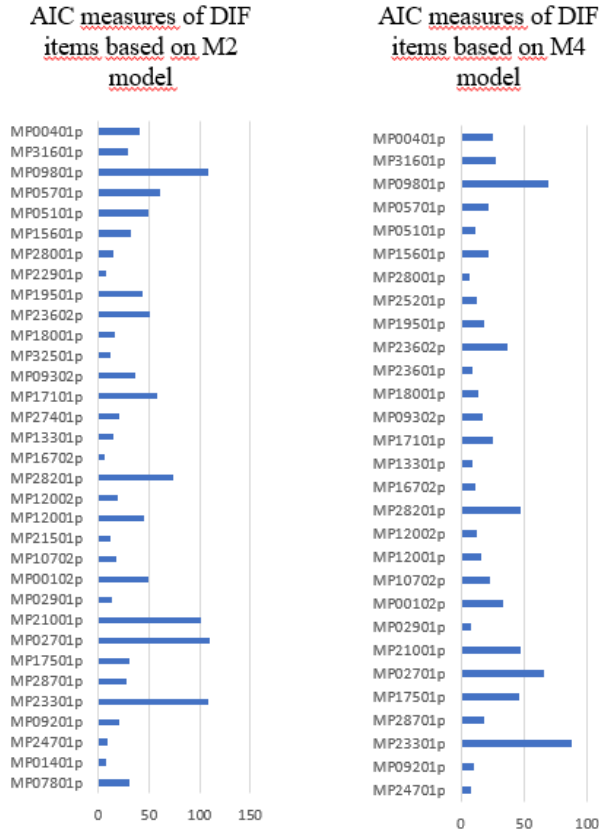


Figure 3: The absolute value of the difference in AIC values of the DIF-loaded items explored based on the two models of the two group

Before performing the Chi-square test, two independent coders performed content analysis on the items (inter-rater reliability = 0,896), if the two codings differed from each other, we decided between the two choices with the involvement of a third person. Based on the Chi-square test, there is a moderate correlation between the cognitive procedure types and the DIF revealed based on the M4 model (Chi-square = 9,611; p =

0,008; Cramer-V = 0,411). Based on the pairwise comparison, items belonging to the Factual Knowledge simple operations type are more likely to have items with a DIF.

	Content	Cog. Operation	Difficulty level	There is/there is not a spreadsheet	Open/closed item	Content code
M2 DIF	3,320 ($p=0,344$)	2,589 ($p=0,274$)	0,499 ($p=0,779$)	0,483 ($p=0,508$)	0,489 ($p=0,484$)	1,661 ($p=0,645$)
M4 DIF	1,628 ($p=0,653$)	9,611 ($p=0,008$)	1,450 ($p=0,484$)	0,160 ($p=0,688$)	0,14 ($p=0,704$)	1,196 ($p=0,753$)

Figure 4 Comparison of the characteristics and content of the items with the variables M2dif and M4dif, which record the occurrence of DIF (using Chi-square test)

DISCUSSION

The main question of the present study is to discover the main influencing factors of the performance of the cumulatively disadvantaged students participating in the National Assessment of Basic Competencies. Our aim was to explore items with differential item functioning (DIF) and to identify possible common item characteristics. This is based on the results of the research of Kispál, Gergely, T. Kárász & Takács (2020, manuscript).

As a negligible number of children in Hungary are cumulatively disadvantaged in public education, we found it important from a pedagogical point of view to explore the factors that cause them difficulties in a knowledge assessment. All previous studies of factors affecting the performance of disadvantaged students (e. g. Fejes, 2013; Mendi, 1999; Keller & Mártonfi, 2006) have approached the issues from different perspectives, but in no case have there been any examples of exploring differences at the item level between the performance of CDC and non-CDC students.

The examination of differential item functioning has been performed in many cases in the case of international knowledge assessments (e.g. Yildirim, 2019; Da Costa & Araújo, 2020 etc.), but so far no such tests have been performed for the items of the National competence measurement. The research mentioned above has in many cases revealed DIF items in the PISA and TIMSS tests.

Statistical analyses were performed on the item-level database of 10th year students of the National Assessment of Basic Competencies in 2017/2018. We used two models, which were chosen based on the manuscript of Gergely, T. Kárász and Takács (2020). These models are M2 (the measurement model currently used in competence measurement) and M4 (multidimensional IRT model, which considers the multidimensionality within items).

A multigroup MIRT procedure was used in the statistical analyses, after which we used DIF for multidimensional IRT models (Chalmers, Counsel & Flora, 2016) and a sequential-elimination algorithm. In order to perform the Chi-square test, we coded the contents using two independent encoders and recorded the parameters and other information of the items and whether DIF was experienced in a separate database.

In the case of the M2 and M4 models, items loaded with DIF appeared almost equally, however, only with the help of the M4 model can we reveal the reasons for the differences. Based on the information curve of the M4 model test, it can be concluded that it provides much more information about the reading comprehension dimension than about mathematics, thus showing that for non-CDC and CDC students, the item measures completely different abilities.

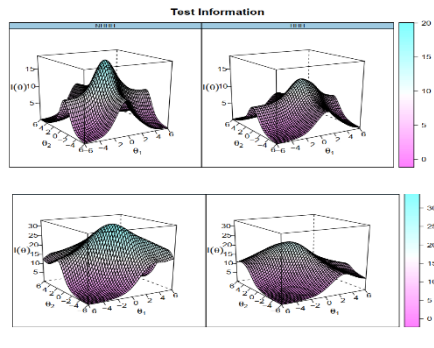


Figure 5: Test information curve for M2 and M4 models.

When examining item characteristics and content coding, only the Cognitive Operations Area of Fact Knowledge and Simple Operations shows a significantly higher chance that a given item is loaded with DIF. This may be explained by the fact that in many cases, cumulatively disadvantaged students grow up in a truncated family, do not have an intellectual role model, and are thus unable to acquire these skills (Mendi, 1999; Horváth, 2015). School segregation may also be behind the lag (Horváth, 2015).

In summary, we can say that students with multiple disadvantages rely mostly on their comprehension knowledge when solving mathematical problems. It is worth noting that among the tasks, Fact Knowledge and Simple Operations has the highest chance of appearing in DIF for cumulatively disadvantaged students, so teachers may place more emphasis on practicing and understanding these tasks for cumulatively disadvantaged students.

BIBLIOGRAPHY

1997. évi XXXI. tv. a gyermekek védelméről és a gyámsági igazgatásról. Letöltve: 2021. 04. 27. <https://netjog-tar.hu/jogszabaly?docid=99700031v>
- Balázs I., Balkányi P., Ostorics L., Palincsár I., Rábainé Szabó A., Szepesi I., Szipőcsné Krolopp J. & Vadász Cs. (2014). Az Országos Kompetenciamérés Tartalmi Keretei: Szövegértés, matematika, háttérkérdőívek. Budapest: Oktatási Hivatal
- Beese, J. & Liang, X. (2010). Do resources matter? PISA science achievement comparisons between students in the United States, Canada and Finland. *Improving Schools*, 13(3), 266-279.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 57, 289-300.
- Buda, M. (2017). Hátrányos helyzet és iskola. Merre keressük a kiutat a motiválatlanság csapdájából?. In Polonyi, T., & Aban, K. (eds). *Digitális tanulás és tanítás*. (pp. 9 - 26). Debreceni Egyetem.
- Belinszki, B., Szepesi, I., Takácsné Kárász, J., & Vadász, Cs. (2019). Országos kompetenciamérés: Országos jelentés. Budapest: Oktatási Hivatal.
- Cs. Czachesz, E. & Radó, P. (2003). Oktatási egyenlőtlenségek és speciális igények. In Falász, G. & Lannert, J. (szerk). *Jelentés a magyar közoktatásról 2003*. (pp. 351 - 376). Budapest: Országos Közoktatási Intézet.
- Chalmers, P., Pritikin, J., Robitzsch, A., Zoltak, M., Kim, K., Falk, C. F. & Oguzhan, O. (2016). Package: mirt: Multidimensional Item Response Theory - Version 1.29.
- Chalmers, R. P., Counsel, A., & Flora, D. B. (2016). It might not make a big DIF: Improved differential test functioning statistics that account for sampling variability. *Educational and Psychological Measurement*, 76(1), 114-140.
- Csapó B., Fejes J. B., Kinyó L., & Tóth E. (2014). Az iskolai teljesítmények alakulása Magyarországon nemzetközi összehasonlításban. In Kolosi Tamás & Tóth István György (szerk). *Társadalmi riport 2014 - Tanulmánykötet* (pp. 110 - 136). Budapest: TÁRKI.
- Da Costa, P. D., & Araújo, L. (2012). Differential item functioning (DIF): What functions differently for Immigrant students in PISA 2009 reading items. *JRS Scientijic and Policy Reports*. Luxembourg: European Commission.
- Deutsch, M. (1964). Early social environment: Its influence on school adaptation. In Schreiber, D. (ed). *The school dropout* (pp. 89 - 100). Washington DC: National Educational Association
- Ehmke, T., van den Ham, A. K., Silzer, C., Heine, J., & Prenzel, M. (2020). Measuring mathematics competence in international and national large scale assessments: Linking PISA and the national educational panel study in Germany. *Studies in Educational Evaluation*, 65, 100847.
- Faluvégi, A. (2004). A társadalmi-gazdasági jellemzők területi alakulása és várható hatásai az átmenet időszakában. Műhelytanulmányok Discussion Papers, 2004/5. Budapest: MTA Közgazdaságtudományi Kutatóközpont.
- Fejes, J. B. (2006). Miért (nem) fontosak a hátrányos helyzetű tanulók?. *Új Pedagógiai Szemle*, 56(7-8), 17-26.
- Fejes, J. B. (2013). Miért van szükség deszereggregációna?. In Fejes József, Balázs & Szűcs Norbert (szerk). *A szegedi és békéscsabai deszereggregáció támogatási Hálójának Mentorprogram. Öt év tapasztalatai*. (pp. 15 - 35). Szeged: Belvedere Mentoriális.
- Gergely, B., Takácsné Kárász, J., Takács, Sz. (2020). Többdimenziós IRT modellek alkalmazása az Országos Kompetenciamérésben (IDK, kézirat)
- Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019). Az iskolai teljesítménykülönbség és a szocioökonómiai státusz összefüggései – az országos kompetenciamérés eredményeinek vizsgálata a szülők munkájának rendszerességére, az észlelt társadalmi helyzet és a lakókörnyezet vonatkozásában, *Psychologia Hungarica Caroliensis*, 16, 148-221.
- Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019). Születni tudni kell: Az Országos kompetenciamérés eredményeinek vizsgálata a szülők munkájának rendszerességére, észlelt társadalmi helyzet és a lakókörnyezet vonatkozásában, *Psychologia Hungarica Caroliensis*, VII/2, 64-85.
- Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Smohai, M., Simon, G., Takács, N., Takács, Sz. (2019). The relationship of school achievement with parents' employment status, perceived social status, and living environment as reflected in findings of the 2017 National Assessment of Basic Competencies (NABC), *Psychologia Hungarica Caroliensis*, VII/2, 167-189.
- Holland, P. W., & Wainer, H. (2012). *Differential item functioning*. Routledge.
- Horváth, A. (2015). Hátrányos helyzetű diákok tanulási motiváció. *Studia Mundi - Economica*, 2(3), 97-112.
- Huang, X., Wilson, M., & Wang, L. (2014). Exploring plausible causes of differential item functioning in the PISA science assessment: language, curriculum or culture. *Educational Psychology*, 36(2), 378-390.
- Kasik, L., Gut, K., & Gáspár, Cs. (2014). Hátrányos és nem hátrányos helyzetű diákok szociálisprobléma-megoldó gondolkodása. *Magyar Pedagógia*, 114(1), 49-63.

- Keller, J., & Mártonfi, G. (2006). Oktatási egyenlőtlenségek és speciális igények. In Halász G. & Lannert J. (szerk). *Jelentés a magyar közoktatásról 2006.* (pp. 377 - 411). Budapest: Országos Közoktatási Intézet.
- Kertesi, G., & Kézdi, G. (2012). A roma és nem roma tanulók teszteredményei közti különbségekről és e különbségek okairól. Budapesti Munkagazdaságtani Füzetek, 5, 1-84. Budapest: Magyar Tudományos Akadémia Közgazdaság- és Regionális Tudományi Kutatóközpontjának Közgazdaságtudományi Intézetében működő Munkapiaci Kutatások, Budapesti Corvinus Egyetem Emberi Erőforrások Tanszéke.
- Kispál, S., Gergely, B., Talkácsné Kárácz, J. & Takács, Sz. (2020). Eltérő itemműködés vizsgálata az Országos kompetenciamérésben halmozottan hátrányos helyzetű diákok körében többdimenziós IRT modellek segítségével (IDK, kézirat).
- Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms bei der DFG. *Zeitschrift für Pädagogik*, 52, 876 - 890.
- Koltói, L., Harsányi, Sz. G., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019a). A szülői tanulmányokba való bevonódásának összefüggése az iskolai teljesítménnyel, *Psychologia Hungarica Caroliensis*, VII/2, 86-103.
- Koltói, L., Harsányi, Sz. G., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019b). Az iskolai szülői bevonódás iskolai szintű vizsgálata megyei és regionális szinten az országos kompetenciamérés 2017-es és 2018-as adatai alapján, *Psychologia Hungarica Caroliensis*, 16, 222-258.
- Koltói, L., Harsányi, Sz. G., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Nyitrai, E., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019c). The relationship between school achievement and paternal involvement in children's school activities as judged by headmasters in the 2017 National Assessment of Basic Competencies (NABC), *Psychologia Hungarica Caroliensis*, VII/2, 190-212.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale NJ: Lawrence Erlbaum Associates.
- McCandless, B. (1952). Environment and intelligence. *American Journal of Mental Deficiency*, 56, 674 - 691.
- Mendi, R. (1999). Felsőoktatásban tanuló roma fiatalok pályaszocializációs és személyiségvizsgálata. *Bölcsészeti szakdolgozat, ELTE Pszichológiai Intézet.* Letöltve: 2021. 04. 27. <http://mek.oszk.hu/02000/02034/html/index.htm>
- Nyitrai, E., Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Mátay, G., Nagybányai-Nagy, O., Pusker, M., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019a). Szülői bevonódás és az iskolai teljesítmény kapcsolata az országos kompetenciamérés 2017-es és 2018-as adatainak tükrében, *Psychologia Hungarica Caroliensis*, 16, 7-51.
- Nyitrai, E., Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019b). Iskolai teljesítmény és szülői bevonódottság, *Psychologia Hungarica Caroliensis*, VII/2, 7-28.
- Nyitrai, E., Harsányi, Sz. G., Koltói, L., Kovács, D., Kövesdi, A., Nagybányai-Nagy, O., Simon, G., Smohai, M., Takács, N., Takács, Sz. (2019c). Relations between Parental Involvement and School Performance in the Light of Data from National Assessment of Basic Competencies (NABC) 2017, *Psychologia Hungarica Caroliensis*, VII/2, 115-130.
- Országos Kompetenciamérés Technikai Leírás (2012). Letöltve: 2020. 08. 17. https://www.oktatas.hu/pub_bin/dload/kozoktatasi/meresek/unios_tanulmanyok/OKM_Technikaileiras.pdf
- R Core Team (2013). R: A language and environment for statistical computing.
- Radó, P. (1997). Jelentés a magyarországi cigány tanulók oktatásáról, szakértői tanulmány a Nemzeti Etnikai Kisebbségi Hivatal számára. Budapest.
- Rajnai, J. (2012). Lehet másképp?! – Avagy pedagógiai alternatívák a hátrányos helyzetű roma fiatalok nevelésében. *Új Pedagógiai Szemle*, 62(11-12), 54–75.
- Siller, J. (1957). Socioeconomic status and conceptual thinking. *Journal of Abnormal and Social Psychology*, 55, 365 - 371.
- Shanmugam, S. K. S. (2020). Gender-Related Differential Item Functioning of Mathematics Computation Items among Non-native Speakers of English. *The Mathematics Enthusiast*, 17(1), 108-139.
- Szabó, L. D., Szepesi, I., Talkácsné, K. J., & Vadász, Cs. (2018). Országos Kompetenciamérés 2017 Országos jelentés. Budapest: Oktatási Hivatal
- T. Kárácz Judit (2019a): Hibabeccslés eljárások véletlen jelenségek paramétereinek becslésére, *Psychologia Hungarica Caroliensis*, VII/2, 104-114.
- T. Kárácz Judit (2019b): Estimation methods on standard error of different statistical parameters, *Psychologia Hungarica Caroliensis*, VII/2, 213-220.
- Trier, U. P., & Peschar, J. (1995). Cross-Curricular Competencies: Rationale and Strategy for Developing a New Indicator. In OECD (ed), *Measuring what Students Learn* (pp. 99 - 109). Paris: OECD
- Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods*, 17(2), 228.
- Walker, C. M., & Beretvas, S. N. (2003). Comparing multidimensional and unidimensional proficiency classifications: Multidimensional IRT as a diagnostic aid. *Journal of Educational Measurement*, 40, 255 - 275.

HÁTRÁNYOS HELYZETŰ DIÁKOK TÖBBDIMENZIÓS IRT MODELL VIZSGÁLATA

- Weinert, F. E. (2001). Concept of competence: a conceptual clarification. In Rychen, D. S., & Salganik, L. H. (eds), *Defining and selecting key competencies* (pp. 5 - 65). Seattle: Hogrefe & Huber Publishers.
- Yıldırım, O. (2019). Detecting Gender Differences in PISA 2012 Mathematics Test with Differential Item Functioning. *International Education Studies*, 12(8), 59 - 71.