



## A KOLLABORATÍV PROBLÉMAMEGOLDÓ KÉPESSÉG MÉRÉSÉNEK ELMÉLETI ÉS MÓDSZERTANI MEGFONTOLÁSAI: EGY PILOT KUTATÁS EREDMÉNYEI

**Pásztor-Kovács Anita**

*Szegedi Tudományegyetem Neveléstudományi Doktori Iskola*

Az infokommunikációs technológia (IKT) rendkívüli sebességű fejlődése, az ezzel – többek között – összefüggésben jelentkező nagymértékű gazdasági és társadalmi változások a munkaerő-piaci igényeket is átrendezik (Kozma, 2009; Molnár, 2011). A 21. század egyre inkább újszerű, 21. századi képességeket követel munkavállalóitól (Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci, & Rumble, 2012). A kollaboratív vagy csoportos problémamegoldó képesség egyike ezeknek a modern munkahelyeken nélkülözhetetlen képességeknek (Neubert, Mainert, Kretzschmar, & Greiff, 2015). Vitathatatlan fontosságú fejlesztéséhez megfelelő mérőeszközökre is szükség van, melyekkel a fejlődés mértékét nyomon követhetjük. Annak ellenére, hogy aktuálisan több nemzetközi kutatócsoportnak is célja a konstruktum empirikus meghatározása, nincsen konszenzus arról, hogy pontosan mit értünk alatta. Kevés empirikus tapasztalattal rendelkezünk a képesség kapcsán, illetve számos vitás kérdés megválaszolása, módszertani akadály leküzdése előtt állunk egy jól működő, a képesség mérését lehetővé tevő mérőeszköz kialakításának útján.

Jelen tanulmányban ismertetjük a kollaboratív problémamegoldó képesség mérésének módszertani kihívásait, illetve az eddig született megoldásokat, mérési irányzatokat. Ezzel összefüggésben saját, a képességre irányuló kutatásunk módszertani megfontolásait és első eredményeit, tanulságait is közöljük.

### **Mit nevezünk kollaboratív problémamegoldó képességnek?**

A kollaboratív problémamegoldást mint folyamatot, illetve az azt lehetővé tevő képességet többféle módon, eltérő összetettséggel és mélységgel definiálják. O’Neil, Chuang és Chung (2003) Zhang (1998) meglehetősen tág értelmezését fogadják el, miszerint a kollaboratív problémamegoldás olyan problémamegoldó tevékenységekre utal, amelyek egyének egy csoportjának interakcióit foglalják magukban.

Az ACT21S (*Assessment & Teaching of 21<sup>st</sup> Century Skills*) projekt, ami a 21. századi képességek feltérképezésére és vizsgálatuk különböző eljárásainak kidolgozására vállalkozott, szintén célul tűzte ki e képesség mérését (Griffin, McGaw, & Care, 2012; Griffin & Care, 2015). Ebben a leírásban a kollaboratív problémamegoldás közös tevékenységet jelent, mely során párok vagy kisebb csoportok lépések sorozatát hajtják végre, hogy egy

adott állapotból a kitűzött célállapotba jussanak (Griffin, Care, Bui, & Zoanetti, 2013; Hesse, Care, Buder, Griffin, & Sassenberg, 2015).

A képesség a 2015-ös PISA<sup>1</sup> negyedik, kiegészítő területének is tárgyát képezi a 2012-es kreatív problémamegoldás után (Greiff, Holt, & Funke, 2013). A mérés támpontjául szolgáló elméleti keret magát a folyamatot nem, csak a képességet definiálja: Kollaboratív problémamegoldó képességnek<sup>2</sup> az egyén azt a fajta kapacitását értjük, aminek segítségével hatékonyan bekapcsolódik egy folyamatba, ahol kettő vagy több résztvevő megkísérel megoldani egy problémát a megoldáshoz szükséges megértés és erőfeszítés megosztásával, tudásuk, készségeik és erőfeszítéseik egyesítésével (OECD, 2013, p. 6.).

Bár a szerzők minden esetben építenek a kollaboráció, kooperáció, kollaboratív és kooperatív tanulás fogalmakra, a kooperatív és a kollaboratív problémamegoldás között nem tesznek különbséget. Mérészközeikben keverten alkalmazzák azok elemeit, például a pozitív interdependencia (kölcsonös függés) kondíciót a mozaik típusú problémákban, mely elem kifejezetten a kooperatív tanulás sajátja, holott egy kooperatív, illetve kollaboratív munka- vagy tanulási folyamat gyökeresen más elemekkel, csoportdinamikai folyamatokkal jár. Ezekkel mind a definícióban, mind a mérőeljárás összeállításánál számolnunk szükséges. Ennek megfelelően saját definíciónk a következőképpen hangzik: „Kollaboratív problémamegoldó képességnek azt a komplex képességet nevezzük, amely az egyént alkalmassá teszi arra, hogy egy közös cél érdekében, melynek elérése új, a csoporttagok eddig bevált stratégiáival nem kezelhető akadályok leküzdését igényli, egy vagy több személlyel együttműködő tevékenységet végezzen. Az együttműködő tevékenység megengedi, de nem írja elő a probléma részekre bontását, egyéni szerepek vállalását, a csoporttagok egymás munkájáért nem vonhatók felelősségre” (Pásztor-Kovács, 2015, p. 9.).

A képességet rendszerint részképességek gazdag halmazaként jellemzik, melyek két fő, kognitív (vagy problémamegoldás) és szociális (vagy kollaboráció) komponens alá sorolhatóak be (Hesse et al., 2015; OECD, 2013; O’Neil et al., 2003). Ezek a modellek nem feltételeznek kapcsolatot a két komponens között, illetve szimmetrikusnak, azonos súlyúnak kezelik a kettőt. Saját modellünkben (l. Pásztor-Kovács, 2015) feltüntetünk kapcsolódási pontokat a két komponens között, illetve hangsúlyozzuk azt, hogy nem tekinthető törvényszerűnek a két komponens mellérendelt viszonya. Ezen kérdés, a két komponens egymással való viszonyának vizsgálata mindazonáltal empirikus módszereket igényelne, ennek hiányában mindössze feltételezésekkel élhetünk. A kollaboratív problémamegoldást érdemes önálló konstruktként kezelni, nem tekinthetjük azt pusztán egyének problémamegoldásának csoportos kontextusban (Care & Griffin, 2014).

### A képesség mérésének módszertani problematikája

A képesség vizsgálata módszertani szempontból számos tekintetben nehézségekbe ütközik (Krkovic, Pásztor-Kovács, Molnár, & Greiff, 2014; Pásztor-Kovács, 2013; Ras,

<sup>1</sup> Programme for International Student Assessment

<sup>2</sup> A competency-t képességnek fordítottuk, ennek indoklását lásd Pásztor-Kovács (2015).

Krkovic, Greiff, Tobias, & Maquil, 2014). Ez elsősorban a képesség komplexitásából fakad, a több különböző képességelem egyidejű mérése nehezen kivitelezhető. Ugyanis a fő komponenseket jelölő kognitív és szociális képességelemek, eltérő jellegük miatt, teljesen más mérési hagyományokkal rendelkeznek.

A problémamegoldó gondolkodás egyik legelterjedtebb vizsgálati eljárása napjainkban a teszt módszer. Számos problémamegoldó teszt feladatsorral gazdagodott a szakirodalom az elmúlt évtizedekben (l. pl. Molnár, Greiff, & Csapó, 2013; Molnár, Greiff, Wüstenberg, & Fischer, 2014; Wüstenberg, Greiff, Molnár, & Funke, 2014; OECD, 2010). A szociális képességeket ezzel szemben jellemzően megfigyelésekkel, interjúval vagy kérdőívvel vizsgálják (O’Neil et al., 2003). Ennek tükrében a problémamegoldó, illetve a szociális teljesítmény mérésének összehangolása rendkívül összetett feladat. Feltehetően ennek is köszönhető az a tendencia, hogy az elmúlt évtizedek csoportos problémamegoldásra irányuló empirikus vizsgálatai hagyományosan a csoport és nem az egyén teljesítményére fókuszáltak. Rendkívül kevés olyan mérési tapasztalattal rendelkezünk, amely egy csoportos problémamegoldó helyzetben az egyén képességszintjét próbálná meg valamilyen módon meghatározni (Greiff, 2012).

Az általánosíthatóság problémája a képesség mérése kapcsán szintén hatványozottan jelentkezik, hiszen a teljesítményt minden esetben befolyásoló faktorok mellett (pl. érzelmi jólét, bioritmus) a csoportösszetétel hatásával is számolni kell. Egy kollaboratív problémahelyzet végtelen számú kifejtést vehet fel. Annak érdekében, hogy standardizált teszt környezetet teremtsünk, azaz képesek legyünk megegyező ingereket adni a vizsgálati személyeknek, mindezt egy megoldás születt: egy olyan mesterséges problémamegoldó helyzet kialakítása, ahol a másik felet nem egy valós kollaboráló partner, hanem egy számítógépes „avatar” jelenti. A szakirodalom „Human-Agent” eljárásnak nevezi ezt az elrendezést (OECD, 2013; Graesser, Foltz, Rosen, Forsyth, & Germany, in press; Rosen, 2014a; Rosen & Tager, 2013).

Bár egyelőre az egyetlen standardizált tesztelési lehetőséget jelenti, a Human-Agent eljárás validitás szempontjából komoly kétségekre ad okot, ugyanis ezek a szoftverek limitált, előre programozott válaszkészletből generálják reakcióikat, az érzelmek csupán szűk keresztmetszetének kifejezésére képesek. Holott az érzelmek kifejezése, az adekvát reagálás azokra, a konfliktusok kezelése nem elhanyagolható részét képezik a kollaboratív kommunikációs mintázatoknak. Kérdés tehát, hogy milyen mértékben képezi le a valóságot ez az elrendezés. Mindazonáltal, még ha nem is kívánunk avatart beépíteni a mérésbe, egyértelműen olyan eljárást érdemes kialakítani, amely kihasználja a modern IKT-eszközök nyújtotta előnyöket (pl. gazdaságosság, gyors, objektívabb visszacsatolás, innovatív itemtípusok alkalmazásának lehetősége) az egyre inkább korszerűtlennek minősített papír-ceruza teszteléssel szemben (Csapó, Ainley, Bennett, Latour, & Law, 2012; Molnár, 2010; Scheuermann & Björnsson, 2009).

Adatelemzés szempontjából a legpraktikusabb megoldást az írásban, azaz chaten vagy e-mailen keresztül történő kommunikáció jelenti. Azonban ez a kondíció mellőzi a nem verbális kommunikációs csatornákat, ami komoly veszteséget jelent annak tudatában, hogy kommunikációnk jelentős mértéke, egyes szerzők szerint kétharmada ilyen jelek használatán keresztül zajlik (Birdwhistell, 1970). Ezen csatorna megléte vagy hiánya tehát

nyilvánvalóan befolyással bír a csoportos problémamegoldás kimenetelére. A másik arcának, hangjának ismeretében egyéb befolyásoló tényezők is közbeszólhatnak viselkedésünkbe (pl. a kapcsolat mintázata, amennyiben már ismerjük az illetőt, vonzalom, holdudvarhatás) (l. pl. Dijksterhuis & Bargh, 2001; Moore, Filippou, & Perret, 2011).

A face-to-face kommunikáció alternatív, online útját videochat-szoftver alkalmazása biztosíthatná. Számos információ elvesztését kerülhetnénk el ezzel a megoldással, kérdés azonban, hogy mit tudnánk kezdeni azzal a hatalmas mennyiségű adattal, amit egy ilyen típusú mérés szolgáltatna, hogyan strukturálnánk, hogyan értékelnénk az interakciókat. Elérhetőek olyan szoftverek, amelyek felhasználhatók lehetnének bizonyos érzelmek detektálására arckifejezések elemzésével. Ezek az eszközök komoly potenciált hordoznak a jövőbeli kutatások kapcsán (Csapó, Lőrincz, & Molnár, 2012), egyelőre azonban felhasználhatóságuk meglehetősen korlátozott.

További problémát jelent, ha mérőeszközünket nagymintás mérésekre tervezzük, hiszen a megfelelő sávszélességű internet, nagy felbontású monitor, fejhallgató, esetleg mikrofon elérhetősége nem feltétlenül egyformán biztosított az adatfelvételi helyszíneken (Molnár & Pásztor-Kovács, 2015). Komoly erőforrásokat emészthet fel ezeknek az eszközöknek a beszerzése vagy kiköszvetítése, esetleg a vizsgálati személyek utaztatása. Hasonlóan nagy ráfordítást igényelne az interakciók szövegének begépelése, ami az elemzésekhez elkerülhetetlennek látszik. Mindezek ismeretében az írásbeli kommunikáció tűnik a legkézenfekvőbb megoldásnak. Mivel az online kommunikáció, sőt problémáink társas megoldása online kommunikációs úton egyre természetesebb része a hétköznapijainknak, a chatsatorna alkalmazása a 21. században már nem tekinthető életszerűtlen eljárásnak (Hermann, Rummel, & Spada, 2001).

Ha nagymintás mérést tervezünk, olyan eljárás kidolgozása szükséges, amely a lehető legkevesebb erőforrást igényli az adatok feldolgozásához. Mérőeszközünk optimális esetben képes az automatikus kódolásra, ennek kapcsán akár azonnali visszacsatolás nyújtására is. Amennyiben azonban nem korlátozzuk semmilyen módon az üzenetváltást, az interakciók kézi kiértékelése nem tűnik kiküszöbölhetőnek. Tartalomelemző szoftverek segítségével ez a folyamat némileg meggyorsítható, ám így is rendkívül hosszú időt vehet igénybe az adatelemzés.

Mindezidáig egy olyan eljárás ismeretes a kollaboratív problémamegoldó képesség vizsgálatára, amely alkalmas az automatikus kódolásra: az előre definiált, kötött üzeneteken keresztül zajló kommunikáció. A vizsgálati személyeknek meghatározott üzenetkészletből kell kiválasztaniuk, hogy melyik felel meg leginkább aktuális mondandójuknak, a kötött üzenetek váltásán kívül/felül nincs mód a szabad kommunikációra. Az automatikus értékelés előnyén túl ez a módszer is problémáktól terhes, hiszen rendkívül rugalmatlan kommunikációhoz vezet, frusztrálhatja a vizsgálati személyeket, nem beszélve az üzenetek esetleg sugalmazó hatásáról, ami befolyásolhatja a megoldás menetét (Chung, O'Neil, & Herl, 1999; Krkovic et al., 2014; O'Neil, Chung, & Brown, 1997).

Mindezek fényében alapos mérlegelés és kompromisszum kialakítása szükséges a kutató részéről, ha e képesség vizsgálatára vállalkozik. A következőkben – az eddigiekkel szoros összefüggésben – ismertetjük, milyen kompromisszumok születtek eddig, azaz a képesség mérésének milyen megoldásai, trendjei bontakoztak ki.

### **A kollaboratív problémamegoldó képesség mérésére irányuló empirikus tapasztalatok**

A kollaboratív problémamegoldó képesség kapcsán több szempontból is megemlíthető a kaliforniai CRESST (*National Center for Research on Evaluation, Standards, and Student Testing*) központ kutatássorozata (Chung et al., 1999; Hsieh & O'Neil, 2002; O'Neil et al., 1997, 2003). Méréseiket már az 1990-es évek végén számítógép alapon folytatták. A két- vagy háromfős csoportok chaten keresztül kommunikálva oldották meg a kezdetben konszenzus kiépítését igénylő problémát, a későbbi vizsgálatokban pedig egy tudástérkép felépítésének feladatát természettudományos fogalmakból. Elsőként alkalmazták kutatásaikban a kötött üzenetváltást hol vagylagos, hol kizárólagos üzenetserelehetőségként, és számoltak be ennek a feltételnek a működőképességéről. Hátrányait is számba vették, például a vártnál alacsonyabb problémamegoldó teljesítményt, amit a hosszú üzenetlista böngészésére fordított többlet kognitív kapacitásnak tulajdonítottak, vagy a vizsgálati személyek konzekvensen jelentkező frusztrációját, amit a kommunikáció szigorú korlátozása okozott. Az üzeneteket kollaboratív problémamegoldó modelljük kollaboráció komponensének hat eleméhez kötötték. A szociális komponens ezzel a módszerrel – egyedülálló módon – a csoportos mellett egyén szintjén is vizsgálhatóvá vált, még hozzá számítógépen keresztül, automatikusan kódolhatóan. Ugyanakkor a kognitív vagy problémamegoldó komponens objektív vizsgálatát és automatikus kiértékelését továbbra is csak csoportszinten tették lehetővé. Egyén szintjén mindössze egy önkitöltős kérdőívből nyertek adatokat az önmagunkba vetett hittel és a probléma megoldására tett erőfeszítéssel, valamint a tervezéssel és az önmonitorozással összefüggésben, melyek a problémamegoldás komponens elemeit képezik. A kutatássorozat bár nem kifejezetten a kollaboratív problémamegoldó képesség egyénszintű mérését valósította meg, eredményei komoly jelentőségűek és jól felhasználhatóak a jelenleg folyó vizsgálatok tervezésénél, kivitelezésénél.

Az ATC21S projekt nagymintás vizsgálata jelenleg az egyetlen olyan kutatás a sajátunkkal párhuzamosan, amely „Human-Human” elrendezést alkalmazott, azaz humán interakciót, gyerekek kollaborációját vizsgálta számítógép alapon, avatar nélkül (Griffin & Care, 2015). Méréseikben párok dolgoztak együtt dinamikus, interdependens, vegyesen tartalomfüggetlen és matematikai, valamint természettudományos tartalmú problémák megoldásán (Care, Griffin, Scouler, Awwal, & Zoanetti, 2015; a problémátípusokról részletesen l. Molnár, 2001, 2006a; Pásztor-Kovács, 2015). A vizsgálati személyek ezúttal is, akár csak a többi vizsgálatban, chaten keresztül kommunikáltak egymással. Az üzenetváltás azonban nem korlátozódott előre definiált üzenetek cseréjére, a párok szabadon begépelhettek és elküldhettek egymásnak bármely szöveget a chat segítségével. A kérdés ez esetben is az, hogy miként használható fel a szabad kommunikációból származó hatalmas mennyiségű adat, megvalósulhat-e bármilyen formában automatikus értékelése a korlátlan üzenetváltásnak. A szerzők komoly lépéseket tettek ebbe az irányba, ám úgy tűnik, az automatikus értékelés egyelőre a chatszövegek felszíni jegyei alapján fog megvalósulni későbbi méréseikben. Meghatározták, hogy az egyes problémák modelljük mely részképességeit tették mérhetővé, az adott részképességekhez pedig változókat rendeltek, pél-

dául mennyi az összes megszólalás száma, mekkora az első megszólalásig eltelt idő, megjelenik-e egy megfelelő kulcsszó a válaszban, szerepel-e kérdőjel az üzenetekben bizonyos pontokon és hány darab (Adams, Vista, Scoular, Awwal, Griffin, & Care, 2015). Ezeknek a tulajdonságoknak a felhasználása valóban kézenfekvő, hiszen relatíve könnyedén kvantifikálhatóak, azonban kizárólag ezekre alapozva döntést hozni az egyén képességszintjéről anélkül, hogy a társalgás tényleges tartalmát ellenőriznénk, torz következtetésekhez vezethet. A Rasch- és parciáliskredit-modellezés, mellyel a kutatók mérésük validitását vizsgálták, pusztán a változók belső működéséről ad információt, azt továbbra sem garantálja, hogy a kódolásból nyert adatok ténylegesen a célképeséget mérik (Griffin, Care, & Harding, 2015). Nem ismertetnek külső validációs eljárást sem, ami alternatív megoldásként felmerülhetne. Mindezek tükrében, bár a projekt tanulságai jelentősen hozzájárulnak az automatikus kódolás, ezzel együtt egy majdani optimális mérőeszköz kialakításához, eredményei felhasználásakor figyelembe kell vennünk a kutatás korlátait is.

Harmadik empirikus tapasztalatunk már a Human-Agent elrendezést követi. Az egye-sült államokbeli Pearson cég kutatói pilot vizsgálatukban egyelőre egy darab tartalom-független, dinamikus itemmel mérték a vizsgálati személyek képességét páros kollaborációs helyzetben (Rosen, 2014b; Rosen & Foltz, 2014; Rosen & Tager, 2013). A pár ezúttal is chaten keresztül, előre definiált üzenetekkel kommunikált. Ezek száma limitált volt, 5-7 üzenet közül választhatott az üzeni kívánó. Azt is előre programozták, hogy milyen szituációkban milyen üzenetek közül választhassanak a vizsgálati személyek, a lista ugyanis folyamatosan módosult. A kommunikációt ily módon két szempontból is determinálták a kutatók, amit egyrészt elfogadhatunk jól felfogott célként is az automatikus kiértékelés érdekében, másrészt viszont ezúttal is felmerülhet kétség a validitás kapcsán. A limitált üzenetkészlet nem tartalmazta például a való életben szükségszerű *miért* kérdést, azaz nem volt lehetőség a megoldási javaslat indoklásának kérésére, ahogy adására sem. Nem volt információ továbbá arról, hogyan történt az üzenetek kialakítása, zajlott-e olyan előzetes vizsgálat, ahol Human-Human kondícióban, szabad kommunikációval történt a problémamegoldás. Módszertani szempontból ugyanis az lenne optimális, ha az előre definiált üzenetek egy ilyen jellegű pilot vizsgálat eredményeire épülnének. A pontozás logikája sem teljesen egyértelmű, nem derült ki, hogy a részpontok miért az adott súllyal szerepeltek az összpontszám kialakításában (pl. a problémamegoldó teljesítmény miért csak 26%-a az összpontszámnak). A vizsgálat jelentősége mindazonáltal vitathatatlan, hiszen elsőként ülteti át a Human-Agent eljárást működőképesen a gyakorlatba.

Az eddig kivitelezett kis számú vizsgálat, bár bizakodásra ad okot, jól demonstrálja a kollaboratív problémamegoldó képesség mérésének módszertani nehézségeit is. Számos alapkutatásra lenne még szükség ahhoz, hogy egy általánosan elfogadott, kielégítő mértékben érvényesnek tekinthető mérőeljárás szülessen a képesség vizsgálatára.

## Módszerek

### A kutatás célja

Méréssorozatunkban a hiányzó alapkutatások egyikét kívántuk megvalósítani. Arra a kutatási kérdésre kerestük a választ egy majdani összehasonlító vizsgálat keretein belül, hogy az előre definiált üzenetsere a szabad kommunikációval szemben hogyan befolyásolja a csoportok problémamegoldó teljesítményét, illetve miként alakítja, módosítja a kommunikációs mintázatát a kollaborációnak, milyen mértékben tekinthető a két különböző kondícióban zajló együttműködés ekvivalensnek. Alapvető kíváncsi lenni, validitási szempontokat figyelembe véve, hogy a két típusú kommunikáció gyökeresen ne térjen el egymástól, azonban e feltétel teljesülése empirikusan nem ellenőrzött.

A kérdés vizsgálatára első lépésben fel kellett építenünk egy online tesztkörnyezetet, be kellett vezetnünk a hazai mérés-értékelési gyakorlatba egy teljesen új chat funkciót. Jelen kutatás célja a számos új, innovatív elemet tartalmazó teszt működőképességének, megbízhatóságának, problémái konzisztenciájának vizsgálata volt, ezen felül annak feltérképezése, hogy a vizsgálati személyek milyen attitűdöket, reakciókat mutatnak az újszerű teszteljárással kapcsolatban.

### Minta, eljárás

Kutatásunkban 71, a Szegedi Tudományegyetem különböző alapképzéseiben részesülő hallgató vett részt. A teszt kitöltése során a résztvevők a problémákat csoportban oldották meg, összesen 17 csoport eredményeit regisztráltuk.

A bemutatott képességet vizsgáló eljárások mindegyike két vagy három fő kollaborációját mérte. Az eddigiekkel ellentétben mindenképpen célszerűbbnek tartottuk a nagyobb létszámú csoportot, több okból. A kétfős elrendezést nem tartottuk szerencsésnek, hiszen nem meríti ki egyértelműen a csoport fogalmát, egyes szerzők három főben határozzák meg a csoport minimális létszámát (pl. Keyton, 2002). Szempont volt továbbá a már ismertett általánosíthatóság problematikája. A csoport összetétele az egyik legdominánsabb változó az egyén viselkedésének alakulásában. Minél nagyobb a csoport, annál hatékonyabb kommunikációt igényel az egyéntől, hiszen több taghoz szükséges az alkalmazkodás (Levine & Moreland, 1998). Ezzel összefüggésben az is feltételezhető, hogy a nagyobb csoportlétszám általánosíthatóbb eredményhez vezethet, hiszen az egyén viselkedésspektruma is szélesebb spektrumon mozoghat. Nem lehet azonban túl nagy sem a csoport, hiszen az kontrollálhatatlan méréshez, túl sok változóhoz vezetne. Mindezt a négyfős csoportméret mellett döntöttünk. A csoportok többsége (12 db) ennek megfelelően négyfős volt, azonban mindhárom mérési pontban keletkeztek ettől eltérő létszámú csoportok is, hiszen az aktuális résztvevőszám nem volt négyvel osztható egy esetben sem. A rendszer ezt úgy kezelte, hogy négy ötfős és egy háromfős csoportot is létrehozott.

Az adatfelvétel 3 mérési pontban zajlott. A vizsgálati személyek a mérések ideje alatt ugyanabban a számítógépes teremben tartózkodtak, ennek ellenére csak virtuálisan kommunikáltak egymással.

## **Eszközök**

A mérőeszközt az eDia online diagnosztikus platformon (Molnár & Csapó, 2013) keresztül közvetítettük ki. A vizsgálati személyek, miután előre kiosztott mérési azonosítójukat beírva beléptek a mérés felületére, megismerkedhettek a mérés céljával, leírást kaptak a problémamegoldás és a csoportos kommunikáció módjáról. Tovább lépve, a szerver random módon kettő, három vagy négy másik személyhez rendelte őket. A csoportok megalakulása után a teszt következő oldalán a chat funkció kipróbálása következett, a csoporttagoknak 90 másodperc állt rendelkezésükre a chat kipróbálására, ezzel együtt az ismerkedésre.

A 90 másodperces ismerkedő oldal után a teszt automatikusan továbbléptette a csoport tagjait a következő oldalra. Ez az oldal azzal kapcsolatban tartalmazott leírást, hogy hogyan lehetséges a megoldások csoportos megadása, illetve egy rendkívül egyszerű pilot feladatot (ki kellett választaniuk négy opció közül Magyarország fővárosát) prezentált, melyen keresztül az olvasottakat ki is próbálhatták.

Ezt követően, tovább lépve, megkezdődött a teszt problémamegoldó része. Négy problémát kellett a csoportoknak együttesen megoldani anélkül, hogy a megoldásra szánt időt korlátoztuk volna. A problémamegoldás után egy kérdőív következett, ezen a ponton a vizsgálati személyek már szétkapcsolódtak, önállóan fejezték be a teszt kitöltését. A mérés a csoport problémamegoldó teljesítményéről nyújtott százalékos visszacsatolással zárult.

### **A csoportkommunikáció módja**

Az üzenetek – hasonlóan a közismert chatelő szoftverek (pl. Skype, Windows Live Messenger, Gmail chat) működéséhez – a képernyő jobb oldalán, egy chatablakban jelentek meg, beírásukra az ablak alatti szövegdoboz szolgált, elküldésük az Enter billentyű leütésével volt lehetséges.

A csoporttagok a csoportok alakulásakor a Tanuló 1, Tanuló 2 stb. nevet kapták, ez a név jelent meg a problémamegoldás során végig, ha egy csoporttag üzenetet írt be. Olyan elnevezést kívántunk biztosítani a vizsgálati személyeknek, amelyek várhatóan nem generálnak semmiféle érzelmi töltést (akár annak tudatosulása nélkül), hiszen a teszt végén egymás teljesítményéről is véleményt kértünk tőlük. Az volt a célunk, hogy ezt a véleményt a lehető legkevesebb kontrollálhatatlan változó befolyásolja. Ezért, ha a másfél perces ismerkedés alatt nem mutatkoztak be egymásnak, nem derült ki a csoporttagok neve, neme. Annak érdekében, hogy megjegyezhetővé tegyük a csoporttagokat – ennek szintén a véleményadásnál volt jelentősége – a chatablakban Tanuló 1, Tanuló 2 stb. üzenetei adott, kötött színnel (pl. Tanuló 1 barnával, Tanuló 2 zölddel stb.) jelentek meg a problémamegoldó folyamatban.

### **A megoldás megadásának módja**

A csoporttagok mindegyike – a pilot feladatban és a problémamegoldás során is – a bevezetésben ismertetett definícióknak megfelelően ugyanazt a feladtleírást kapta, meg egyező információkészlettel. Miután megismerkedtek az instrukcióval, a chat segítségével



megvitathatták, milyen megoldást adjanak a problémára. A megoldási alternatívák közül a kiválasztottakat a csoporttagok drag & drop („fogd és vidd”) funkció segítségével húzhatták a Megoldás ablakba. Ezen ablak mellett szereplő Megosztás gombra kattintva a chatablakban megjelent egy üzenet arról, hogy egy adott csoporttag megosztott egy megoldást (pl. „Tanuló 1 megosztotta ezt a megoldást: Bécs”).

Csak akkor léphettek tovább saját felületükön a csoporttagok, ha egy probléma kapcsán mindegyikük megosztotta ugyanazt a megoldást, a Tovább gomb ebben az esetben aktiválódott az adott oldalon. A konszenzus mint feltétel indokoltságát az információkészlet exponálásának módjában láttuk. Mivel a vizsgálati személyek szimultán szembesültek a teljes, a probléma megoldásához szükséges információkészlettel, konszenzus igényének hiányában könnyedén előfordulhatott volna, hogy nem bontakozik ki csoportmunka, azaz a csoporttagok egyénileg hozzák és adják meg döntéseiket anélkül, hogy azt közösen megvitatnák. A Tovább gombra lépve a csoporttagok egy újabb oldalon találták magukat, a már továbblépők kilépéséről a még az előző oldalon maradt csoporttagok szintén üzenetet kaptak a chatablakban (pl. Tanuló 2 kilépett a beszélgetésből).

### A problémák

A vizsgálati személyeknek négy analitikus (statikus), tartalomfüggetlen, jól definiált problémát kellett csoportosan megoldaniuk. Az első egy pontot, a második kettő, a harmadik és a negyedik pedig négy pontot érő probléma volt, összesen 11 egy pontot érő ítemet tartalmaztak.

Az analitikus probléma olyan problémátípust jelöl, amelyben a változókészlet a probléma megoldása során – a dinamikus problémákkal szemben – nem változik. Példa erre a 2003-as PISA-mérés problémamegoldás minorterületén alkalmazott feladatsor (Molnár, 2006b). Az volt a célunk, hogy a számos kontrollálhatatlan változót minimálisra csökkentsük. Úgy ítéltük meg, hogy erre a célra alkalmasabbak az analitikus feladatok, illetve hiánypótlónak is találtuk ezt a választást, hiszen az eddigi és a tervezett kutatások dinamikus problémákat használnak. Jelen kutatás problémái a 2003-as PISA-feladatok döntéshozatali problémáihoz hasonlíthatóak leginkább. A vizsgálati személyeknek értelmezniük kellett a problémahelyzetet, össze kellett vetniük a lehetséges kimeneteleket a korlátozó tényezőkkel, majd mindezek alapján a megoldási alternatívák közül kiválasztani a lehető legmegfelelőbbet (1. ábra). A korlátozó tényezők számának növelésével és az azok egyre nehezebben kiszűrhető, rejtett közlésével mi is folyamatosan fokoztuk a problémák komplexitását.

A rosszul definiált problémákkal szemben a jól definiált jelleg arra utal, hogy a probléma célállapota egyértelmű, az elérési útvonal az ismeretlen (Molnár, 2001, 2006a; Pásztor-Kovács, 2015). Számítógép alapú problémamegoldás vizsgálatokban egyelőre az ilyen típusú problémák alkalmazása a kézenfekvő. A rosszul definiált problémák megoldásának nyomon követése, értékelése – akár egyénileg, akár csoportban – lényegesen nagyobb kihívást jelentene.


Ágit szülei két ajándékkal szeretnék meglepni a születésnapján. Hogy biztosan jól válasszanak, számba veszik, hogy szabadidejében mivel foglalkozik, és mi az, amivel nem szívesen tölti idejét. Összeírnak egy rövid listát is:


**Szeret:**


- ✓ kirándulni
- ✓ moziba járni
- ✓ táncolni
- ✓ énekelni
- ✓ jókat nevetni
- ✓ új dolgokat kipróbálni
- ✓ a barátaival lenni
- ✓ szereti az állatokat


**Nem szeret:**


- ✗ olvasni
- ✗ énekelni
- ✗ nem szereti az erőszakot
- ✗ fél a kígyóktól
- ✗ nem szereti a komolyzenét
- ✗ tériszonya van


 mozijegy kedvenc színésze legújabb, 120 perces, feliratos fimjének hazai premierjére


 belépő egy karaoke estre a barátaival


 DVD a kedvenc vígjátéksorozatáról

 belépő a közeli tv-torony panorámateraszára

 CD lemez Vivaldi zenéjével

 belépő a helyi tinidisco-ba

 karate bérlet a helyi fitness terembe

 belépő az országos vándor kisállat- és hullókiállításra

Válasszátok ki, és húzzátok a megoldókeretekbe annak a két ajándékötletnek a képét, amelyekről azt gondoljátok, hogy Ági **biztosan örülne** nekik!

Megoldás 1

Megoldás 2

Megosztás

Tovább

**EDIA Chat Ablak**

Tanuló - 1: a kiállítás! annak biztos örülne

Tanuló - 2: A Vivaldi CD-t elvethetjük, mert nem szereti a komolyzenét

Tanuló - 3: de fél a kígyóktól... ezek meg hülyék

Tanuló - 4: ok, a kiállítás jó ötlet!

Tanuló - 4: tényleg... :)

Tanuló - 3: igen, Vivaldi nem jó

Tanuló - 1: és a mozijegy?

Tanuló - 2: mit szólnátok a DVD-hez, szeret moziba járni,

Tanuló - 2: Oh, valóban, van mozijegy is, igazad van

Tanuló - 3: de várjatok, feliratos a film, és nem szeret olvasni

Tanuló - 1: megosztotta ezt a megoldást: tv-torony belépő

Tanuló - 4: Várj még, a tv-torony nem jó, mert tériszonya van!

1. ábra

*Korlátozó tényezők mérlegelése az általunk használt eszköz második problémájának megoldása közben*

A tartalomfüggetlen, azaz nem tantárgyakhoz köthető tudást igénylő problémákra abból a megfontolásból esett a választásunk, hogy az előzetes tudás eltérő mennyiségét mint lehetséges csoportmunkát befolyásoló tényezőt kiküszöböljük. Olyan problémahelyzeteket vázoltunk fel, amelyek későbbi, felső tagozatos célcsoportunk hétköznapijait reálisan jellemezhetik (pl. születésnapi ajándék vagy nyári tábor választása, iskolaválasztás adott paraméterek megfontolásával).

Ahogy a való életben is ritka a vegytiszta tartalomfüggetlen probléma, számunkra is nehéz volt a tantárgyi tartalmak teljes kiiktatása, így két probléma is igényelt számolási készséget. A harmadik probléma megoldásához szükséges volt néhány fejben elvégezhető alpművelet, a negyedik probléma kivitelezéséhez átlagot is kellett számolni, ehhez egy tesztbe épített számológép állt a résztvevők rendelkezésére (2. ábra).

The screenshot shows a web application interface for a collaborative problem-solving task. The main content area displays the profiles of four students: Csabi, Ildi, Bea, and Norbi. Each profile includes a name, a brief description, and a dropdown menu for selecting a subject. The dropdown menu for Norbi is open, showing a list of subjects: Matematika, Fizika, Kémia, Biológia, Földrajz, Rajz, Technika, Informatika, Ének-zene, and Testnevelés. To the right of the main content area is a chat window titled 'EDIA Chat Ablak' showing messages from students. At the bottom of the interface is a calculator and a row of buttons for different subjects: Kossuth, Petőfi, Blaha, II. Rákóczi, Balassi, József Attila, and Karinthy. There are also buttons for 'Megosztás' (Share) and 'Tovább' (Next).

2. ábra

*Számológép használata a negyedik probléma megoldása közben*

## A kérdőív

A kérdőív célja kettős volt. Egyrészt a vizsgálati személyek általános attitűdjeit vizsgálta a tesztel kapcsolatosan. Ebből a célból három – ötfokú Likert-skálán értékelt – kérdést tettünk fel a résztvevőknek, sorrendben: Hogy érezted magad a problémák közös megoldása közben? Mennyire találtad nehéznek a problémákat? Mennyire érzed sikeresnek az együttműködéseket? Felkínáltuk továbbá a lehetőséget bármely egyéb, a tesztel (a technológiával, a problémákkal, a csoporttársakkal) kapcsolatos visszajelzés megadására. A vizsgálati személyek egy erre a célra kialakított szövegdobozban szabadon kifejezhették véleményüket.

A kérdőív további része validálási célokat szolgált. A vizsgálati személyeket csoporttársaik viselkedéséről kérdeztük. Olyan állítások igazságáról kellett döntést hozniuk, amelyek jól leírták az általunk felvázolt képességmodell elemeit (I. Pásztor-Kovács, 2015). Eredeti kérdőívünkben a kitöltőknek 20 állítással kapcsolatban minden csoporttársukra nézve meg kellett határozniuk egyetértésük mértékét egy 1-től 5-ig terjedő skálán (pl. Kimondottan aktív volt a problémamegoldás során, sokat hozzászólt a beszélgetésünkhöz. Volt, hogy várni kellett rá, mert nem követte az eseményeket. A legtöbb ötletéről kiderült,

hogy nem működőképes). Ám előzetes méréseink során azt a visszajelzést kaptuk, hogy a kérdőív ebben a formában túllépi a tolerálható hosszúságot, illetve a kitöltők nem emlékeznek olyan pontosan csoporttársaik viselkedésére, hogy arról érdemben ilyen részletes-ségű értékelést tudjanak adni. Ezért a kérdőív jelen változatában megtartottuk a 20 állítást, azonban a résztvevőknek csak arról kellett döntenüik, hogy melyik társukra (többet is bejelölhettek) volt jellemző az adott viselkedés.

Az adatfelvétel után ezúttal is arról számoltak be szóban a vizsgálati személyek, hogy nem emlékeztek egymás viselkedésére, illetve nem tudták azt a Tanuló 1, Tanuló 2 stb. elnevezésekhez kötni, ezért gyakorlatilag megfontolás nélkül, random módon töltötték ki a kérdőívet. A chatablakban adott csoporttagokhoz tartozó üzenetek kötött színe sokaknak – saját bevallásuk szerint – fel sem tűnt, ez a megoldás tehát nem jelent elegendő kognitív munkát a társak megjegyzéséhez. Mindezek alapján a kérdőív második feléből nyert eredményeket nem elemeztük.

### **A visszajelzés**

A teszt végén a köszönetnyilvánítás mellett visszacsatolást nyújtottunk a vizsgálati személyeknek egy százalékos érték formájában, amit a csoport problémamegoldás során szerezhető maximális pontszámához képest számolt ki a rendszer. Mivel a csoporttagok minden esetben ugyanazt a választ adták, minden csoporttag esetében megegyező volt ez az érték.

Kérdéses volt, hogy a teszt mely pontján adjuk a visszajelzést, közvetlen a problémamegoldást követően, a kérdőív előtt vagy a teszt végén, a kérdőív után. Tartottunk attól, hogy a visszajelzést azonnal megkapva a vizsgálati személyek elvesztik érdeklődésüket a teszt iránt, és a kérdőívet már csupán felületesen, érdektelenül töltik ki. Ugyanakkor fennállt az a veszély is, hogy a végeredmény iránti kíváncsiságuk okán, ha azt a teszt végén helyezzük el, szintén gyorsan, megfontolás nélkül töltik ki a kérdőívet. Végül utóbbit ítéltük meg a kisebb kockázattal bírónak, azaz a visszajelzés a teszt végére került. A problémák megoldása után előrevetítettük, hogy a kérdőívet követően az eredményről is kapnak információt.

## **Eredmények**

### **A problémamegoldás eredményei**

Az elemzéshez 69 vizsgálati személy adatait használtuk fel, két személy eredményeit figyelmen kívül kellett hagynunk, mert valamilyen hiba folytán ugyanazzal a mérési azonosítóval dolgoztak, és bizonytalan volt, mikor melyikük választ rögzítette a rendszer.

A Cronbach- $\alpha$  érték a 11 itemre nézve 0,78. A négy problémán nyújtott teljesítmény együttjárását az 1. táblázat szemlélteti.

1. táblázat. Szignifikáns korrelációs kapcsolatok a teszt problémái között

Probléma sorszáma	1.	2.	3.
2.	0,68*	–	–
3.	0,41*	0,53*	–
4.	0,29**	0,27**	n. s.

Megjegyzés: \*  $p < 0,01$ . \*\*  $p < 0,05$ . n.s.=nem szignifikáns

Az összesen 11 pontos teszten a vizsgálati személyek átlagosan 8,91 pontot értek el ( $SD=2,50$ ), a négy probléma megoldására fordított idő átlaga kb. 30 perc ( $M_{\min}=29,99$ ;  $SD=5,96$ ) volt. Szignifikáns pozitív korrelációt találtunk az összpontszám és a probléma-megoldó idő hossza között ( $r=0,529$ ;  $p < 0,01$ ).

#### A kérdőív kitöltéséből nyert eredmények

Az eredmények alapján a vizsgálatban résztvevők több mint fele jól érezte magát a problémák közös megoldása közben: 44,8%-uk nyilatkozott úgy, hogy élvezte, 14,9%-uk úgy, hogy nagyon élvezte a tesztet, 32,8%-uk volt, amit élvezett a tesztben, volt, amit nem, 6%-uk nem, illetve 1,5%-uk egyáltalán nem élvezte a csoportos problémamegoldást (2. táblázat). A résztvevők túlnyomó többsége közepesen nehéznek találta a problémákat: 70,1%-uk szerint a problémák között volt könnyű és nehéz is, a minta negyede, 25,4%-a könnyűnek minősítette őket, 1,5%-a nagyon könnyűeknek, 3%-a ellenben nehezeknek. A résztvevők több mint 85%-a eredményesnek élte meg a csoportjában zajló munkát: 70,2%-uk sikeresnek, 16,4%-uk nagyon sikeresnek értékelte az együttműködést, 10,4%-uk valamennyire sikeresnek találta azt, mindössze 3%-uk ítélte együttműködésüket sikertelennek.

A problémamegoldó teljesítményt leíró összpontszám nem mutatott szignifikáns kapcsolatot sem a teszt általános élvezeti értékét, sem a problémák nehézségét, sem a csoport sikerének szubjektív megítélését vizsgáló kérdésből nyert változóval. Szignifikáns pozitív együttjárást ( $r=0,50$ ;  $p < 0,01$ ) tapasztaltunk a Hogy érezted magad a problémák közös megoldása közben? és a Mennyire érzed sikeresnek az együttműködéseket? kérdésekre adott válaszok értékei között.

A résztvevők közül 37 használta ki a lehetőségét annak, hogy véleményét szabadon is megfogalmazza a tesztrel kapcsolatban. A válaszokat tartalmuk alapján hat kategóriába rendeztük, a 37-ből 11 választ kettő vagy három kategóriához is hozzárendeltünk. A legnagyobb kategória a tesztrel kapcsolatos általános attitűdök kategóriája. 16 választ soroltunk ide, ebből 15 pozitív attitűdöt kommunikált (pl. „Szerintem nagyon jó volt!”, „Én élveztem a tesztet, gyakrabban kellene ilyen kitöltetni a tanulókkal.”, „...alapvetően élvezetes teszt, tetszett, újszerű.”, „A programot jó ötletnek találtam és modern eszközökkel próbálta a feladatok megoldását elősegíteni.”), egy negatívát („Nem szívesen dolgoznék így.”).

2. táblázat. A tesztre irányuló kérdésekre adott válaszok gyakorisága

<i>Kérdés</i>	<i>Válaszlehetőség</i>	<i>Válaszgyakoriság (%)</i>
1. Hogy érezted magad a problémák közös megoldása közben?	Nagyon élveztem	14,9
	Élveztem	44,8
	Volt, amit élveztem benne, volt, amit nem	32,8
	Nem élveztem	6,0
	Egyáltalán nem élveztem	1,5
2. Mennyire találtad nehéznek a problémákat?	Nagyon könnyűek voltak	1,5
	Könnyűek voltak	25,4
	Volt nehéz is, könnyű is	70,1
	Nehezek voltak	3,0
	Nagyon nehezek voltak	0
3. Mennyire érzed sikeresnek az együttműködéseket?	Nagyon sikeres volt az együttműködésünk	16,4
	Sikeres volt az együttműködésünk	70,2
	Valamennyire sikeres volt az együttműködésünk	10,4
	Nem volt sikeres az együttműködésünk	3,0
	Egyáltalán nem volt sikeres az együttműködésünk	0

A második legnagyobb tartalmi kategóriához az a 12 válasz tartozik, amely a csoport munkájára vagy a csoportban betöltött saját szerepre reflektál számos különböző szempontból (pl. „Amikor valamelyikünk kihagyott valamit, a másik figyelt... jó csapat volt!”, „Sikerült megoldani, tehát sikeres volt, de nem volt semmi csapatmunka benne. Aki hamarabb rájött, az próbálta néha már kissé erőszakosan is kiharcolni a többiektől, hogy haladjunk.”, „Érdekes, valakinek fel kell lépni vezetőként, és meg kell győznie a többit, vagy irányítania kell őket. A bizonytalan csapatemberekkel nehéz bánni, hiszen nincs személyes, verbális kommunikáció. Izgalmas lehet olyan emberek esetében, akik ismerik egymást”).

A chatre mint kommunikációs eszközre vonatkozott tíz válasz, közülük hat jónak vagy érdekesnek, négy valamilyen szempontból zavarónak vagy kevésbé hatékonynak minősítette azt (pl. „Nagyon érdekes volt chat-en kommunikálni a csoporttársakkal.”, „A virtuális kommunikáció számomra kevésbé hatékony, nehezebben tudom kifejezni magam”). Technikai jellegű problémákat, tanácsokat tartalmazott nyolc válasz (pl. „Egy csoporttársat technikai okokból elvesztettünk: a rendszer kiléptette valamiért.”, „...azonban a megoldásoknál a megosztást nagyon nehéz volt lekövetni és ellenőrizni”).

Öt megjegyzés a teszt problémáira referált, érdekesnek vagy élvezetesnek, illetve ötből négy esetben összetettnek minősítette őket (pl. „Érdekesek voltak a feladatok nagyon, figyelmesnek kellett lenni, mert sokszor becsapósak voltak, és minden részletre ügyelni kell!”, „Tetszett, hogy nagyon megdolgoztatta az agyat, a feladatok nagyon komplexek...”). Az egyéb kategóriában szereplő egyetlen válasz nem kapcsolódott szorosan a teszthez.

### **További, a csoportok interakcióiból nyert eredmények**

Az interakciókkal kapcsolatos elemzéseink már nem egyén-, hanem csoportszintűek. Az elemzések során egy interakciós egységként az Enter billentyű lenyomásával elküldött üzeneteket tekintettük, illetve a Megosztás gomb lenyomásával adott megoldási javaslatokat, amelyek szintén megjelentek üzenetként a chatablakban.

Nincs szignifikáns összefüggés a csoportok teszten elért összpontszáma, illetve aközött, hogy átlagosan hány üzenetet váltottak egymással a problémák megoldása során. Megvizsgáltuk, hogy az interakciók milyen mértékben szorítkoznak pusztán a megoldások közlésére anélkül, hogy arról tényleges eszmecsere, vita alakulna ki. A következő elemzést végeztük: az egy probléma megoldásához tartozó teljes interakciószámot összevetettük a megosztások számával is. Egy 0-nál nagyobb, 1-nél kisebb arányszámot nyertünk mind a négy probléma esetében, minél közelebb van az érték az 1-hez, annál kevesebb interakciót folytattak a csoportok a megosztáson kívül. Az így nyert értékeket átlagolva a megosztások és az összes interakció aránya az első probléma esetében átlagosan 0,42 (szórás=0,29), a másodiknál 0,3 (szórás=0,14), a harmadiknál 0,17 (szórás=0,12), a negyediknél 0,18 (szórás=0,15).

A problémák megoldását megelőző 90 másodperces ismerkedésre szánt interakciókat is megvizsgáltuk arra fókuszálva, hogy milyen mértékben használták ki a csoportok a rendelkezésükre álló időt. Minden csoport esetében regisztráltuk azokat a tíz másodpercnél hosszabb intervallumokat, amelyek üzenetváltás nélkül teltek el. A 17 csoportból 15-nél találtunk ilyen szünetet, az esetek többségében egynél többet is (átlag<sub>szünetek száma</sub>=1,67, szórás=0,72), a szünetek együttes hosszának átlaga 33,87 másodperc (szórás=18,45). Szignifikáns negatív korrelációt találtunk az interakciós szünetek hossza és a csoportok mérete között ( $r=-0,64$ ;  $p=0,01$ ).

## **Összegzés**

Tanulmányunkban ismertettük a kollaboratív problémamegoldó képesség mérésének módszertani kihívásait, valamint az eddigi kihívásra adott válaszokat, vizsgálati trendeket (Griffin & Care, 2015; Hsieh & O'Neil, 2002; Rosen & Foltz, 2014). Bemutattuk saját kutatásunk kivitelezéséhez szükséges online mérőeljárásunkat az első próbamérések adatainak tükrében. Eredményeink arra utalnak, hogy mérőeszközünk, annak elrendezése, az újszerű online kommunikációs feltétel elfogadható és élvezhető tesztkörnyezetet teremt felhasználóinak, illetve problémáink megfelelő szintű megbízhatósággal vizsgálják a konstruktumot. Pilot kutatásunk ugyanakkor számos módosításra érdemes pont felismeréséhez is hozzásegített. A módosítások leginkább adekvát kivitelezése további kutatásokat, számos kérdés alapos mérlegelését igényli a jövőben: döntést kell hoznunk többek között a problémák összetételéről, prezentálásának körülményeiről, a mozaik elrendezés szükségességéről.

Az alacsony probléma-, illetve résztvevőszámot, valamint a válaszok kis variációját tekintve (hiszen a csoporttagok pontszámai minden esetben megegyeztek, így gyakorlatilag 69 fő 17 egybehangzó eredményével dolgoztunk) a 11 itemre nyert 0,78-as Cronbach- $\alpha$  értéket elfogadhatónak találjuk.

A problémák pontszámai közötti pozitív együttjárások szintén arra engedhetnek következtetni, hogy a problémák ugyanazt a konstruktumot vizsgálják. A negyedik probléma viselkedik a legkevésbé konzisztensen a három másikkal, az első kettővel mutat ugyan egy enyhe pozitív korrelációt, a harmadik problémával azonban nem mutatható ki szignifikáns együttjárása. Ennek több oka is lehet. Elképzelhető, hogy a teszt végére csökkent a vizsgálati személyek motivációja, nem gondolkodtak már érdemben a problémákon. Az is feltételezhető, hogy a valamelyest eltérő szerkezet okozta a teljesítménybeli változást. Ebben a feladatban ugyanis felajánlottuk egy beépíthető számológép használatának lehetőségét. Ez önmagában eltérést okozhatott a csoport eddigi folyamataiban, hiszen ezúttal kézenfekvővé vált, hogy feladatokat, szerepeket osszanak ki egymás között a csoporttagok, például a számológépet kezelő szerepét. A három másik problémához messze nem asszociálódott ilyen mértékben a munkamegosztás célszerűsége. További tényezőként húzódhat meg a negyedik probléma eltérő viselkedése mögött annak némileg hangsúlyosabb matematikai tartalma, amit már a számológép mint opcionális segédeszköz használata is előrevetít. Azt feltételezzük, illetve a csoporttagok interakcióit elemezve explicit módon is találkoztunk azzal az attitűddel, hogy a matematikai tartalom negatívan, szorongáskeltően hatott a vizsgálati személyekre (pl. „jajj, átlagot kell számolni”, „hurrá, matek”, „ajjj, számolás”).

Többen már a harmadik probléma kapcsán, amely a negyedik problémánál lényegesen kevesebb és könnyebb műveletet, minimális fejszámolást igényelt, azonnal kijelentették, hogy ők inkább a háttérbe vonulnak, mert inkompetensek, nem matematika szakosok stb. (pl. „itt akkor Jani a matekos, hallgatunk :D”, „légyszi, Dóri segíts majd a matekos részeknél”, „ááá, éljen a BTK”). Ez a viselkedés tehát ismételten befolyással bírhatott a vizsgálati személyek motivációjára. Mindezek fényében a negyedik probléma használatát szükséges átgondolni, további felhasználása csak jelentős módosítások után javasolt.

Az összpontszám meglehetősen magas átlagértéke arra utal, hogy egyetemisták számára a teszt könnyedén teljesíthetőnek bizonyult. Ezt az eredményt azonban erősen mintafüggőként szükséges kezelnünk. Terveink között szerepel fiatalabb korcsoportok vizsgálata is, az, hogy az alapszakos hallgatói minta sem produkált plafonhatást a teszten, bizakodásra ad okot a teszt kihívásértékét, differenciáló erejét illetően. A meglehetősen erős pozitív korrelációs kapcsolat, amelyet a problémamegoldással eltöltött idő és az összpontszám között találtunk, arra utal, hogy azok a csoportok, akik nagyobb mértékben involválódtak a problémamegoldásba, jobban is teljesítettek.

Az attitűdvizsgálat eredményei alapján a teszt fogadtatása alapvetően pozitív. A vizsgálati személyek nagyon alacsony (7,5) százaléka nyilatkozott úgy, hogy nem vagy egyáltalán nem élvezte azt, a többiek legalább részben jól érezték magukat a teszt megoldása során. Összhangban a problémák összetettségére referáló kvalitatív válaszokkal, a Likert-skálás kérdésre adott válaszok alapján is úgy tűnik, hogy a teszt már a mintát alkotó fiatal felnőttek számára is rendelkezett némi kihívásértékkel.



A csoport sikerének szubjektív megítélése – elvárásunkkal ellentétben – teljes mértékben független a tényleges teljesítménytől (bár utóbbiról a kérdőív kitöltése után kaptak csak visszajelzést, ezen a ponton még nem volt róla információjuk). Feltételezhető, hogy az eredmény ismeretében másként válaszoltak volna, ugyanakkor az is lehetséges, hogy a résztvevők számára a konkrét output sokkal kevésbé volt lényeges a sikeresség odaítélésénél megfontolt szempontok között, mint maga a folyamat, ami a megoldás adásához vezetett. Ezen kérdések kutatásra érdemesek lehetnek későbbi vizsgálataink során. A teszt élvezetének és a siker megítélésének pozitív együttjárása kevésbé váratlan eredmény, ha a teszt élvezetét a vélt siker következményeként értelmezzük.

A kvalitatív válaszok szintén a teszt alapvetően pozitív megítélésére engednek következtetni. Azt, hogy a 69 vizsgálati személy közül 59 nem tartotta kommentálásra érdemesnek a chat funkciót, önmagában jelentheti azt, hogy használata nem jelentett gondot, illetve a válaszadók több mint fele is pozitív véleményének adott hangot az újdonsággal kapcsolatban.

A csoportmunkával kapcsolatos számos értékes megjegyzés közül gyakorlati szempontból számunkra az a legfontosabb, ami az időkorlát mint csoportmunkát potenciálisan befolyásoló tényező szerepére utalt. Valóban releváns kutatási kérdésre adna választ a továbbiakban egy olyan kísérleti elrendezés, ahol lehetséges összehasonlítani a csoportmunka jellegét időkorlát alkalmazásával és nélküle.

A technológiai jellegű észrevételek rendkívül hasznosnak bizonyultak, több ponton is szükséges a beavatkozás. Indokoltnak érezzük például a megjegyzések nyomán egy olyan ablak kialakítását a tesztfelületen, ahol a megosztások az üzenetektől izoláltan nyomon követhetők, leegyszerűsítve ezzel a közös megoldás megadásához szükséges visszakeresést a szövegben. Az interakciók elemzéseiből nyert eredmények arra utalnak, hogy nem teljesítettek jobban azok a csoportok, akik több üzenetet váltottak egymással, azaz ténylegesen megvitatták a potenciális megoldási lehetőségeket. Ez a tendencia némileg ellentmondásos annak tükrében, hogy a problémamegoldással eltöltött időtartam hossza erős pozitív összefüggést mutat az összpontszámmal. Mindebből arra következtetünk, hogy a problémák megoldásával eltöltött időt a csoporttagok nem a problémák megvitatására, hanem az önálló átgondolására fordították. Alátámasztja ezt a feltételezést az az eredmény is, hogy az első probléma esetén a csoporton belüli interakciók átlagosan 42%-a, a második probléma esetén 30%-a kizárólag a megoldási javaslatok megosztásából áll. Ilyen méretű csoportok esetében, ahol akár négy ellentétes vélemény is ütközhet, ez az arány nagyon magas. Az összetettebb harmadik és negyedik probléma kapcsán már több interakciót tapasztaltunk, a megosztások száma ezeknél a problémáknál már csak az interakciók 17, illetve 18%-át jelenti.

Két tanulsággal is szolgálnak ezek az eredmények. Egyrészt a kollaboráció tényleges kibontakozásához megfelelő komplexitású problémák prezentálása szükséges. Az, hogy mi számít megfelelő komplexitású problémának, mindazonáltal erőteljesen mintafüggő kérdés, elképzelhető, hogy eltérő összetételű minta esetében az első két probléma is több megvitátást igényelne, ezen kérdés ismételten további kutatásokra szorul. A második tanulság a problémaadással kapcsolatos elrendezésre vonatkozik. Bevezetőnkben ismertettük a képesség általunk alkotott definícióját, melynek értelmében a teljes információkészlet egyidejű bemutatása szükséges az összes csoporttagnak a jigsaw vagy mozaik típusú

problémaadással szemben. Bár továbbra is indokoltnak tartjuk az említett elméleti megfontolásból a kutatásban alkalmazott elrendezést, további kutatási céljaink, azaz kommunikációs mintázatok összehasonlításának szempontjából beláthatóan kevésbé szerencsés egy olyan elrendezés, amely nem feltétlen igényel tényleges, kellő elemzési alapot nyújtó interakciót.

Nem váratlan az az eredmény, hogy a nagyobb méretű csoportok jobban kihasználják az ismerkedésre felkínált időt, az azonban igen, hogy a 90 másodperc feleslegesen hosszú időnek bizonyult, 60 másodperc elegendőnek tűnik a bemutatkozásra.

Az interakciók tartalmát figyelembe véve a jövőben a problémák természetét is érdemes módosítani. A bevezetőben ismertettük, hogy jól definiált, ugyanakkor komplex, a korlátozó tényezőket nem explicit módon közlő döntéshozatali igénylő problémákat állítottunk össze, amivel arra törekedtünk, hogy azok a lehető leginkább tartalomfüggetlenek maradjanak. Azonban azt tapasztaltuk, hogy bár számunkra egyértelmű volt, hogy az adott problémáknak melyek az objektív, jól körülírható szempontokat figyelembe véve optimális megoldásai, a vizsgálati személyek sokszor teljesen más szempontok alapján mérlegelték döntéseiket, és gondolatmeneteiket nem minősíthetjük egyértelműen hibásnak.

A résztvevőknél, bár szembesültek az objektív kritériumokkal, gyakorta bekapcsolt egy saját tapasztalataikra, véleményükre alapuló döntéshozó mechanizmus, amelyet egyértelműen a problémák hétköznapi, ismerős kontextusa hívott elő. Néhány példát kiemelve: az első probléma esetében, amikor azt kellett eldönteniük, hogy négy objektív kritérium alapján kinek jár jutalomkönyv a négy jellemzett tanuló közül, megjelent az a vélemény, hogy bár adott tanuló nem teljesíti mind a négy kritériumot, csak kettőt, ösztönzőképpen mindenképpen neki kellene odaítélni a könyvet. Egy másik, szintén nem a megoldást jelentő tanuló kiválasztása mellett olyan érvet hozott fel egy hallgató, hogy miként érezné magát, ha nem ő, az osztály messze legeminebb diákja kapná a jutalmat, hanem valaki más. Jó példa a harmadik problémánál több ízben megjelenő vélemény is, miszerint annak ellenére, hogy az egyik kisfiú, akinek a megfelelő, számára biztosan élményt nyújtó nyári tábort kellett kiválasztani, nem szeret mozogni, olyan táborba menjen, ahol hosszú túrákat kell tennie, elvégre muszáj megszeretnie a testmozgást.

A fenti jelenség, azaz az erőteljesen egyéni tapasztalatra, véleményre alapuló gondolkodás bekapcsolásának elkerülésére megoldás lehetne például olyan természettudományos gondolkodást igénylő problémák alkalmazása, ahol a kivitelezéshez szükséges információkat újfent biztosítanánk, ilyen formán előzetes tudást nem igényelnének. Ez esetben azonban a tartalom kapcsán potenciálisan jelentkező ellenállás okozhat torzítást az adatfelvétel során, azaz a saját – vélt vagy valós – inkompetencia érzése vagy egy általános természettudományos tantárgyakkal szemben mutatott negatív attitűd, ami a vizsgálati személyek jelentős hányadát demotiválhatja.

#### *Köszönetnyilvánítás*

Köszönettel tartozom *Molnár Gyöngyvér* szakmai támogatásáért. A kutatás az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával a TÁMOP 3.1.9-11/1-2012-0001 azonosító jelű „*Diagnosztikus mérések fejlesztése*” című kiemelt projekt keretében valósult meg. A tanulmány megírását támogatta a TÁMOP-4.2.2.B-15/1/KONV-2015-0006 azonosítószámú, „*A tehetség*

értékének kibontakoztatása a Szegedi Tudományegyetem kiválósága érdekében” c. projekt keretében az Európai Unió.

## Irodalom

- Adams, R., Vista, A., Scoular, C., Awwal, N., Griffin, P., & Care, E. (2015). Automatic coding procedures for collaborative problem solving. In P. Griffin & E. Care (Eds.), *Assessment & teaching of 21st century skills: Methods and approach* (pp. 115–132). Dordrecht: Springer. doi: [10.1007/978-94-017-9395-7\\_6](https://doi.org/10.1007/978-94-017-9395-7_6)
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Martin, R., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment & teaching of 21st century skills* (pp. 17–66). New York: Springer. doi: [10.1007/978-94-007-2324-5\\_2](https://doi.org/10.1007/978-94-007-2324-5_2)
- Birdwhistell, R. L. (1970). *Kinesics and context. Essays on body-motion communication*. University of Philadelphia: Pennsylvania Press. doi: [10.9783/9780812201284](https://doi.org/10.9783/9780812201284)
- Care, E., & Griffin, P. (2014). An approach to assessment of collaborative problem solving. *Research & Practice in Technology Enhanced Learning*, 9(3), 367–388. doi: [10.1007/978-94-017-9395-7\\_4](https://doi.org/10.1007/978-94-017-9395-7_4)
- Care, E., Griffin, P., Scoular, C., Awwal, N., & Zoanetti, N. (2015). Collaborative problem solving tasks. In P. Griffin & E. Care (Eds.), *Assessment & teaching of 21st century skills. Methods and approach* (pp. 85–104). Dordrecht: Springer.
- Chung, G. K. W. K., O’Neil, H. F., & Herl, H. E. (1999). The use of computer-based collaborative knowledge mapping to measure team processes and team outcomes. *Computers in Human Behavior*, 15(3), 463–493. doi: [10.1016/s0747-5632\(99\)00032-1](https://doi.org/10.1016/s0747-5632(99)00032-1)
- Csapó, B., Ainley, J., Bennett, R., Latour, T., & Law, N. (2012). Technological issues of computer-based assessment of 21st century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment & teaching of 21st century skills* (pp. 143–230). New York: Springer. doi: [10.1007/978-94-007-2324-5\\_4](https://doi.org/10.1007/978-94-007-2324-5_4)
- Csapó, B., Lőrincz, A., & Molnár, G. (2012). Innovative assessment technologies in educational games designed for young students. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: foundations, innovations, and perspectives* (pp. 235–254). New York: Springer. doi: [10.1007/978-1-4614-3546-4\\_13](https://doi.org/10.1007/978-1-4614-3546-4_13)
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, 33, 1–40. doi: [10.1016/s0065-2601\(01\)80003-4](https://doi.org/10.1016/s0065-2601(01)80003-4)
- Graesser, A. C., Foltz, P., Rosen, Y., Forsyth, C., & Germany, M. (in press). Challenges of assessing collaborative problem solving. In B. Csapó, J. Funke, & A. Schleicher (Eds.), *On the nature of problem solving*. Paris: OECD.
- Griffin, P., McGaw, B., & Care, E. (Eds.). (2012). *Assessment & teaching of 21st century skills*. Dordrecht: Springer. doi: [10.1007/978-94-007-2324-5](https://doi.org/10.1007/978-94-007-2324-5)
- Griffin P., Care, E., Bui, M., & Zoanetti, N. (2013). Development of the assessment design and delivery of collaborative problem solving in the assessment and teaching of 1st century skills project. In E. McKay (Ed.), *Epedagogy in online learning: New developments in web mediated human computer interaction* (pp. 55–73). Hershey: IGI Global. doi: [10.4018/978-1-4666-3649-1.ch004](https://doi.org/10.4018/978-1-4666-3649-1.ch004)
- Griffin, P., & Care, E. (2015). *Assessment & teaching of 21st century skills. Methods and approach*. Dordrecht: Springer. doi: [10.1007/978-94-017-9395-7](https://doi.org/10.1007/978-94-017-9395-7)
- Griffin, P., Care, E., & Harding, S-M. (2015). Task characteristics and calibration. In P. Griffin & E. Care (Eds.), *Assessment & teaching of 21st century skills. Methods and approach* (pp. 133–181). Dordrecht: Springer. doi: [10.1007/978-94-017-9395-7\\_7](https://doi.org/10.1007/978-94-017-9395-7_7)

- Greiff, S. (2012). From interactive to collaborative problem solving: Current issues in the Programme for International Student Assessment. *Review of Psychology*, 19(2), 111–121.
- Greiff, S., Holt, D. V., & Funke, J. (2013). Perspectives on problem solving in educational assessment: Analytical, interactive, and collaborative problem solving. *Journal of Problem Solving*, 5(2) 71–91. doi: [10.7771/1932-6246.1153](https://doi.org/10.7771/1932-6246.1153)
- Hermann, F., Rummel, N., & Spada H. (2001). Solving the case together: The challenge of net-based interdisciplinary collaboration. Retrieved from [http://pdf.aminer.org/000/248/605/a\\_comparison\\_of\\_collaborative\\_problem\\_solving\\_using\\_face\\_to\\_face.pdf](http://pdf.aminer.org/000/248/605/a_comparison_of_collaborative_problem_solving_using_face_to_face.pdf)
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment & teaching of 21st century skills. Methods and approach* (pp. 37–56). Dordrecht: Springer. doi: [10.1007/978-94-017-9395-7\\_2](https://doi.org/10.1007/978-94-017-9395-7_2)
- Hsieh, I. L., & O'Neil, H. F. Jr. (2002). Types of feedback in a computer-based collaborative problem solving group task. *Computers in Human Behavior*, 18(1), 699–715. doi: [10.1016/s0747-5632\(02\)00025-0](https://doi.org/10.1016/s0747-5632(02)00025-0)
- Keyton, J. (2002). *Communicating in groups: Building relationships for effective decision-making*. New York: McGraw-Hill.
- Kozma, R. (2009). Assessing and teaching 21st century skills: A call to action. In F. Schueremann, F. & J. Bjornsson (Eds.), *The transition to computer-based assessment: New approaches to skills assessment and implications for large scale assessment* (pp. 13–23). Brussels: European Communities.
- Krkovic, K., Pásztor-Kovács, A., Molnár, G., & Greiff, S. (2014). New technologies in psychological assessment: The example of computer-based collaborative problem solving assessment. *International Journal of e-Assessment*, 1(1), online.
- Levine, J. M., & Moreland, R. L. (1998). Small groups. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 415–469). Boston: McGraw-Hill.
- Molnár, G. (2001). Az életszerű feladathelyzetekben történő problémamegoldás vizsgálata. *Magyar Pedagógia*, 101(3), 347–373.
- Molnár, G. (2006a). *Tudástranszfer és komplex problémamegoldás*. Budapest: Műszaki Kiadó.
- Molnár, G. (2006b). A tudáskonceptió változása és annak megjelenése a PISA 2003 vizsgálat komplex problémamegoldás moduljában. *Új Pedagógiai Szemle*, 56(1), 75–86.
- Molnár, G. (2010). Technológia-alapú mérés-értékelés hazai és nemzetközi implementációi. *Iskolakultúra*, 20(7–8), 22–34.
- Molnár, G. (2011). Az információs-kommunikációs technológiák hatása a tanulásra és oktatásra. *Magyar Tudomány*, 172(9), 1038–1047.
- Molnár, G., & Csapó, B. (2013, April). *Az eDia online diagnosztikus mérési rendszer*. Paper presented at the XI. Pedagógiai Értékelési Konferencia, Szeged.
- Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: relations and development. *Thinking skills and creativity*, 9, 35–45. doi: [10.1016/j.tsc.2013.03.002](https://doi.org/10.1016/j.tsc.2013.03.002)
- Molnár, G., Greiff, S., Wüstenberg, S., & Fischer, A. (in press). Empirical study of computer-based assessment of domain-general complex problem solving skills. In B. Csapó, J. Funke, & A. Schleicher (Eds.), *On the nature of problem solving*. Paris: OECD.
- Molnár, G., & Pásztor-Kovács, A. (2015). A számítógépes vizsgáztatás infrastrukturális kérdései: az iskolák eszközparkjának helyzete és a változás tendenciái. *Iskolakultúra*, 15(4), 49–61. doi: [10.17543/iskult.2015.4.49](https://doi.org/10.17543/iskult.2015.4.49)
- Moore, F. R., Filippou, D., & Perrett, D. (2011). Intelligence and attractiveness in the face: Beyond the attractiveness halo effect. *Journal of Evolutionary Psychology*, 9(3), 205–217. doi: [10.1556/jep.9.2011.3.2](https://doi.org/10.1556/jep.9.2011.3.2)

- Neubert, J., Mainert, J., Kretschmar, A., & Greiff, S. (2015). The assessment of 21st century skills in industrial and organizational psychology: Complex and collaborative problem solving. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 8(2), 238–268. doi: [10.1017/iop.2015.14](https://doi.org/10.1017/iop.2015.14)
- OECD (2010). *PISA 2012 problem solving framework*. Paris: OECD.
- OECD (2013). PISA 2015 draft collaborative problem solving assessment framework. Retrieved from <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>
- O’Neil, H. F., Chuang, S., & Chung, G. K. W. K. (2003). Issues in the computer- based assessment of collaborative problem solving. *Assessment in Education*, 10(3), 361–373. doi: [10.1080/0969594032000148190](https://doi.org/10.1080/0969594032000148190)
- O’Neil, H. F., Chung, G. K. W. K., & Brown, R. S. (1997). Use of networked simulations as a context to measure team competencies. In H. F. O’Neil (Ed.), *Workforce readiness: Competencies and assessment* (pp. 411–452). Mahwah: Lawrence Erlbaum.
- Pásztor-Kovács, A. (2013, August). *Methodological challenges in the assessment of collaborative problem solving*. Paper presented at the 15th European Conference for the Research on Learning and Instruction, Munich, Germany.
- Pásztor-Kovács, A. (2015). Kollaboratív problémamegoldó képesség: egy új, integratív elméleti keret. *Iskolakultúra*, 15(2), 3–16. doi: [10.17543/iskult.2015.2.3](https://doi.org/10.17543/iskult.2015.2.3)
- Ras, E., Krkovic, K., Greiff, S., Tobias, E., & Maquil, V. (2014). Moving towards the assessment of collaborative problem solving skills with a tangible user interface. *TOJET: The Turkish Online Journal of Educational Technology*, 13(4) 95–104.
- Rosen, Y. (2014a). Assessing collaborative problem solving through computer agent technologies. In M. Khosrow-Pour (Ed.), *Encyclopedia of information science and technology* (pp. 94–102). Hershey: Information Science Reference, IGI Global. doi: [10.4018/978-1-4666-5888-2.ch010](https://doi.org/10.4018/978-1-4666-5888-2.ch010)
- Rosen, Y. (2014b). Comparability of conflict opportunities in human-to-human and human-to-agent online collaborative problem solving. *Technology, Knowledge and Learning*, 19(1–2), 147–164. doi: [10.1007/s10758-014-9229-1](https://doi.org/10.1007/s10758-014-9229-1)
- Rosen, Y., & Foltz, P. (2014). Assessing collaborative problem solving through automated technologies. *Research and Practice in Technology Enhanced Learning*, 9(3), 389–410.
- Rosen, Y., & Tager, M. (2013). *Computer-based assessment of collaborative problem-solving skills: Human-to-agent versus human-to-human approach*. Boston, MA: Pearson Education.
- Scheuermann, F., & Björnsson, J. (Eds). (2009). *The transition to computer-based assessment. New approaches to skills assessment and implications for large-scale testing*. Luxemburg: Office for Official Publications of the European Communities.
- Wüstenberg, S., Greiff, S., Molnar, G., & Funke, J. (2014). Determinants of cross-national gender differences in complex problem solving competency. *Learning and Individual Differences*, 29(1), 18–29. doi: [10.1016/j.lindif.2013.10.006](https://doi.org/10.1016/j.lindif.2013.10.006)
- Zhang, J. (1998). A distributed representation approach to group problem solving. *Journal of the American Society for Information Science*, 49(9), 801–809. doi: [10.1002/\(SICI\)1097-4571\(199807\)49:9<801::AID-ASIS>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-4571(199807)49:9<801::AID-ASIS>3.0.CO;2-Q)

Pásztor-Kovács Anita

## ABSTRACT

### THEORETICAL AND METHODOLOGICAL CONSIDERATIONS IN ASSESSING COLLABORATIVE PROBLEM SOLVING COMPETENCE: RESULTS OF A PILOT STUDY

Anita Pásztor-Kovács

The great relevance of the development of collaborative problem solving (ColPS) competence as a 21st-century skill is widely recognized. However, the creation of a reliable, valid measurement to follow improvement of this competence seems extremely challenging from a number of theoretical and methodological aspects. The aim of this paper is twofold: (1) to provide a review of problematic issues in the measurement of ColPS competence and of assessment trends to date; and (2) to present the first results of a trial of a new ColPS assessment tool. The pilot study whose results we report investigated the psychometric properties of our online test and the attitudes toward its new, unusual testing context. 71 bachelor's students completed our test in 17 three-, four- and five-member groups via the eDia (Electronic Diagnostic Assessment) platform. The test consisted of four analytical, content-general, decision-making problems with 11 items in all. To make their joint decision, group members discussed their ideas via chat. After the problem solving part of the test, the students were asked for their opinions on the test via five-point Likert scale items and a qualitative question. The reliability index of the eleven-item test was Cronbach's  $\alpha=.79$ . We found significant positive correlations between the first and the second ( $r=.66$ ;  $p<0.01$ ), the first and the third ( $r=.41$ ;  $p<0.01$ ), and the second and the third problem ( $r=0.53$ ;  $p<0.01$ ); the fourth problem, however, showed only a weak correlation with the first ( $r=.29$ ;  $p<0.05$ ) and the second problem ( $r=.27$ ;  $p<0.05$ ) but did not correlate with the third one. 92.5% of the participants reported partly enjoying, enjoying or very much enjoying this collaborative effort. 94% of the answers expressing general attitudes toward the test clearly suggested a positive opinion. 60% of the comments referring directly to the chat function found this mode of communication modern, interesting and great, while 40% described it as less effective, disturbing or frustrating. Based on our results, we consider our new, third-generation test reliable and acceptable in general. Nevertheless, our findings also provided some very useful conclusions about the necessity of reconstructing the fourth problem and/or reconsidering some core conditions like the content or the way problems are presented in our assessment tool with regard to our future research aims.

Magyar Pedagógia, 116(1). 51–72. (2016)  
DOI: 10.17670/MPed.2016.1.51

Levelezési cím / Address for correspondence: Pásztor-Kovács Anita, Szegedi Tudományegyetem Neveléstudományi Doktori Iskola, H–6722 Szeged, Petőfi Sándor sgt. 30–34.