



MAGYAR TUDOMÁNYOS AKADÉMIA
KÖZGAZDASÁGTUDOMÁNYI INTÉZET

BUDAPESTI MUNKAGAZDASÁGTANI FÜZETEK

BWP – 2010/1

**NEM-KOGNITÍV KÉSZSÉGEK MÉRÉSE
AZ OKTATÁSI INTEGRÁCIÓS
PROGRAM HATÁSVIZSGÁLATÁBAN**

SURÁNYI ÉVA – KÉZDI GÁBOR

Budapest Working Papers On The Labour Market
Budapest Munkagazdaságtani Füzetek
BWP – 2010/1

Nem-kognitív készségek mérése az oktatási integrációs program hatásvizsgálatában

Magyar Tudományos Akadémia Közgazdaságtudományi Intézet
Budapesti Corvinus Egyetem, Emberi Erőforrások Tanszék

Szerzők:

Surányi Éva
kutatási igazgató
MATT Humán Tanácsadó Kft.
E-mail: suranyi@matt-od.com

Kézdi Gábor
tudományos főmunkatárs
Magyar Tudományos Akadémia Közgazdaságtudományi Intézet
associate professor
Közép-európai Egyetem
E-mail: kezdig@ceu.hu

2010. április

ISBN 978 963 9796 97 3
ISSN 1785 3788

Kiadja

a Magyar Tudományos Akadémia Közgazdaságtudományi Intézete

Nem-kognitív készségek mérése az oktatási integrációs program hatásvizsgálatában

Surányi Éva – Kézdi Gábor

Összefoglaló

A nem-kognitív készségek mérése növekvő szükséglet az oktatáspolitikai intervenciók hatásvizsgálataiban. A mérőeszközök kiválasztása azonban nehéz feladat. A kognitív fejlődéssel ellentétben kevés olyan megfelelően validált és standardizált teszt áll rendelkezésre, amely nagymintás felmérésekben, longitudinális vizsgálatokban, és más olyan kutatási helyzetekben alkalmazható, ahol az idő jelentősen korlátozza az adminisztrálható tételek számát. Tanulmányunk az Országos Oktatási Integrációs Hálózat 2003 őszen indult iskolai integrációs programjának hatásvizsgálata során használt mérőeszközök kidolgozásának elméleti hátterét és fejlesztési részleteit ismerteti. Ezen keresztül általános érvényű szempontokat vet fel a mérőeszközök kidolgozásának általános nehézségeivel és kihívásaival kapcsolatban. A program hatását vizsgáló kutatásunk módszereit és eredményeit a Magyar Tudomány 2010. márciusi számában súlyos bírálatok érték. Ez a tanulmány a nem-kognitív készségeket mérő eszközök kidolgozására irányuló bírálatokra adott válaszainkat tartalmazza. Külön cikkben válaszolunk a hatásvizsgálat mintavételét és elemzési módszereit ért kritikákra. Tanulmányainkban részletesen kimutatjuk, hogy a bírálatok megalapozatlanok.

Tárgyszavak: integrált oktatás, program hatásvizsgálat, nem-kognitív készségek, mérőeszközök kidolgozása

JEL: I20, I38, C81

Válaszcikk Kabai Péternek a Magyar Tudomány című folyóirat 2010. márciusi számában megjelent „Egy sikeres iskolai integrációs program elemzése“ /Analysis of a Study on a Successful School Integration Program/ című – a szerzők általános iskolai integráció hatását elemző kutatását bíráló – cikkére.

Measuring Non-Cognitive Skills in the Evaluation of the National Education Integration Program of Hungary

Éva Surányi – Gábor Kézdi

Summary

Measuring non-cognitive skills in evaluations of education policy interventions is an increasingly important task. At the same time, choosing the adequate measures is difficult. Contrary to cognitive development, there is a shortage of reliable and valid measures of personality and behavioral traits that can be used in large scale longitudinal evaluations or in any other research contexts where time constraints limit the number of adoptable questionnaire items. Our study presents the theoretical background and the details of the design of the measurement instrument adopted in the evaluation of the education integration program launched by the National Integration Network (OOIH) of Hungary in 2003. It also demonstrates general points of the design of such measurement instruments. The results and methods of the impact evaluation research were criticized in a paper published (in Hungarian) in *Magyar Tudomány*, March 2010. This study answers the criticism on the measurement of non-cognitive skills. A separate study answers the criticism of sampling and methodology. In both papers we show that the critiques are unfounded.

Keywords: integrated education, program evaluation, non-cognitive skills, measurement

JEL: I20, I38, C81

BEVEZETŐ

Az Országos Oktatási Integrációs Hálózat 2003 őszén egy olyan, általános iskolás gyerekekre irányuló programot indított útjára, melynek célja a hátrányos helyzetű gyerekek - otthoni gondozói környezetből eredő - lemaradásainak kompenzálása és továbbtanulási esélyeinek növelése volt egy integrált¹, magas minőségű oktatási/nevelési környezet biztosításával. A program 45 iskola részvételével indult el, lényegi eleme az integrált környezetben folyó hatékony és előítéletektől mentes oktatás elősegítése és hazai meghonosítása volt. Ennek érdekében a pedagógusok olyan képzéseken vettek részt, melyeken módjuk volt elsajátítani a modern, kompetenciaorientált, gyermekközpontú és kooperációt segítő oktatási módszerek használatát. Az iskolák emellett egy koordinátori hálózat működtetésén keresztül segítséget kaptak az iskolai élet hatékonyabb szervezéséhez, valamint az iskola és környezete kapcsolatának fejlesztéséhez. Kutatásunk e program hatását vizsgálta a tanulók fejlődésére és iskolai eredményességére.

A gyerekek fejlődésének nyomon követésében nagy jelentőséget tulajdonítottunk annak a modern szemléletnek, mely – szakítva azzal a kutatói gyakorlattal, ami az oktatási programok hatását kizárólag a gyerekek kognitív (tudásbeli) fejlődése mintázatainak elemzésével próbálja feltárni - felismeri és elfogadja a gyermeki fejlődés komplexitását és a szociális kompetenciák mérésének jelentőségét. Tanulmányunk így túlmutat a szóban forgó kutatásban használt mérőeszközök elméleti háttérének és fejlesztési részleteinek ismertetésén, azokat tágabb kontextusban értelmezve szempontokat vet fel a mérőeszközök kidolgozásának általános nehézségeivel és kihívásaival kapcsolatban is.

A NEM-KOGNITÍV KÉSZSÉGEK MÉRÉSÉNEK JELENTŐSÉGE

Bár a gyerekek intellektuális fejlődése fontos tényező iskolai előmenetelük alakulásában, ez életpályájuk és társadalmi beilleszkedési sikereik korántsem egyedüli meghatározója.

¹A programban résztvevő iskoláknak az osztályok újraszervezésével meg kellett szüntetniük a párhuzamos osztályok tanulói összetételében megmutatkozó különbségeket, ha voltak ilyenek. Az intervenció az egyes iskolákra fókuszált, természeténél fogva nem célozta az iskolák közötti szegregáció problémáját.

A gyerekek személyes és társas kompetenciája, egyéni fejlődéstörténete, tanulással kapcsolatos érzelmei és az iskolai helyzetekkel kapcsolatos önértékelése szintén nagy jelentőséggel bír iskolai pályájuk alakulásában. Maga a kognitív fejlődés is a más területeken történő fejlődéssel szimultán történik, a tanulást közvetlenül segítő készségek alakulásával egy időben fejlődik a gyermek én-ismerete, társas interakciós készsége, felelősségtudata, és számos más olyan képessége, mely a világban való eligazodását segíti. Mára az is bizonyított lett, hogy ezeken a közgazdasági irodalom által nem-kognitív készségek² gyűjtőfogalom alá sorolt területeken történő fejlődés menete nagyobb hatással van a gyermekek későbbi munkaerőpiaci sikereinek alakulására, mint maguknak a kognitív kompetenciáknak a formálódása (Heckman, Stixrud, Urzua, 2006).

A nem-kognitív készségek mérése így növekvő szükséglet az oktatási/nevelési programok hatásvizsgálataiban, vizsgálatuk mégis sokáig figyelmen kívül maradt az ilyen irányú kutatásokban, nemcsak a hazai, de a nemzetközi szintén is (Heckman és Rubinstein, 2001). Ennek legfőbb oka a mérési eszközök kiválasztásának nehézségében rejlik. Egyrészt a nem-kognitív képességek összessége nem egy jól definiált, mérhető ismérvekből álló gyűjtőfogalom, inkább olyan mögöttes egyéni kompetenciákat jelent, amely azután számos más területen érezteti hatását (Collins és Gunnar, 1990), s amelyek ebből kifolyólag nem összegezhetőek egyetlen (vagy néhány) olyan faktorról, amely e sokszínű képességeket magába foglalná (Cawley és mtsai, 2001). A közgazdasági irodalom ugyanakkor magára hagyta az empirikus kutatókat annak eldöntésében, melyek is valójában a vizsgálandó nem-kognitív készségek.

Másrészt a kognitív fejlődéssel ellentétben az itt releváns dimenziók mérésére sokkal kevesebb megfelelően validált és standardizált teszt áll a kutatók rendelkezésére. Néhány létező, a gyermekek szociális és emocionális fejlődésének mérésére jól kidolgozott módszer nem alkalmas sok gyermeket, iskolát érintő beavatkozások hatásvizsgálatára, mert a teszt felvételének elsajátítása hosszas képzést, maga a felvétel pedig jól körülhatárolt laboratóriumi körülményeket igényel. Más, kifejezetten osztálytermi környezetre kidolgozott tesztek ugyanakkor gyakran azért nem kivitelezhetők, mert túlságosan szubjektívek vagy „tolakodók”, ami a válaszadások nagymértékű csökkenését vonja maga után. Aktuális probléma, és éppen ezért nem lehet eléggé hangsúlyozni, hogy a nagymintás felmérésekben használható (nem pedig diagnosztikai céllal készült) személyiség- és viselkedéses vonások megfelelően validált és reliábilis mutatóinak hiánya még mindig az ilyen irányú kutatások legfőbb akadálya (Surányi és Danis, 2009).

²Erről a közgazdasági irodalom azt feltételezi, hogy merőleges a kognitív teljesítményre, és tartalmaz minden olyan tényezőt, ami különbözik az egyén kognitív képességeitől, de hatással van későbbi munkaerőpiaci sikereire. Ennek ellenére a legtöbb pszichológus valószínűleg vonakodna attól, hogy ezeket a tényezőket minden kétséget kizáróan nem-kognitívnek nevezze (lásd Heckman (2000) és Bowles és Gintis (2001) átfogó irodalmi összefoglalóját).

A VIZSGÁLT FEJLŐDÉSI TERÜLETEK KIVÁLASZTÁSA A HATÁSVIZSGÁLATBAN

Az oktatási programok hatásvizsgálatainak egyik legnagyobb nehézségét maguknak a vizsgálni kívánt fejlődési területek körének meghatározása jelenti, két ok miatt is. Egyrészt, mivel a hatásvizsgálatok eredményei jelentik a legfontosabb kiindulópontot a kormányzati intervenciók fenntartható működtetésével kapcsolatos döntések meghozatalában, nagy jelentőséget kap az, hogy a kutatás olyan fejlődési területekre irányuljon, amikben a pozitív elmozdulások nemcsak az egyének (egy-egy gyerek), hanem várhatóan a *társadalom szélesebb rétegei számára is hasznot jelentenek*, főként a szociális kiadások csökkenésén keresztül (pl. a felnőttkori munkavállalás valószínűségének növekedésében vagy a különböző deviáns viselkedésmódok rizikójának csökkenésében).

Másrészt, bár az oktatási/nevelési programok hatása a legtöbb esetben csak hosszú távon (sokszor többéves késéssel) bontakozik ki, a hatásvizsgálatoknak általában ezek élettartamához és hatásainak érvényesüléséhez képest csak nagyon rövid táv áll rendelkezésére (egy-két év) arra, hogy a beavatkozás menetéről adekvát információt szolgáltatassanak. Ezért a hatásvizsgálatok általában arra törekednek, hogy olyan közvetlen, rövidtávon is érvényesülő hatásokat mérjenek, melyek feltehetően az intervenció hosszú távú hatásainak *mediátor változói* (Surányi és Danis, 2009). Ennek technikai szempontból igen fontos következménye viszont az, hogy pusztán abból a tényből, hogy a vizsgált tényezők *milyen aránya* szignifikáns egy hatásvizsgálatban, nem lehet következtetéseket levonni magának a programnak a hatékonyságára vonatkozóan. Az ebben az életkori szakaszban történő beavatkozások eredményességének közvetítő mechanizmusai ma még nem ismertek pontosan, maga a pszichológiai irodalom sem egységes például abban, hogy az iskolai integrációval járó megnövekedett csoportközi kontaktus várhatóan az előítéletesség kognitív, affektív, vagy viselkedési komponensére hat-e inkább (Pettigrew, 1998). Ugyanez a gondolatmenet magyarázza azt is, hogy *miért nem lehet a hatásnagyságok abszolút értékéből sem következtetni a gyermekkori programok sikerességére*. A hátrányos helyzetű gyermekek fejlődését támogató, a gondozói környezet minőségének javítását célzó beruházások megtérülését vizsgáló közgazdasági kutatások arra mutattak rá, hogy a különböző életkori szakaszokban történő tudásbeli „befektetések” egymással komplementer viszonyban állnak (vagyis a támogató korai évek teszik lehetővé az iskolai befektetések későbbi jobb érvényesülését, ami azután a munkaerő növekvő minőségében térül meg), így a gyermekkorban történő *látszólag kis fejlődési elmozdulás is jelentős hosszú távon érvényesülő társadalmi haszonnal járhat*. A költségvetési erőforrások gyermekkora történő összpontosítása így amellet, hogy a társadalmi egyenlőtlenségek csökkentésének hatékony eszköze, megtérülő befektetés a társadalom gazdasági hatékonysága tekintetében is

(Heckman és Masterov, 2007; lásd még Surányi és Danis, 2010 részletes összefoglalása a témáról). Félrevezetőnek gondoljuk ezért, amikor Kabai Péter az olvasókat több helyen is arról próbálja meggyőzni, hogy a kutatásban tapasztalt pozitív elmozdulások az integrációs programban részt vevő és a kontroll csoportot alkotó gyerekek között szignifikánsak ugyan, de „bizonytalan” mértékűek³.

Mindezen megfontolások szem előtt tartása mellett hatásvizsgálatunk a gyermeki nem-kognitív fejlődés három területének vizsgálatára tért ki. Az első a gyerekek *átfogó önértékelésének*⁴ két összetevője, az énkép és a fogalmilag és tartalmilag is ehhez társítható kontrollhely (lásd később). Valószínűsíthető ugyanis az, hogy a pozitív gondozói környezet hosszú távon is megmutatkozó hatása - részben legalábbis - a gyermekek önértékelésében bekövetkező pozitív változásokra vezethető vissza, hiszen az énfogalom kialakulása központi jelentőségű, funkcionális szerepet tölt be a gyermekek társas és érzelmi fejlődésében (Kagen, Moore, Bredekamp, 1995). Két nagymintás, szigorú módszertani követelményeknek megfelelő longitudinális kutatás is⁵ egyöntetűen bizonyítja emellett, hogy az énkép és a kontrollhely hosszú távon is hatással vannak az iskolai végzettség alakulására, a várható keresetekre, a hosszú ideig tartó munkanélküliségben töltött életszakaszokra, de magára a szakmaválasztásra és néhány deviáns magatartásforma megjelenési kockázatára is (Feinstein, 2000; Heckman, Stixrud, Urzua, 2006; Blanden, Gregg, Macmillan, 2006).

Az iskolai integráció hatásának nyomon követésére kiválasztott második vizsgálati terület kutatásunkban a gyerekek *csoportközi toleranciájának illetve előítéletes gondolkodásának* vizsgálata volt. Ennek indokoltsága Gordon Allport mára klasszikussá vált munkájára vezethető vissza, aki már 1954-ben felvetette, hogy az ellenséges viszonyban levő csoportok közötti fizikai távolság csökkentése – bizonyos feltételek mellett - a csoportok közötti előítéleteket csökkentheti. A deszegregációs programok így a különböző társadalmi csoportok kapcsolatba kerülése révén, a megismerés, megértés növelésével, barátságok kialakulásával, képesek lehetnek hosszú távon csökkenteni az etnikai konfliktusokat és ellenségeskedést (Taylor, Rickett, 1981). Allport korai munkája nyomán széles körű kutatás-sorozat bontakozott ki, melyek azóta részletes képet rajzolnak a csoportközi kapcsolatok és előítéletesség összefüggésének előfeltételeiről és közvetítő mechanizmusairól (lásd Pettigrew és Tropp (2000) összefoglalását).

E két fejlődési terület mérése kutatásunkban felvetett egy további problémát. A pozitív énkép és előítélet-mentesség olyan „szociálisan kívánatos” tulajdonságok, melyek a külső elvárásoknak nagymértékben megfelelni akaró személyek esetében jelentősen csökkenthetik

³ A statisztikai eljárásokkal kapcsolatos kritikákat a mintavételről és becslési eljárásokról szóló cikkünk válaszolja meg.

⁴ core self evaluation (Judge, Locke and Durham, 1997)

⁵ Az National Longitudinal Survey of Youth (NLSY) 1979, amely 1979 óta gyűjt adatokat az akkor 14-21 éves korú fiatalok iskolai teljesítményére, későbbi munkaerőpiaci aktivitására és keresetere nézve vonatkozóan; illetve a *British Cohort Study*, mely egy 1970-ben (akkor újszülött gyerekekkel) elkezdődött, napjainkban is folyó longitudinális kutatás Angliában.

a válaszok megbízhatóságát (Reynolds, 1979). Fokozottan jelentkezik ez a probléma a gyerekek esetében, akiknél a társas elvárásoknak megfelelő válaszadásra törekvésük normális fejlődésük része. Kérdőívünk ezért tartalmaz a gyerekek *társas kívánatosságát* mérő kérdéseket, melyek lehetővé teszik az ebből eredő torzítások kiszűrését.

A MÉRŐESZKÖZÖK KIALAKÍTÁSÁNAK ÁLTALÁNOS ALAPELVEI

A hatásvizsgálatoknak nemcsak a fejlődési területek kiválasztásában, hanem az ezek mérésére alkalmas eszközök kidolgozásában is komoly kihívásokkal kell szembenézniük. Az adatfelvétel nagy költségigénye miatt ugyanis a hatásvizsgálatoknak a gyerekek komplex fejlődési folyamatának értékelésére *egyetlen időpontban felvett adatok alapján* kell következtetniük, ellentétben például a pedagógusok és gondozók munkájának segítése céljából kidolgozott tesztelési eljárásokkal, ahol a fejlődési jellegzetességek megfigyelésére hosszú időszak áll rendelkezésre, s a döntések következménye is kevésbé súlyos: ha a pedagógusok tévesen ítélik meg egy alkalommal a gyermek képességeinek szintjét, később azonnal korrigálni tudnak a gyerekekkel való napi kapcsolat során. Ráadásul a program településenkénti/intézményenkénti megvalósulásának várhatóan nagy heterogenitása miatt a hatásvizsgálat szükségszerűen *nagy elemszámú mintával* dolgozhat csak annak érdekében, hogy az intézményenként és gyerekenként is eltérő pedagógiai munka hatását megbízhatóan nyomon tudja követni. Egy nagy elemszámú, közel 4000 gyermekre kiterjedő hatásvizsgálatban ugyanakkor, nincs idő-, energiaforrás, valamint „intellektuális” kapacitás arra, hogy hosszú, időigényes, részletes és nehezen kitölthető kérdőíveket, illetve fejlődési teszteket vegyünk fel, ami tovább korlátozza a tesztek megalkotóinak mozgásterét.

Szerencsére a gyermeki fejlettség pontos mérésének nehézségeit részben kompenzálja az adatok aggregát természetete. Fontos annak megértése ugyanis, hogy a hatásvizsgálatok módszertana azt használja ki, hogy az adatgyűjtés ekkor *nem az egyénekre* (az egyes gyerekekre) irányul, *hanem gyerekek aggregált csoportjaira*, mivel a hatásvizsgálat célja a programmal kapcsolatos politikai döntéshozatal támogatása, oktatási és szociális programok tervezésének empirikus alátámasztása, nem pedig magukról a gyermekekről alkotott véleményformálás. Ezt kihasználva, a hatásvizsgálatok ahelyett, hogy a validitás és a reliabilitás biztosítása érdekében mély, diagnosztikus jellegű felmérést végeznének a programban részt vevő gyermekekkel, *megfelelő számosságú gyerekekről történő adatgyűjtéssel biztosítják a mérés pontosságát* és magas megbízhatóságát, az egyes gyerekek túlzott terhelése nélkül (Surányi és Danis, 2009).

Mindebből következően a tesztelési eljárások kidolgozása csak részben építhet azokra a már publikus magyar és nemzetközi mérőeszközökre, amik eltérő céllal (pl. fejlődési lemaradások diagnosztizálása, iskolaérettség) és technikai igényel lettek kidolgozva és

standardizálva. A hatásvizsgálatokban a mérőeszközök kidolgozása általában ezek megfelelő módosításával (redukciójával) történik, melyek így hasznos helyettesítői lehetnek az eredeti kérdőíveknek különböző kutatási kontextusokban, elsősorban nagymintás felmérésekben, longitudinális vizsgálatokban, és más olyan kutatási helyzetekben, ahol az idő jelentősen korlátozza az adminisztrálható tételek számát. E rövidített változatok előnye ezen felül az, hogy csökkentik a tételek redundanciáját, illetve a kifáradás, a frusztráció és az unalom megjelenésének valószínűségét, ami gyakran előfordul, amikor hasonló kérdésekre kell válaszolni nagy számban, különösen, ha figyelembe vesszük a gyerekek figyelmének területi korlátait. Az eredmények hibásan értelmezett, félrevezető interpretálásából eredő téves döntések elkerülése érdekében viszont hangsúlyoznunk kell még egyszer, hogy az így megkonstruált mérőeszközök nem alkalmasak arra, hogy következtetéseket vonjunk le belőle a programban részt vevő *egyedi gyermekekre* illetve gondozókra vonatkozóan.

Szeretnénk itt azonban elébe menni egy lehetséges kifogásnak, ami látszólag jelentős pszichometriai indokot szolgáltat ahhoz, hogy szkeptikusak legyünk a rövidített változatok használatának helyességével, még azok gyakorlati érényei ellenére is. A Spearman-Brown formula és a klasszikus reliabilitás teória szerint ugyanis a kérdésekre adott válaszok két részből tevődnek össze, egy random mérési hibából és a valós válaszok pontszámainak szóródásából. A több-iteemes skálák előnye ebből a nézőpontból így az lehet, hogy a tételek összegzésével a random hibák kioltják egymást, így a több-iteemes skálák megbízhatósága nagyobb. Ezzel kapcsolatban azonban két dolgot szükséges látni. Az egyik, hogy jó okunk van feltételezni, hogy a fogalmilag egydimenziós, nem terület-specifikus vonások esetében a tételek redundanciája következtében a mérési hiba aligha csökkenthető a tételek számának emelésével, ráadásul a mérési hibák nem feltétlenül random módon jelennek meg. Sőt, tíz alkalommal feltenni többé-kevésbé ugyanazt a kérdést inkább növelheti a szisztematikus hiba megjelenésének valószínűségét (pl. a szociálisan kívánatos válaszok megjelenését), ami a Spearman-Brown formula érvényességét megkérdőjelezi ezekben az esetekben (lásd pl. Robins, Hendin, Trzesniewski, 2001 érvelése a Rosenberg skála kapcsán). Másrészt, bár nem szeretnénk azt mondani, hogy azokban a diagnosztikus eljárásokban, amelyek az *egyed gyermekek* fejlődési deficitjeinek felismerésére irányulnak, nincs jelentősége akár egy-egy vizsgált dimenzióknak is a hosszas, sok kérdésre kiterjedő vizsgálatára, azokban a nagymintás kutatásokban, melyek elsődleges célja az eltérő gondozói környezetben nevelkedő gyerekek fejlődése *közötti* heterogenitás megértése, nem pedig az egyes gyermekek fejlődésére vonatkozó következtetések levonása, a *fogalmilag egydimenziós* konstrukciók jól mérhetőek egy rövidített skálával úgy, hogy közben azok tartalmi validitása nem sérül, és a kérdőív itemeinek számának csökkentéséből eredő információ-veszteség elhanyagolható mértékű.

Meglepő ugyanakkor, hogy amikor Kabai Péter kétségeket próbál ébreszteni a rövidített kérdőívek megfelelőségével kapcsolatban, a kérdőív tételek felsorolását kéri számon a

zárótanulmányban, hiszen amint írja: „megítélni a módosítás utáni tesztek validitását lehetetlen, mert a szerzők nem közlik a tesztkérdések megfogalmazását”. A valóságban, aki ismeri a diagnosztikus (egyedi gyerekek mérésére kidolgozott) tesztek rövidítésének szigorú módszertani eljárását, ami valóban garantálja azt, hogy az eredeti mérőeszközök tartalmi validitása ne sérüljön az itemek számának csökkentésével, az tudhatja, hogy ezek az eljárások hosszadalmasak, dokumentációjuk önmagukban is terjedelmes és technikai jellegű, mint amilyen a mi kutatásunkban is volt. A kérdőívek validitásának megfelelőségét pusztán a kérdőív tételek ismerete alapján így lehetetlen megítélni. Szakmailag nem helytálló ezért például Kabai Péter intuitív érvelése sem a harter-i tanulási énkép többdimenziósságára vonatkozóan, melyben különbséget tesz a kérdések *ránézésre* „inkább iskolai előmenetelre” illetve „inkább kognitív készségekre” rákérdező csoportjai között, vagy a Rosenberg skála „különböző töltetű kérdései” között. Az, hogy e kérdések valóban más tartalmi elemekre kérdeznek-e rá, vagy a kérdésfeltevés módja befolyásolja-e a válaszok mintázatát, egy empirikusan tesztelhető kérdés, ami meg is történt elemzéseink során (lásd a későbbiekben).

Visszatérve a kérdőívek kidolgozásának módszertanához, a fenti érvelésben természetesen nagy hangsúly van azon, hogy a rövidített formák használatának helyessége azon a feltételezésen alapul, hogy a vizsgált konstruktum egy-dimenziós vonás. Ennek ellenőrzését a nemzetközi kutatási eredmények és elméleti megfontolások részletes áttekintése mellett egy elővizsgálat során minden esetben empirikusan is teszteltük; a kérdőív-itemek tartalmi dimenzióit főkomponens-analízis segítségével ellenőriztük, s a kérdőívek rövidített verzióiban minden dimenzióból azokat a tételeket tartottuk meg, amelyek a legjobban kapcsolódtak az adott konstruktumot alkotó kérdésekből álló eredeti faktorokhoz⁶. Egyetlen vizsgálati terület - az előítéletes gondolkodás látens formáit mérő eszközök áttekintése - esetében azonban egyáltalán nem találtunk olyan kérdőívet, ami pszichometriai megbízhatóságát tekintve megfelelhett volna a hatásvizsgálat követelményeinek, így ennek tartalmi vonatkozásainak megismerése után a gyerekekkel folytatott fókuszcsoporthoz beszélgetések alapján dolgoztunk ki mi magunk egy mérőeszközt.

Az elővizsgálatot⁷ egy olyan településen végeztük el, ahol az önkormányzat egy évvel kutatásunk előtt két peremkerületi városrészben szüntetett meg általános iskolát, ahonnan a gyerekeket a jogutód intézményekbe azóta iskolabusszal szállítják. Minden jel arra mutat, hogy a jogutód iskolába azok a gyermekek mentek, akiknek a (zömében hátrányos helyzetű és roma) szülei nem tudtak a szabad iskolaválasztás jogával élni. A befogadó iskolában az alsó tagozatos évfolyamokat egyenlő képességű heterogén csoportba sorolták (vagyis az osztályokat újraszervezték), míg felső tagozatban az eredeti osztály-struktúrák érintetlenek maradtak. Az elővizsgálat elvégzéséhez az iskola így alkalmas terepet szolgáltatott, mivel itt

⁶ Az egyes dimenziókon belül általában két kérdést tartottunk meg, ami az instrumentális változók módszerének használata segítségével lehetővé teszi a mérési hibák további csökkentését (Hu, Schennach, 2008).

egy intézményen belül találkozhattunk egy iskolaösszevonás különböző stratégiáival (beintegrálás, illetve az osztályok érintetlenül hagyása). Végül pedig annak érdekében, hogy az így kialakított mérőeszközökről országosan összehasonlítható adatokhoz jussunk, a kérdőíveket két országos mintán is standardizáltuk; egy országosan reprezentatív mintán, illetve egy olyan mintán, mely az ország azon iskoláira reprezentatív, ahol a roma tanulók aránya 10 százalék felett van.

AZ ÉNKÉP MÉRÉSE

A ma létező több száz énkép-teszt közül a leginkább széles körben használt a *Rosenberg Self-esteem Scale* (RSE), mely az *átfogó önértékelés egydimenziós mérésére* jött létre (Rosenberg, 1965). Az RSE egyik vonzereje egyszerűségében rejlik, az eredeti skála mindössze tíz, a személy önmagával kapcsolatos véleményét megfogalmazó kijelentést tartalmaz, és a személynek egyszerűen el kell döntenie, hogy egyetért-e a kijelentésekkel vagy sem. Az RSE-n több pszichometriai és validálási vizsgálatot végeztek, mint bármely más önértékelést mérő skálán, amik egyöntetűen támasztották alá a skála megbízhatóságát és pszichometriai érvényességét (Gray-Little és mtsai, 1997). A skála magas belső konzisztenciáját és egydimenziósságát elővizsgálatunk eredményei is megerősítették; azzal a már korábbi kutatásokban is tapasztalt jellegzetességgel (pl. Marsh, 1996), hogy a kérdésmegfogalmazás iránya (pozitív/negatív tartalom) hatással van a válaszok mintázatára. Emiatt az RSE kutatásunkban használt rövidített formájában két pozitív és két negatív tartalmú tételt is szerepeltettünk.

Rosenberg elképzelésével ellentétben viszont, aki az énképet egydimenziós vonásnak felételezte, a modern pszichológiai elméletek arra mutattak rá, hogy az általános önértékelés (pl. „értékes ember vagyok”) mellett terület-specifikus (kognitív, szociális, fizikai, stb.) önértékelések is elkülöníthetőek egymástól. Mai ismereteink szerint tehát *az énkép jól differenciált, hierarchikus jelleget ölt*, amely specifikus önértékelési komponensek egymásra épülő rendszeréből áll. Az empirikus munkák ugyanakkor arra is rámutattak, hogy az általános és terület-specifikus önértékelések között csak gyenge kapcsolat tapasztalható, ezek így bár kapcsolódó, kérdőíves kutatásokban egymással semmiképpen sem helyettesíthető vagy felcserélhető fogalmak (Blascovich és Tomaka, 1991). Hatásvizsgálatunkban az énkép mérése ezért kiterjedt egy másik, a terület-specifikus önértékeléseket is feltérképező mérőeszköz alkalmazására is. A gyerekek részére kidolgozott „összegző” skálák közül ma a Susan Harter (1985) által kifejlesztett *Self Perception Profile for Children* (SPPC) a legelterjedtebb az iskoláskorú gyermekekre irányuló programok hatásvizsgálataiban⁸, így mi

⁷ Az elővizsgálatban összesen 186 gyerek vett részt.

⁸ Néhány, a SPPC-t használó hatásvizsgálat példaként: Across Ages, Big Brothers/Big Sisters, Thunderbirds Teen Center, Woodrock Youth Development Program, de ezt használja az NLSY is.

is ezt használtuk. Az SPPC egy 36 tételes skála, amely a gyerekek terület-specifikus készségeiknek ön-megítélése mellett (iskolai kompetencia /társas elfogadottság /fizikai kompetencia /külső megjelenés /viselkedés, fegyelmezettség) az általános énképre vonatkozó kérdéseket is tartalmaz. Elővizsgálatunkban az SPPC faktorstruktúrája többségében jól reprodukálta Susan Harter eredményeit, néhány egyedi kivételtől eltekintve (például a gyerekek az általános énképre vonatkozó „elégedett vagyok önmagammal” kérdést a külső megjelenésre vonatkoztatták). Kabai Péter intuitív érvelését azonban a tanulási énkép több-dimenziósságáról, vagy a kérdésfeltevés módjának hatásáról sem a teszt megbízhatóságát értékelő nagyszámú nemzetközi kutatás, sem pedig saját hazai kutatási eredményeink nem erősítik meg.

Végül kutatásunk kitért egy az énképpel szoros összefüggésben álló, azzal konceptuális és elméleti hasonlóságot mutató vonás, a *külső/belső kontrollhely* mérésére is. A belső kontrollhely azt a hitet jelenti, hogy képesek vagyunk saját magunk irányítani sorsunkat, míg a külső kontrollal rendelkező személyek az események okait külső körülményekben keresik, úgy gondolják, életük önálló alakítására nincs lehetőségük. A siker vagy a kudarc okának önmaguk cselekedeteiben való keresését a gyerekek úgy képesek elsajátítani, ha felismerik a viselkedésük és annak következményei közötti ok-okozati kapcsolatot, vagyis a cselekedeteik következményeiért vállalt felelősségen keresztül (Judge és mtsai, 2003). Kutatásunkban a Rotter-skála (Rotter, 1966) rövidített, 4-iteemes verzióját használtuk. Ennek kritikai vitatását azért érezzük különösen balszerencsésnek Kabai Péter részéről, mert a kérdőív rövidítését ebben az esetben nem mi végeztük el, hanem egy a miénkkel analóg módszertannal kidolgozott, már használatban levő rövidített verziót vettünk át a külföldi kutatások gyakorlatából. Ezt a rövidített formát ma nem kisebb felmérések használják rendszeresen, mint például a több tízezer amerikai fiatalra kiterjedő NLSY⁹ longitudinális kutatássorozat.

⁹ <http://www.nlsinfo.org/nlsy79/docs/79html/79text/attitude.htm>

A RASSZIZMUS MÉRÉSE

A rasszizmus mérésében két eltérő fogalom is megjelenik függő változóként, melyek bár közös elméleti alapról erednek, egymástól jól elhatárolhatók, s egyben elkülönülten fejlődő kutatási irányok is kapcsolódnak hozzájuk: a *sztereotípiák* a külső csoport tagjaival kapcsolatos vélekedések kognitív komponensei, míg az *előítéletek* a csoportközi kapcsolatok érzelmi összetevői.

A *sztereotípiákat* (egy szociális csoport tagjairól alkotott leegyszerűsített és túláltalánosított megítéléseket) mérő eszközök alapvető elemei minden esetben az egyéni hiedelmek: a kutatók egy adott szociális csoportról alkotott vélekedéseikről kérdezik a válaszolókat. A legmodernebb mérési eljárások olyan bipoláris vonás-osztályozásokat (*szemantikus differenciál* skálák) használnak, melyek inkább relatív ítéleteket fejeznek ki a kategorikus elutasítás vagy egyetértés helyett. A sztereotípia mérésében a valódi nehézséget természetesen a vizsgálni kívánt tulajdonság-párok meghatározása jelenti. Iskoláskorú gyermekek etnikai sztereotípiáinak mérésére kidolgozott, megfelelően validált mérőeszköz ma sajnos nem áll rendelkezésre. Felnőttekre vonatkozóan létezik ugyan néhány széles körben használt sztereotípia-kérdőív, ezek hatásvizsgálatunkban nem jelentenek nagy segítséget, mivel elsősorban politikai jellegű felmérésekben használják őket, így kérdéseik is általában valamely etnikai csoport politikai megítélésére vonatkoznak; ráadásul a két leggyakrabban használt mérőeszköz¹⁰ kidolgozói nem sok magyarázattal szolgálnak arra vonatkozóan sem, hogy miért és hogyan választották ki a kérdőívben használt attribútumokat. A hatásvizsgálat kérdőívének kidolgozása során elméleti vezérfonalként így az énkép-kutatásokhoz fordultunk. Susan Harter 8-18 éves gyermekekkel folytatott kutatásai során ugyanis azt találta, hogy a gyerekkori önértékelés legfontosabb meghatározója az, hogy a fiatalok mennyiben érzik magukat kompetensnek mindazonon a területeken, amelyek számukra nagy jelentőséggel bírnak (Harter, 1983). Ez arra inspirálta a szerzőt, hogy énkép-tesztje kidolgozása során meghatározza mindazokat a kompetenciaterületeket, melyek a gyerekek önpercepciójának domináns összetevői ebben az életkori szakaszban. Harter így alakította ki énkép-mérőeszköze korábban már ismertetett kategóriáit, s ezeket a kategóriákat adaptáltuk mi is a sztereotípia kérdések kidolgozása során. Ennek további hozadéka, hogy így a gyerekek énképe és a kortárs sztereotípiák közvetlenül összefüggésbe hozhatók egymással¹¹.

¹⁰ *Prejudice Index* (Bobo, Kluegel, 1993), *Racial Stereotype Measure* (Peffley és mtsai, 1997)

¹¹ Ennek jelentőségére Steele és Aronson (1995) hívta fel a figyelmet. Sztereotípia-félelem elméletük szerint a negatív sztereotípiák romboló hatással vannak a gyerekek szociális identitására, ami azután iskolai teljesítményükben is megmutatkozik, mivel hajlamosá válnak rosszabb teljesítményt nyújtani valós képességeiknél a róluk kialakított sztereotípiával konzisztens módon.

Az **előítéletek** a csoportközi kapcsolatok *affektív aspektusai*, az erre fókuszáló kutatási irányvonal így az egyének különböző társadalmi csoportokkal kapcsolatos érzelmeinek az összességét próbálja feltérképezni. Az előítéletek mérésére használt kérdőívek korai formái a kisebbségek meglehetősen durva és nyilvánvaló elutasítására kérdeztek rá, ez azonban mára jelentősen megváltozott; a napjainkban használatos mérőeszközök már sokkal kevésbé támadó jellegűek és reaktívak, mint korábbi változataik. A változás háttérében az állt, hogy a rasszizmus különböző formáinak kutatási eredményeiben egymásnak ellentmondó hosszú távú mintázatok jelentek meg. Míg a véleménykutatások rendre pozitív változásokat jeleztek előre a faji kérdések tekintetében és a kisebbségekkel kapcsolatos negatív sztereotípiák számának csökkenését mutatták, a többségi és kisebbségi csoportok közötti valós kapcsolatokban nem jelentkeztek a várt változások: a területi és iskolai szegregáció szintje magas maradt, és az etnikai csoportok közötti jövedelmi különbségek sem mérséklődtek (McClelland, Linnander, 2006). A magyarázatot sok kutató a rasszizmus új formáinak megjelenésében vélte megtalálni. Az új *rasszizmus* elméletek újrafogalmazták az etnikai attitűdök koncepcióját, lehetőséget adva olyan látens formák megjelenésére, melyek túlmutatnak az alárendelt csoportok iránt érzett haragon és gyűlöleten. Bár a számos ide sorolható irányzat mindegyike független kutatási területet képvisel, közös vonásuk, hogy tagadják a kisebbségi csoportról kialakított negatív sztereotípiák, a velük szemben megnyilvánuló ellenséges érzelmek, és általában is az etnikai diszkrimináció létét, nem csupán elméletben, hanem a mindennapi élet gyakorlatában is. A társadalmat nyitottnak és igazságosnak vélik, a faji egyenlőtlenségekért ők ezért - explicit vagy implicit módon - magukat az alárendelt csoportokat okolják (McClelland, Linnander, 2006). Az előítéletek nyílt és látens formái között megnyilvánuló elméleti kettősség az alkalmazott mérőeszközökben is megjelenik.

A külső csoporttal szembeni általános megítélések régóta használt, máig is időálló eszköze Bogardus (1928) Társadalmi Távolság (Social Distance) Skálája, ami azt hivatott mérni, hogy az egyén milyen mértékű szociális távolságot kíván fenntartani egy adott társadalmi csoport tagjaival, illetve hogy a velük kialakított kapcsolatokban milyen mértékű intimitást tart elképzelhetőnek. A kérdőív eredeti formájában olyan kérdések szerepelnek mint házasságkötés, munkatársi, szomszédi kapcsolatok létesítése, állampolgárság megosztása a célzott csoporttal. A mérőeszköznek ugyanakkor nincs egyetlen, kizárólagosan használt formája, maga Bogardus (1928) is több változatról számol be, később pedig számos más kutató is adaptálta, kiegészítette, vagy módosította az egyes tételeket. Néhányan úgy vélik, a kérdőív inkább a diszkriminatív magatartás mérését szolgálja (Triandis, 1964), az azonban vitathatatlan, hogy a válaszoknak erőteljes érzelmi összetevői is vannak (Biernat és Crandall, 1999). Technikai értelemben a Társadalmi Távolság Skála egy ún. Guttman-skála, melynek tételeire bináris (igen-nem) válaszok adhatók, és aminek egymás utáni kérdései az intimitás növekvő fokát jelzik, így ha valakinek a skálán az összpontszáma 0, az azt mutatja,

hogy az illető nem szeretne semmiféle kapcsolatot kialakítani a kérdéses társadalmi csoport tagjaival, míg a magasabb pontok nagyobb elfogadottságot és feltehetően kedvezőbb attitűdöket jeleznek a kérdéses társadalmi csoport irányában. Évente többszáz kutatás használja a Bogardus skálát, töretlen népszerűsége ellenére viszont nehéz találni két olyan kutatást, ami ugyanazt a formáját alkalmazná (Biernat és Crandall, 1999). Kutatásunkban mi a kérdőív számos létező verziója közül olyan tételket használtunk fel, melyek a kisiskolás (4. osztályos) gyerekek számára is könnyen értelmezhetőek, jól érthetőek voltak. Kabai Péter felvetésére ugyanakkor, miszerint az általunk használt skála itemei „nem egy dimenziót képviselnek, így a skála nem kumulatív, a tényleges válaszok alapján nem lehet a kérdéseket sorrendbe állítani”, két megjegyzést szeretnénk tenni, mivel véleményünk szerint e kritikában két tényező némileg összemosódik. Az általunk használt skála egy-dimenziósságát az elővizsgálat empirikus eredményei megerősítették, a skála ráadásul igen magas belső konzisztenciával rendelkezik (Cronbach alfa=0.89). A kérdőív validálása során a kumulativitás előzetes tesztelésének szükségességéről ugyanakkor megoszlik a kutatók véleménye és gyakorlata. Néhányan valóban ennek szükségességét emelik ki, és ennek megfelelően a megfelelő skála-indexnek az első „nem” válasz sorszámát tekintik, (azaz egy „elfogadási szintet” – ezt javasolja Kabai Péter is), ami így a skála ezt követő kérdéseire adott válaszokat már nem értékeli. Mi magunk azonban azoknak a kutatóknak a véleményét osztjuk, akik szerint nem szerencsés a válaszok sorrendiségét előzetesen „ráerőltetni” a válaszadókra, vagy ez alapján tesztelni a kérdéseket, mert ez jelentős információvesztéshez vezethet. Így például nem derült volna fény arra a Kabai Péter által kiemelt, és véleményünk szerint is érdekes tényre, hogy a nem roma gyerekek több, mint fele lenne szívesen egy roma tanuló barátja, de csak negyede lenne szívesen padtársa. Véleményünk szerint egy ilyen eredmény azonban nem a kérdésfeltevés helytelenségét mutatja, hanem azt, hogy a vizsgált attribútum háttérben olyan tényezők lehetnek, amikre csak a téma további kutatásai deríthetnek fényt. Ehhez a megfontoláshoz alkalmazkodik az index-készítésnek az a módja is, melyben a kutatók a teljes skálára adott „igen” válaszok összpontszámát/arányát mérik (pl. Brewer, 1968), mint ahogyan azt mi is tettük. Ezen lehet szakmailag vitatkozni, a kutatói gyakorlat valóban megoszlik ezzel kapcsolatban (lásd Biernat és Crandall (1999) összefoglalója a Bogardus skála alkalmazásának gyakorlatairól), de Kabai Péternek azt a következtetését, miszerint „az összesített skála nehezen értelmezhető, mert nem tekinthető elfogadási szintnek” nem gondoljuk egy bírálattól méltán elvárható objektív véleménynyilvánításnak.

Két újabb, kérdőíves kutatásokban gyakran használt mérőeszköz létezik emellett a nyílt etnikai előítéletesség mérésére. Az első a Pettigrew és Meertens (1995) nevéhez fűződő *Blatant Prejudice Scale* (BPS), ami két dimenzióra vonatkozó kérdéseket tartalmaz: a kisebbségi csoporttal szembeni *haragot*, illetve a kapcsolatok kialakításának *visszautasítását*. A második a Sidanius és Pratto (1993) által kidolgozott, majd többször módosított (Pratto és

mtsai, 1994) *Social Dominance Orientation* (SDO) skála, ami nem közvetlenül méri az etnikai attitűdöket, inkább az egyén előítéletes gondolkodásra való hajlamát tükrözi. A mérőeszközök a társadalmi hierarchia és egyenlőtlenség elfogadásában és előnyben részesítésében rejlő egyéni különbségeket próbálja feltérképezni. A válaszadó magas pontszáma a skálán azt jelzi, hogy a megkérdezett szerint egyes csoportok értékesebbek, mint mások, hogy a társadalmi csoportok közötti hierarchia elkerülhetetlen és jó, és hogy a dominancia szükségszerű. A gyakorlatban problémát jelent ugyanakkor, hogy etnikai kutatásokban a skálának számos verziója ismert és használt, maga Sidanius és Pratto is a skála összesen hat változatát fejlesztette ki úgy, hogy a tételek közötti átfedés meglehetősen kicsi. A hatásvizsgálat kérdőívében a Sidanius és Pratto által használt mérőeszközök felső tagozatos gyerekek számára jól értelmezhető tételeit szerepeltettük, összesen 6 kérdést. Vissza kell utasítanunk Kabai Péternek azt a kritikai megjegyzését, miszerint a kérdőívünkben feltett kérdések „lényegesen mást mérnek, mint az *'eredeti'* SDO kérdőívek”; azok ugyanis szó szerint megegyeznek a Sidano és Pratto kérdőívében megtalálható itemekkel. A Kabai Péter által kritikával illetett két kérdést példaként hozva: „Vannak olyan emberek, akik értékesebbek, mint mások” (*Some people are more worthy than others*) illetve „A nyereség fontosabb annál, mint ahogyan a játékot játsszuk” (*Winning is more important than how the game is played*) (Sidanius és Pratto, 1993). A nyílt előítéletesség mérésére használt három mérőeszközből adódó eredmények közötti kapcsolatot (Bogardus skála, BPS, SDO) végül az elővizsgálatban külön is vizsgáltuk az esetleges redundáns tételek kiszűrése érdekében, aminek eredményeképpen a BPS „*visszautasítás*” dimenzióját a kérdőív végső, rövidített változatában már nem szerepeltettük, mivel az arra adott válaszokat jól prediktálta a Bogardus skálára adott válaszok mintázata. Ezt Kabai Péter a kérdések „*kicsemegőzésének*” címkézte.

Az előítéletesség nyílt formáitól a *kisebbségi csoportok látens elutasítása* irányába történt változás igen megnehezíti ennek kérdőíves mérését; számos kutató felhívja a figyelmet az elmélet és gyakorlat közötti diszkrepanciára a témában és hiányolja az egyetemlegesen elfogadott mérőeszközök létét (Fiske, 1998). Ráadásul, mivel az új rasszizmus elméletek általában egymástól függetlenül fejlődtek, az egyes kutatásokban használt kérdések variabilitása tovább nőtt. A kérdőívek emellett kivétel nélkül felnőttek részére lettek kialakítva, sajnos egyáltalán nem találtunk a témában releváns, általános iskoláskorú gyerekekre vonatkozó mérőeszközöket vagy kutatási eredményeket. Természetesen így az sem volt egyértelmű, hogy az 'új rasszizmus' által megfogalmazott problémakörrel való gondolkodás egyáltalán kialakul(hat)-e már ebben az életkorban, s ha igen, milyen témakörök mentén és megfogalmazásban kérdezhető ez meg a gyermekektől. A kérdőív kialakításának első lépéseként ezért fókuszcsoportos beszélgetéseket szerveztünk különböző életkori csoportokba tartozó általános iskolás gyermekekkel, melyek témája a hazai etnikai problémák megítélése volt, különös tekintettel a család és a társadalom felelősségvállalására. Röviden

összefoglalva, e beszélgetések a következő tapasztalatokkal zárultak. Körülbelül 12 éves korig a gyermekek nem vagy nehezen értik a társadalmi felelősségvállalás kérdését, szegényebb körülmények között élő kortársaik helyzetéért csak a szülőket okolják, nem ismerik még fel, és nem fogadják el a társadalmi környezet meghatározó erejének lehetőségét. A változás ebben körülbelül 13 éves kortól következik be, amikor a gyermekek elkezdnek politikai döntéshozókról, történelmi múlttól és lemaradó társadalmi csoportokról gondolkodni, megértik ezek potenciális meghatározó erejét. A fókuszcsoporthoz tartozó beszélgetések tükrében az általunk kialakított kérdőívben négy kérdést fogalmaztunk meg a gyermekek számára a modern rasszizmus gondolköre nyomán, melyek arra kérdeznék rá, hogy a gyerekek mennyiben érzik felelősnek hátrányos helyzetükért a szülőket, illetve a társadalmat.

Az előítéletesség mérése kapcsán végül beszélni kell arról a nagyszámú kutatásról is, melyek a csoportközi kontaktus és az előítélet közötti közvetítő mechanizmusokra próbálnak fényt deríteni. Ezek a vizsgálatok általában az érzelmi tényezőkre hívják fel a figyelmet, és a *csoportok közötti szorongás* csökkentésének jelentőségét emelik ki a csoportközi viszonyok javításában (Dovidio és mtsai, 2002). Hatásvizsgálatunkban a társas szorongás mérésére a La Greca és munkatársai (1988) által kidolgozott, széles körben elismert *Társas Szorongás Skálát* használtuk. Az eredeti skála általános iskolás gyerekekre vonatkozóan 10 itemet tartalmaz, amik két faktor köré csoportosulnak: a *negatív megítéléstől való félelem* és a *társas kirekesztettség*. Elővizsgálatunk ezt a faktorstruktúrát közel azonos módon reprodukálta, azzal, hogy a 'társas kirekesztettség' dimenzió két aldimenzióra oszlott, ezért a rövidített skálában mindhárom faktorból szerepeltettünk kérdéseket.

A TÁRSAS KÍVÁNATOSSÁG MÉRÉSE

A ma legelterjedtebb, gyerekek társas kívánatosságát mérő skála a *Children's Social Desirability Scale* (CSDS) (Candall és mtsai, 1965) a felnőttek részére korábban kifejlesztett *Marlowe-Crowne Social Desirability Scale* (MCSD) (Crowne és Marlowe, 1960) mintájára készült, legtöbb kérdése így követi az abban szereplő itemek témáját. A kérdőívben szereplő itemeik egy része olyan hétköznapi magatartásformákat tartalmaz, amelyek társas kívánatosság szempontjából negatívak, másik része pedig ritkán előforduló, de társas kívánatosság szempontjából pozitív viselkedésformákat ír le. A válaszadóknak egyenként kell eldönteniük, hogy szokták-e a kérdéses viselkedést alkalmazni. Crowne és Marlowe (1964) azt feltételezték, hogy a társas kívánatosság skálaértékek háttérben egy olyan egydimenziós személyiségvonás áll, amit ők "*elismerés iránti igény*"-nek neveztek el, és úgy definiálták mint az ember arra való hajlamának mértékét, hogy keresse az elismerést, és lehetőleg elkerülje az elmarasztalást. A későbbi empirikus vizsgálatok ugyanakkor nem támasztották alá azt, hogy a Marlowe-Crowne skála valóban egyetlen vonást mér, ráadásul sok esetben a

társas kívánatosságra való kontrollálás inkább csökkentette, mint növelte a vizsgált konstruktumok előrejelző validitását (McCrae és Costa, 1983). Erre végül Palhus (1984) talált magyarázatot, aki megmutatta, hogy a társas kívánatosság valójában két vonást tartalmaz: az első, az „önmegettevesztő megerősítés” (self-deceptive enhancement, SDE) a személy őszinte, pozitív irányba torzító válaszait foglalja magában, így egészséges énképének természetes velejárója; míg a második, az ún. „jó benyomás keltés” (impression management, IM) a külső megítélő tudatos megettevesztésére irányuló törekvést jelzi. Ebből adódóan minden olyan társas kívánatosság mérőszám, amely magában foglalja az önmegettevesztő megerősítés kérdéseit is (köztük az MCSD illetve a CSDS skála is), csökkenti azoknak a mérőeszközöknek a validitását, melyek olyan vonásokra irányulnak, amik önmaguk is tartalmazzák a pozitív önmegettevesztést (mint pl. szorongás, teljesítmény, motiváció, dominancia, önértékelés). A kérdőíves kutatásokban így csak a tudatosan irányított benyomáskeltés (IM) kontrollálására van szükség a válaszadás torzításának kontrollálása céljából. Az elővizsgálat során így a CSD kérdőív és a Rosenberg önértékelési skála együttes faktor-struktúrájának elemzésével elkülönítettük az „önmegettevesztő megerősítés”-t mérő kérdéseket és a „jó benyomáskeltés” dimenzióhoz tartozó kérdéseket, s hatásvizsgálatunkban csak ez utóbbit szerepeltettük.

ZÁRÓ GONDOLATOK

A hatásvizsgálatoknak a mérések reliabilitására és validitására vonatkozóan kiemelkedően magas technikai elvárásoknak kell megfelelniük, hiszen ezek eredményeit társadalmi programok elszámoltathatóságára és az azzal kapcsolatos komoly, fiskálisan is jelentős döntések meghozatalára használják. Tisztában vagyunk mindazzal a felelősséggel, amit ez ránk ruház. Úgy érezzük ugyanakkor, hogy Kabai Péter nem vetett számot azzal a felelősséggel, amely a sok esetben nem megfelelően megalapozott kritikáinak társadalmi hatásaival jár. Nem gondoljuk, hogy egy hároméves, 60 iskolára kiterjedő kutatásban nem lehet kritizálható részletkérdést találni, mi is sokat fejlődünk a kutatás befejezése óta, esetleges hibáinkból pedig tanultunk/tanulunk. De az alaptalan, nem megfelelő körültekintéssel megírt kritikák nem csak a kutatásnak és magának a programnak a megtervezésében és kivitelezésében résztvevők számára sértőek, hanem hatással vannak az akadémiai kutatások általános megítélésére éppúgy, mint esetünkben maguknak az etnikai integrációs törekvéseknek a társadalmi fogadtatására is.

FELHASZNÁLT IRODALMAK

- Allport, G. W. (1954/1979). *The nature of prejudice*. Cambridge, MA: Perseus Books.
- Biernat, M., Crandall, C. C. (1999). Racial Attitudes. In: J. P. Robinson, P. R. Shaver, L. S. Wrightman (Eds.): *Measures of Political Attitudes*. Academic Press, 297–411.
- Blanden, J., Gregg, P., Macmillan, L. (2006). *Explaining Intergenerational Income Persistence: Non-cognitive Skills, Ability and Education*. Working Paper No. 06/146. Centre for Market and Public Organisation, University of Bristol, Department of Economics.
- Blanden, J., Gregg, P., Macmillan, L. (2007). *Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education*. IZA Discussion Paper No. 2554.
- Blascovich, J., Tomaka, J. (1991). The Self-Esteem Scale. In: J. P. Robinson, P. R. Shaver, L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. New York: Academic Press, 115-160.
- Bobo, L., Kluegel, J. R. (1993). Opposition to Race-Targeting: Self-Interest, Stratification Ideology, or Racial Attitudes? *American Sociological Review* (58) 443-464.
- Bogardus, E. (1928). *Immigration and Race Attitudes*. Boston: Heath.
- Bowles, S., Gintis, H., (2001). The Inheritance of Economic Status. *Journal of Economic Perspectives*. 16(3) 3–30.
- Brewer, M. B. (1968). Determinants of social distance among East African tribal groups. *Journal of Personality and Social Psychology* (10) 279-289.
- Cawley, J., Heckman, J. J., Vytlačil, E. (2001). Three observations on wages and measured cognitive ability. *Labour Economics* 8(4) 419–442.
- Collins, W. A., Gunnar, M. R. (1990). Social and personality development. *Annual Review of Psychology* (41) 387-416.
- Crowne, D. P., Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology* (24) 349-354.
- Crowne, D. P., Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*. New York: Wiley.
- Dovidio, J. F., Gaertner, S. L., Kawakami, K., Hodson, G. (2002). Why can't we just get along? Interpersonal biases and interracial distrust. *Cultural Diversity and Ethnic Minority Psychology* (8) 88–102.
- Feinstein, L. (2000). *The relative economic importance of academic, psychological and behavioural attributes developed in childhood*. CEP Discussion Paper No. 443.
- Fiske, S. T. (1998) Stereotyping, prejudice, and discrimination. In: Gilbert, D., Fiske, S. T., Lindzey, G. (Eds.): *The Handbook of Social Psychology*, 4th ed., New York: McGraw-Hill, 357-411.
- Gray-Little, B., Williams, V. S. L., Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin* (23) 443-451.
- Harter, S. (1983). Developmental perspectives on the self-system. In: Hetherington, E. H. (Ed.): *Handbook of Child Psychology*. Vol. 4. *Socialization, personality and social development*. New York: John Wiley, 275-385.
- Harter, S. (1985). *Manual of the Self-Perception Profile for Children*. University of Denver: Denver Co.

- Heckman J. J., Rubinstein, R. (2001) The Importance of Noncognitive skills: Lessons from the GED Testing Program. *American Economic Review* 91(2) 145-149.
- Heckman J. J., Stixrud, J., Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics* 24(3) 411-482.
- Heckman, J. J. (2000). Policies to Foster Human Capital. *Research in Economics* 54(1) 3-56.
- Heckman, J. J., Masterov, D. V. (2007). *The Productivity Argument for Investing in Young Children*. NBER Working Paper No. 13016.
- Hu, Y., Schennach, S. (2008). Instrumental Variable Treatment of Nonclassical Measurement Error Models. *Econometrica* (76)1, 195-216.
- Judge, T. A., Erez, A., Bono, J. E., Thoresen, C. J. (2002). Do the traits self-esteem, neuroticism, locus of control, and generalized self-efficacy indicate a common core construct? *Journal of Personality and Social Psychology* (83) 693-710.
- Judge, T. A., Locke, E. A., Durham, C. C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior* (19) 151-188.
- Kagan, S. L., Moore, E., Bredekamp, S. (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Washington, DC: National Education Goals Panel.
- Kézdi, G., Surányi, É. (2008a). *Egy sikeres iskolai integrációs program tapasztalatai*. Kutatási összefoglaló. Educatio Kht., Budapest. <http://bin.sulinet.hu/tanar/szabadido/planetarium/8.1.1.pdf>
- Kézdi, G., Surányi, É. (2008b). Egy integrációs program hatása a tanulók fejlődésére. Educatio, 2008/IV. <http://www.hier.iif.hu/hu/letoltes.php?fid=tartalomsoor/1250>
- La Greca, A. M., Dandes, S. K., Wick, P., Shaw, K., Stone, W. L. (1988). Development of the Social Anxiety Scale for Children: Reliability and Concurrent Validity, *Journal of Clinical Child Psychology* (17) 84-91.
- Marsh, H. W. (1996). Positive and negative self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology* (70) 810-819.
- McClelland, K., Linnander, E. (2006). The Role of Contact and Information in Racial Attitude Change among White College Students. *Sociological Inquiry* 76(1) 81-115.
- McCrae, R. R. Costa, P. T. (1983). Social Desirability Scales: More Substance than Style, *Journal of Consulting and Clinical Psychology* (51) 882-888.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Peffley, M., Hurwitz, J., Sniderman, P. M. (1997). Racial Stereotypes and Whites' Political Views of Blacks in the Context of Welfare and Crime. *American Journal of Political Science* (41) 30-60.
- Pettigrew, T. F. (1998). Intergroup Contact Theory. *Annual Review of Psychology* (49) 65-85.
- Pettigrew, T. F., Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology* (25) 57-75.
- Pettigrew, T. F., Tropp, L. R. (2000). Does intergroup contact reduce prejudice? Recent metaanalytic findings. In: S. Oskamp (Ed.): *Reducing prejudice and discrimination*. Hillsdale, NJ: Erlbaum, 93-114.
- Pratto, F., Sidanius, J., Stallworth, L. M., Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology* (67) 741-763.
- Reynolds, C. R. (1979). What I think and feel: a revised measure of children's manifest anxiety. *Journal of Abnorm Child Psychology* (6) 271- 80.

- Robins, R. W., Hendin, H. M., Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin* (27)2, 151-161.
- Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton University Press: Princeton, NJ.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs* 33(1), 300-303.
- Sidanius, J., Pratto, F. (1993). The inevitability of oppression and the dynamics of social dominance. In: P. M. Sniderman, P. E. Tetlock, E. G. Carmines (Eds.): *Prejudice, politics, and the American dilemma*, Stanford, CA: Stanford University Press 173–211.
- Surányi, É.; Danis, I. (2009). *A Biztos Kezdet program hatásvizsgálatának kutatási terve*. Biztos Kezdet program (TÁMOP 5.2.1.).
- Surányi, É.; Danis, I. (2010). *Családpolitika más szemmel*. Kiadás alatt álló kézirat. Készült a Gazdasági és Szociális Tanács megbízásából.
- Taylor, D. H. , Rickel, A. U. (1981). An Analysis of Factors Affecting School Social Segregation. *Journal of Negro Education*, Vol. 50, No 2.
- Triandis, H. C. (1964). Exploratory factor analyses of the behavioral component of social attitudes. *Journal of Abnormal and Social Psychology* (68) 420-430.

A sorozat korábban megjelent kötetei

2008

Galasi Péter – Nagy Gyula: A munkanélküli-jövedelemtámogatások célzása Magyarországon 2004-ben. BWP 2008/01

Szilvia Hámori: Adult education in the European Union – with a focus on Hungary. BWP 2008/02

Galasi Péter: A felsőfokú végzettségű munkavállalók munkaerő-piaci helyzete és foglalkozásuk-iskolai végzettségük illeszkedése. BWP 2008/03

Anna Lovász: Competition and the Gender Wage Gap: New Evidence from Linked Employer-Employee Data in Hungary, 1986-2003. BWP 2008/04

Zsuzsa Blaskó: Does early maternal employment affect non-cognitive children outcomes? BWP 2008/05

Galasi Péter: A túl- és az alulképzés bérhozama 25 országban. BWP 2008/06

Galasi Péter - Nagy Gyula: Az aktív munkaerő-piaci programokba kerülés esélyei: képzés, bértámogatás, közhasznú munka. BWP 2008/07

Galasi Péter: The effect of educational mismatch on wages for 25 countries. BWP 2008/08

Gábor R. István: Életpálya-kereseti profilok és keresetingadozás. BWP 2008/09

2009

Istvan Gábor R.: Experience-earnings profile and earnings fluctuation: a missing piece in some labour market puzzles? BWP 2009/01

Anna Lovász – Mariann Rigó: Who Earns Their Keep? An Estimation of the Productivity-Wage Gap in Hungary 1986-2005. BWP 2009/02

Köllő János: Miért nem keresnek állást a magyar munkanélküliek? Hipotézisek az Európai Munkaerőfelvétel adatai alapján. BWP 2009/03

Bálint Mónika - Köllő János - Molnár György: Összefoglaló jelentés a KSH-ONYP adatfelvételről. BWP 2009/04

Gábor R. István: "Minimálbér-paradoxon" - versenyzői munkaerőpiacon? Egy gondolkísérlet tanulságai. BWP 2009/05

Kertesi Gábor - Kézdi Gábor: Általános iskolai szegregáció Magyarországon az ezredforduló után. BWP 2009/06

Szilvia Hámori: Employment convergence of immigrants in the European Union. BWP 2009/07

Gábor Kőrösi: Innovation and Rent Sharing in Corporate Wage Setting in Hungary. BWP 2009/08

A Budapesti Munkagazdaságtani Füzetek a Magyar Tudományos Akadémia Közgazdaságtudományi Intézetében működő Munkapiaci Kutatások valamint a Budapesti Corvinus Egyetem Emberi Erőforrások Tanszékének közös kiadványa. A kiadványsorozat angol nyelvű füzetei **“Budapest Working Papers on the Labour Market”** címmel jelennek meg. A kötetek letölthetők az MTA Közgazdaságtudományi Intézet honlapjáról: <http://www.econ.core.hu>