

PAPÍR-CERUZA ÉS SZÁMÍTÓGÉPES INDUKTÍVGONDOLKODÁS-TESZTEK RELIABILITÁSÁNAK ÖSSZEHASONLÍTÓ VIZSGÁLATA

Tóth Krisztina

SZTE Neveléstudományi Doktori Iskola

Kulcsszavak: számítógépes tesztelés; reliabilitás

A papír-ceruza (PP) tesztek számítógépes adaptációja során biztosítanunk kell, hogy egy teszt nyomtatott formában és számítógépes (CB) környezetben megoldva közel azonos teszt-átlagokat és szórásokat eredményezzen (Leason, 2006; Kim és Kim, 2013). Ezt követően – a nemzetközi ajánlások alapján (pl.: International Test Commission, 2005) – igazolnunk kell, hogy az adott mérőeszköz mindkét tesztkörnyezetben egyformán megbízhatóan mér.

A tesztpontszámok eltéréseinek vizsgálatára számos hazai tanulmány fókuszált (pl.: Hülber és Molnár, 2013; Tóth és Hódi, 2011), ugyanakkor a CB és PP mérőeszközök reliabilitásának összehasonlítását kevés tanulmány tűzte ki céljává hazai szinten. Ezért megvizsgáltuk, hogy a különböző inductív gondolkodás-tesztek két tesztkörnyezetben ugyanolyan megbízhatóan mérnek-e.

A vizsgálatot 4-6. osztályos tanulók (N=1753) bevonásával végeztük. A 4. osztályos tanulók egy alsó tagozatosok számára (Csapó, 1994) készült, az 5-6. osztályos diákok pedig egy felső tagozatosoknak (Csapó, 2001) szóló inductív gondolkodás-tesztet oldottak meg PP formátumban és online. Az adatfelvétel során kétféle mintavételi eljárást alkalmaztunk. Az első esetben illesztett mintavételi eljárással összekötöttük a PP és CB tesztmegoldók mintáját a nem, a településtípus és a szülők iskolai végzettsége változók segítségével. Továbbá az 5. osztályos mintán az illesztett mintavétel mellett, ismételt méréssel is elvégeztünk összehasonlító vizsgálatot. E mérés keretében a diákok előbb a papír alapú, majd néhány hét eltéréssel a nyomtatott teszt digitalizált változatát oldották meg. A PP adatfelvétel 2005 és 2010 között zajlott, az online pedig 2010-ben a TAO platformon (Latour és Farcot, 2008) keresztül.

Az eredményeket tekintve a PP tesztek Cronbach- α értékei 0,91-0,93, a CB tesztváltozatoké 0,89-0,91 között vannak. Megállapítottuk, hogy mind a számítógépes ($\chi^2=3,90$, $p=0,27$), mind a PP inductív gondolkodás-tesztek ($\chi^2=5,91$, $p=0,12$) egyformán megbízhatóan mérnek a három évfolyamon. Ugyanakkor 5. évfolyamon a PP teszt reliabilitása szignifikánsan magasabb mindkét minta-kialakítási módszer esetén ($\chi^2=5,76$, $p=0,02$; $\chi^2=17,22$, $p=0,00$), mint az online teszté. Ezzel szemben a 4. és 6. osztályos részmintán nincs jelentős különbség a PP és CB teszt Cronbach- α értékei között. A teljes teszt szintű reliabilitásvizsgálatot kiegészítettük részteszt szintű elemzésekkel, melyek eredményei alapján a zárt típusú feladatokat tartalmazó részteszt feladatain nincs jelentős különbség a részteszt szintű reliabilitásmutatók között, ezzel szemben a nyílt végű feladatok esetén jellemzően szignifikáns különbség adódik.

E vizsgálat keretében igazoltuk, hogy a 4. és 6. osztályos mintán a PP és CB tesztek reliabilitásmutatói statisztikailag megegyeznek. Ugyanakkor a részteszt szintű vizsgálatok szignifikáns eltéréseket azonosítottak, melynek oka lehet például az itemtípus, feladatneheztség, a feladat megoldásához szükséges tevékenység. Az okok feltárásához azonban további vizsgálatok szükségesek.