A Redmenta levonásos pontozásáról

I.

1. Németh Ferenc vagyok, matematika-fizika szakos gimnáziumi tanár. Tanulmányomban arra a kérdésre keresem a választ, hogy a Redmenta nevű közismert webhely (http://www.redmenta.com), amely tesztfeladatok létrehozását és oktatási alkalmazását teszi lehetővé tanároknak, vajon **joggal működött-e** úgy egészen a közelmúltig (vagyis 2022 áprilisáig), hogy **a többválasztásos tesztfeladatokban a *hibásan* bejelölt válaszlehetőségeket alapértelmezésben –*1 ponttal* értékelte.**
2. Előzetesen meg kell jegyeznem, hogy a Redmentán a többválasztásos tesztfeladat mint számonkérési eszköz ez időben úgy volt beállítva, hogy a tanuló a kitöltésekor már tudhatta, hogy az adott feladatban *hány helyes válaszlehetőség található*, és a rendszer gépileg nem is engedte, hogy ennél többet bejelöljön. (Kevesebbet természetesen igen.) Ez a beállítás megakadályozza azt, hogy a tanuló minden válaszlehetőséget beikszeljen, és akár érdemi tudás vagy gondolkodás nélkül is maximális pontszámot szerezzen. (Jelenleg a bejelölhető válaszok száma lehet több is, mint a helyes válaszlehetőségek száma.)
3. Az ilyen környezetben gyakorolt levonásos pontozás ellen számos ellenvetés tehető. Mindenekelőtt az, hogy a hibás válaszokat a *kidolgozandó* feladatokban sosem szokták mínuszponttal értékelni, hanem csakis nullával. Az értékelés célja ugyanis deklaráltan az, hogy a tanuló *tudását* értékeljük, és ezzel a céllal ellentétes az az értékelési mód, amely alapján a teszt kitöltője *duplán veszít pontot* a hibásan megjelölt válaszért. E logikát követve a többválasztásos tesztfeladatban a bejelölt jó válaszoknak 1 pontot kell érniük, a bejelölt hibásaknak (és persze a be nem jelölt helyeseknek) 0 pontot.
4. A tanulók reakciója a levonásos pontozásra szinte egyöntetű volt: minden tanárt kórusban kérleltek, hogy *kézzel bírálja felül a Redmentát* – mert erre régen is lehetőséget adott a rendszer. A tanárok egy része bele is ment a sziszifuszi munkába, és minden többválasztásos tesztkérdésnél hozzáadta a rendszer által levont pontokat az eredményhez. Mások ezt mellőzték, és kitartottak a Redmenta alapértelmezése mellett. Aztán 2022 áprilisában a Redmenta fejlesztői – megjegyezve, hogy ez sokan kérték tőlük – *választhatóvá* tették a levonásos pontozást. Ettől kezdve a következő értékelési módszereket kínálják a teszteket készítő tanároknak:
   - Pontozás: Mindig / Csak hibátlan válasz esetén.
   - Pontszám mehet-e nulla alá? Igen / Nem
   - Rossz válaszért vonjon le pontot? Igen / Nem
5. 2022. május 17-én egy Facebook-csoportban (Redmentás tanárok közössége, https://www.facebook.com/groups/redmenta.tanari) feltettem a kérdést, hogy miért volt éveken át alapértelmezettnek beállítva a levonásos pontozás. Azt a választ kaptam a fejlesztőcsapat jelen lévő képviselőjétől, Visy Zoltántól: *„A pontlevonást azért tettük bele a Redmentába, hogy ne bátorítsuk a diákot tippelgetésre. Ne szerencsén múljon a jó érdemjegy.”* (Ezt a szándékot szintén érinteni fogom elemzésemben.)
6. Ugyanezen a helyen találkoztam olyan tanári véleménnyel is, ami *ezt a pontozást kemény szavakkal ítélte el*: „rossz szájízű,” meg „nagyon negatív pedagógiai mondanivalójú büntetésről,” a pontokkal való „negatív ösztönzésről” és a már

„megszolgált" pontok „levonogatásáról" beszélt. Az e véleményt hangoztató kolléga szerint aki így pontoz, az „pedagógiailag fordítva ül a lovon."

7. Mások viszont rámutattak arra, hogy *a mínuszpontozás nem példátlan az oktatásban,* így például bizonyos fajta érettségi tesztekben is előfordul. (Utánanéztem, és meg is találtam: angolérettségin a hallott szöveg értésének tesztes számonkérésénél szokásosan adnak 2-ből 1- vagy 2-választásos igaz-hamis kérdéseket, és ott az alul- vagy túlikszelésre nulla pont jár.) Az általam ismert klasszikus példa a Zrínyi-matematikaverseny, melyen az ötválaszos tesztfeladatban egyetlen helyes válaszlehetőség van, és a jó tipp 4, a rossz tipp –1, a hiányzó tipp pedig 0 pontot ér. Hivatkozhatom még a középiskolai matematikafelvételi „ikszelgetős" típusfeladatára, melyben a hibá(ka)t tartalmazó kitöltésért egységesen 1 pontot le kell vonni az egyébként elért pontszámból.

8. Bár a kérdést sugalmazott válasz nélkül tettem fel, nem titkoltam, hogy *a levonásos pontozás mellett állok*, úgyhogy kezdettől fogva jó lelkiismerettel alkalmaztam is – vagyis nem változtattam a Redmenta alapértelmezett pontozási rendszerének kimenetén. A Facebook-csoportban érdemi vita nem alakult ki (velem szemben valójában csak a fenti, puszta minősítésekkel fellépő vélemény hangzott el), de megígértem a kérdésemre reagáló kollégáknak, hogy idő jártával részletesebb vizsgálattal jelentkezem, melynek az lesz a kimenetele, hogy a levonásos pontozás jogos. Ez időben ugyanis már elkészültem az itt dokumentált első négy Excel-munkalappal, illetve legalább egy éve használtam érveléseimben azok egyszerűsített változatát, a tanulmány végén idézett 4-2-es „ábécéskönyv"-táblázatot.

9. Kiindulásul megfogalmazok néhány **intuitív érvet** amellett, hogy a levonásos módszer nem feltétlenül jogtalan, azaz legitim célokat is szolgálhat. (Hogy tényleg szolgál-e, azt ezek az érvek még nem igazolják, hiszen az ilyen kijelentéseket csak mennyiségi elemzéssel lehet alátámasztani.)

10. A „kettős büntetés" szemrehányása ellen felhozható az, hogy a –1 pontnak igenis van külön jelentése: a (magabiztosan) *bepipált hibás* tipp és a (szerényen) *kihagyott* tipp értékelése közötti *különbségtételt* szolgálja. Eszerint a tanuló egy pontot veszít az elérhetőből amiatt, mert nem tudta a helyes választ, és egy másikat azért, mert bejelölt egy hibásat.

11. A „kétszeres büntetés" elméleti tilalma nem veszi figyelembe azt a *kitöltésbeli segítséget* sem, amit a diák *a helyes válaszok számát* tudva megkap.

12. A Redmentán elvileg át lehet térni a „Mindig" nevű beállításról a „Csak hibátlan válasz esetén" beállításra, melyben a kifogásolt *„levonás, mínuszpont" kifejezések nem szerepelnek.* Ennek dacára ez az értékelés még a kifogásolt levonásos értékelésnél is szigorúbb volna a diák számára (a legtöbb teszttípusnál). Tehát a Redmentában beállítható pontozási rendszerek szűkmarkúsága nem csupán a levonás meglétén múlik.

13. A teszt mint számonkérési mód eleve a diák „szájába adja" a szükséges kifejezéseket, míg ha kiegészítendő volna a kérdés, akkor a helyes válasz végett erőfeszítést kellene tennie (bizonyos szavakat felidéznie vagy szöveget alkotnia), és nulla tudással garantáltan nulla pontot fog kapni. Így viszont elég válogatnia a felkínált lehetőségek között, és a „vaktalálatok" folytán *nulla tudással is esélye van nullánál több pontra*. Tehát a teszt az értékelés módját nem tekintve, lényegéből fakadóan felülpontoz.*

14. **A levonásos módszer *csökkenti* ezt az eredendő *felülpontozást,* ezáltal közelíti a teszt által szolgáltatott eredményt a valós tudáshoz**.

* Megjegyzés: A szakirodalom ezt a problémát régóta tudatosította már. Itt elégnek tartom, ha hivatkozom Pitlik László 2021-ben közzétett, olvasásra igencsak ajánlható szakdolgozatára: *Tudásszintmérő tesztek: a feladattípusok és a tanulói válaszadási stratégiák vizsgálata.* (https://miau.my-x.hu/matkut/eredmenyek/szakdolgozat-pL-sign.pdf) Ennek 18. oldalán ez olvasható:

„A zárt végű (feleletválasztó) feladatok jellemzői számos ponton fordítottnak tekinthetők: legnagyobb előnyükként a gyors javíthatóságuk; automatizált, számítógépes (online) formában is széleskörű alkalmazhatóságuk kiemelkedő. Hagyományosan kevéssé alkalmasnak tartják őket magasabb rendű tudáselemek mérésére, mivel ráismeréssel, vagy egyenesen találgatással (vaktalálat) nem elhanyagolható esély mutatkozik a helyes válaszra.”

A 20. oldalon Pitlik László kitér a levonásos pontozásra is (egyválasztásos teszt esetén):

„A nemzetközi szakirodalomban a tippelés korrekciójának (correction for guessing) nevezett módszer esetében az alábbi összefüggés adható meg. (KÁDÁRNÉ; 1973, p.22) $T_c = H - R/(L - 1)$ A teszten szerzett korrigált pontszám (T_c) a helyes válaszok (H) számához képest az elrontott feladatok (R) bizonyos hányada – az eredetileg adott válaszlehetőségek (L) száma – alapján pontlevonást alkalmaz úgy, hogy a tippelés várható értéke éppen zérus legyen. Mivel a levonás ténye önmagában ellentmondásos üzenetet hordozhat, így létezik a kihagyásokra (K) jutalmazó, lényegében ekvivalens formula. (EBEL & FRISBIE; 1991, p.212) $T_c' = H + K/L$.

[…] Magyarországi gyakorlatban az alábbi, a jelen kutatásban is alkalmazott szóhasználat terjedt el: „helyes válasz: $+X$ pont, rossz válasz: $-Y$ pont” ($X \geq Y$). Ez utóbbi forma az eltérő szóhasználaton kívül megnyitja a lehetőséget – a tesztkészítő tudatosságától és egyéni preferenciáitól függően – különböző, nemnulla várható értékű pontozási rendszerek alkalmazása előtt. Hosszú távon, matematikailag: pozitív várható érték esetén a tanuló optimális stratégiája mindenkor a következetes tippelés, negatív várható érték esetén pedig a találgatás minimalizálása.”

II.

1. Amint a bevezetőben felvázoltam, innentől az I/14. alatti intuitív kijelentést fogom tételesen, számszakilag igazolni. A hivatkozások e tanulmány kísérődokumentumára, a Redmenta-levonasos-pontozas.xlsx nevű Excel-munkafüzetre vonatkoznak. Mindkettőt elhelyezem a http://leporollak.hu/tudomany/redmenta/REDMENTA.HTM címen, ahol e dokumentumok mindenkori legfrissebb változatai megtalálhatók lesznek.

2. A felvetett kérdés kapcsán mindenekelőtt a mérés *validitásának* kérdése merül fel, vagyis hogy a teszt valóban azt méri-e, amit mérni akarunk vele – ez esetben a tanuló valós tudását. A validitás mértékét a méréselméletben különböző jelzőszámok mutatják; ezek helyett én grafikonokon, százalékokban fogom ábrázolni a teszteredményt a tudás függvényében.

3. Elemzésem során négy redmentás értékelési rendszert vizsgálok e szempontból:

- *Levonás nélküli* pontozás: ahány helyes válaszmegjelölés, annyi pont.
- *Levonásos* pontozás: a pontszám a helyes és a helytelen válaszmegjelölések számának különbsége. Ez *negatív is lehet*.
- *Levonásos* pontozás, de a feladatonkénti eredmény *negatív nem lehet*: az előző

rendszer negatív kimenete esetén nullát ad az illető tesztfeladatra.

- *Csak hibátlan* válasz esetén: csak akkor jár pont, ha mindegyik helyes válaszlehetőséget bejelöli a tanuló. Vizsgálatomban e pontszám egyenlő a helyes válaszlehetőségek számával.

4. A lényegi eltérést az első és a harmadik pontozási rendszer között látom, mert a gyakorlatban ezek a legnépszerűbbek. Ezért a hozzájuk tartozó pontszámokat és százalékokat jelölöm fekete számokkal, a többit – mint kevésbé fontosakat – szürkével.

5. A vizsgált *teszttípusok* közül négyet emelek ki:

- 4 válaszlehetőségből 2 helyes (Teszt 4-2),
- 6 válaszlehetőségből 3 helyes (Teszt 6-3),
- 6 válaszlehetőségből 2 helyes (Teszt 6-2),
- 8 válaszlehetőségből 3 helyes (Teszt 8-3).

III.

Vizsgálatom első része (a „determinisztikus modell") a következő **korlátokkal** rendelkezik:

1. Nem kerítem sorra azt a körülményt, hogy az egyes válaszlehetőségek *függhetnek egymástól*: pl. objektíve egyenértékűek, kizárják egymást, egyikből következik a másik stb., esetleg a tanuló szubjektív tudatában fennállhat efféle összefüggés. Tehát a továbbiakban feltételezem, hogy a felkínált válaszlehetőségek teljesen függetlenek, vagyis az egyik be (nem) jelölése semmiben nem befolyásolja a másiknak a be (nem) jelölését.

2. A tanulók tudásszintjét deklaráltan fekete-fehér jellegűként kezelem, azaz akkor tekintem létezőnek az adott válaszlehetőségről birtokolt tudását, ha *biztosan és helyénvalóan tudja* róla, hogy helyes-e vagy helytelen. Tehát nem vállalkozom olyan részletek elemzésére, hogy a tanulónak rémlik valami az adott válaszlehetőségről, de nincs róla biztos tudása. Azt sem veszem számításba, hogy magabiztosan, de hibásan „tudja" valamely válaszlehetőségről, hogy helyes (vagy nem). Természetesen kizárom a félreütés lehetőségét is, amikor a tanuló a biztosan tudott helyes vagy helytelen válaszlehetőséget elfelejti bejelölni, vagy fordítva jelöli be. Vagyis *a biztosan tudott válaszlehetőségeket garantáltan be (nem) jelöltnek tekintem,* és ekképpen modellezem úgy a tábla bal oldalán, az elméleti számításokban, mint a jobb oldalán, a véletlenszám-generátoros modellezésben. Ezeket az egyedi válaszlehetőségeket a továbbiakban *nemtippeknek* nevezem, a belőlük alkotott rendezett szám-n-eseket pedig nemtipp-kombinációknak.

3. Ezen túlmenően egyazon tudásszint és tudáseloszlás esetén *a nemtippeket teljesen egyenlő nehézségűeknek tekintem egymással*. Ez a megkötés (az „**egyenértékűségi hipotézis**") a legnagyobb hiányossága a determinisztikus modellnek, mert feltételezi, hogy két tanuló, aki az öt hibás (4;5;6;7;8) válaszlehetőségből pontosan hármat tud

biztosan (vagyis ugyanolyan tudáseloszlásuk van a hibás válaszlehetőségek körében), ugyanolyan valószínűséggel tudja biztosan pl. a 4;5;8 és a 6;7;8 nemtipp-kombinációt. Hasonlót feltételez a helyes válaszlehetőségekről is.

4. Ez a valóságban szembeszökően nincs így, hiszen az életben előforduló tesztbeli válaszlehetőségek között vannak könnyebben és nehezebben tudhatók, s ekképpen gyakrabban és ritkábban tudottak. Mégis feltételeztem e modellben a tudott állítások ezen egyenlő nehézségét, mert csak így tudtam determinisztikusan (azaz véletlenszám-generátor nélkül) is dolgozni az egyes tudásszinteken lévő tanulók nemtippjeinek számával, és *így adhattam becslést az illető tudásszintű tanulók által elért tesztbeli pontszámok várható értékére*.

5. A többi, tehát az összes, nem biztosan tudott válaszlehetőségre nézve feltételeztem, hogy a tanuló róluk *teljesen véletlenszerűen dönt*. Ezt az egyedi választ „tippnek” nevezem.

6. Természetesen lehetségesek bonyolultabb (kevert vagy bizonytalan) tudásszerkezetek és válaszadási stratégiák is, de ezek elméleti vizsgálatát nem tartom feladatomnak, hiszen a determinisztikus vizsgálatban azt vállaltam, hogy *a különféle értékelési rendszereket* hasonlítom össze egyenértékű feltételek között. Az itt vizsgált teszttípusok, tudásszintek és stratégiák kiválasztása azt a célt szolgálta, hogy minél változatosabb feltételek között történhessen meg az általam vállalt összehasonlítás. Úgy vélem, joggal feltételezhetem, hogy más észszerű alapfeltevések esetén is *hasonló viszonylagos eredményeket* kaptam volna a tudás és az eredmény között.

7. Éppen e korlátozottság oldása végett készítettem két nem-determinisztikus, paramétervezérlésű, véletlenszám-generátoros *modellezési lehetőséget* az Excel-állomány 5-6. munkalapján (1./2. modellezés 8-3), melyek számításba veszik a válaszlehetőségek különböző nehézségét, és ezekre alkalmasan választott ikszelési mintázatot állítanak elő. Ez a két modell éppen ott tartalmaz némiképp önkényes feltevéseket, ahol *a tudásviszonyokból válaszadási mintákat kell létrehozni*. E feltevések csak valószínűsítő jellegűek, és rajtuk kívül nyilván több más válaszadási stratégia is elképzelhető. De látni fogjuk, hogy az 1. nem-determinisztikus modellben a valószínűsítő erő különösen nagy.

8. A 6-3-as és 6-2-es vizsgálatban kitérek az ún. *merész és óvatos* tippelési stratégiák közti különbségre is. A merészen tippelő tanuló definíció szerint minden lehetséges tippet megtesz, az óvatosan tippelő pedig ennél kevesebbet tesz meg. Ez esetben bizonyos tudásszinteken és tudáseloszlások esetén már nincs mód óvatos tippelésre, hiszen ekkor a biztos tudás alapján a tanuló eleve minden lehetséges bejelölést megtesz.

IV.

Az Excel-táblákban a következő **jelölésmódokat** alkalmazom:

1. A 8-3-as munkalapon szereplő első három válaszlehetőség *mindvégig helyes*, az utánuk következő öt *mindvégig helytelen*. A többi munkalapon hasonló elrendezést alkalmazok.
2. A bal oldalon kettős keretbe foglalom a biztosan tudott válaszlehetőségeket. Ezeket az adott szakaszban *biztosan bejelöltnek* tekintem, ha a *helyes* válaszlehetőségek közé tartoznak, illetve *biztosan be nem jelöltnek*, ha a *helytelen* válaszlehetőségek közé. A kettős keretezést az áttekinthetőség érdekében nem ismétlem meg, helyette a munkalap mindkét oldalán sárga hátteret alkalmazok.
3. Az egyes szakaszokat egyszeres vízszintes vonallal választom el egymástól.
4. A jobb oldalon a zöld oszlopokba tett *1-es számmal* jelzem a tanuló által *bejelölt* válaszlehetőségeket, függetlenül attól, hogy azok bejelölése helyes-e vagy nem.
5. Mindkét oldalon kávébarna színnel jelölöm az értékelőmezőket.
6. A munkalap jobb oldalán „ikszelésnek” nevezem a tippek és nemtippek együttesét. A tippeket zöld, a nemtippeket sárga háttérrel emeltem ki.

V.

Most rátérek a determinisztikus modellben, vagyis az első négy Excel-munkalap bal oldalán alkalmazott **összeszámlálási módszerre**.

1. Először megválasztom a *tudásszintet*. Ez a válaszlehetőségek számától függően lehet kerek negyed, hatod vagy nyolcad alakú tört, százalékban pedig értelemszerűen kerekített szám. A példa legyen *a Teszt 8-3 táblában az 50%-os tudásszint*, ami négy tudott válaszlehetőséget jelent a nyolcból.

2. Ezután a választott tudásszinten belül a helyes és a helytelen válaszlehetőségek közti *tudáseloszlás alapján* különböző szakaszokat hozok létre, kettős kerettel jelölve a *biztosan tudott, és ennek alapján be (nem) jelölt válaszlehetőségeket* („nemtippeket”). E szakaszokat a mondott tudásszinten a következő nyolctagú kódok vezetik be: 111_ _ _ _0, 11_ _ _ _00, 1_ _ _ _000 és _ _ _ _0000. Itt a számjegyek a nemtippeket jelölik, azaz konkrétan az 1-esek a biztos tudás alapján *bejelölt helyes* válaszlehetőségeket, a 0-k pedig azokat a *helyteleneket*, amelyeket a tanuló biztos tudás alapján *kizár* a bejelölhetők közül. A _ jelek pedig a szoftverrel később tudás nélkül, tisztán kombinatorikusan kitöltendő válaszlehetőségeket (a továbbiakban: „*tippeket*”) jelölik. A kódok a konkrét tesztfeladatnak a tanár által a szerkesztéskor sorba állított válaszlehetőségeihez igazodnak, függetlenül attól, hogy a Redmenta rendszerint megkeverve tálalja őket a tanulók elé.

3. E kódok jelentése további magyarázatot igényel, ugyanis a felsorolásból hiányzik pl. az 1_1_ 0_ _0 kód, amely ugyancsak az 50%-os tudásszinthez tartozna. (Valójában az 11_ _ _ _00 kódnak van alárendelve.) Az általam alkalmazott, helytakarékossági célú jelölési konvenció azon a fentebb taglalt **egyenértékűségi hipotézisen** alapul, hogy az egyes nemtipp-kombinációk objektíve *egyenlő valószínűségűnek tekinthetők* az adott tudásszint alá rendelt egyes szakaszokon belül. Vagyis ha azt a szakaszt vizsgáljuk, melyben két tanuló a 8 válaszlehetőségből pontosan két helyesről és két helytelenről tudja a helyességet biztosan eldönteni, akkor nála az 1;2;7;8 nemtipp-kombinációnak ugyanakkora a valószínűsége, mint az 1;3;5;8-nak.

4. Ha mármost az említett egyenlő valószínűség fennáll, akkor a 11_ _ _ _00 kód *egymaga képviselheti* mindazokat a *bejelöléseket*, ahol a három helyes válaszlehetőség közül a tanuló pontosan kettőt jelöl be, és az öt helytelen közül szintén kettőt (röviden: a 2+2 tudáseloszlást). Ezen bejelölések száma annyi, ahányféleképpen 3 helyesből 2-t, és 5 helytelenből 2-t lehet választani, azaz (3 alatt a 2) · (5 alatt a 2) = 3·10 = 30. Ezt a gyakorisági számot neveztem az egyes *biztostudás-kódhoz* tartozó *súlynak*.

5. A súlynak az a jelentés is adható, hogy *ennyi tanulót* tételeztem fel az 50%-os tudásszinten, akik éppen ezt a 2+2 tudáseloszlást birtokolják. Erre az a körülmény adott módot, hogy a valódi életben a Redmentán egyetlen tanuló *végül csak egyféle módon tölti ki* az illető 8-3-as tesztfeladatot, noha kitöltés előtt számos kitöltési lehetőség állt előtte.

6. A *tippelt* válaszokat az összes lehetséges módon a tudott válaszok közé illesztve készítettem egy vagy több *teljes válaszsort* az illető szakaszba, amelyek azonban – az előző pontban mondottak alapján – nem jelentik azt, hogy az illető kódhoz tartozó *egyetlen* tanuló ennyiféle válaszvariációt meg is tenne. Ő *csak egyet* fog ezek közül bejelölni, és hogy melyiket, azt nem tudhatjuk. Viszont kiszámíthattam e variációkból származó *várható pontértéket*, és ezt hozzárendeltem az illető szakaszhoz (biztostudás-kódhoz) a következőképpen:

   a) Az összes lehetséges módon kitöltöttem a _ jelekkel eddig üresen hagyott tippeket. Az 1+3 tudáseloszlásba (vagyis az 1_ _ _ _000 biztostudás-kódba) a négy üres helyre a két hiányzó 1-est és két hiányzó 0-t kellett beírnom az összes lehetséges módon. Ez hat sort (teljes ikszelést) jelentett.

   b) Az így elkészült ikszeléseket sorra alávetettem a négy, általam választott pontozási rendszernek, és gépileg átlagot számoltattam mindegyik alapján a hat sorból. Ez lett az adott biztostudás-kódhoz rendelt *várható pontérték* az illető pontozási rendszerben. Így például a fenti biztostudás-kódhoz az egyes pontozási rendszerek a következő várható értékeket szolgáltatták:

      - a levonás nélküli rendszer 2 pontot;
      - a levonásos, negatívba is menni tudó rendszer 1 pontot;
      - a levonásos, negatívba menni nem tudó rendszer 1,1667 pontot;

- a „csak helyes” rendszer 0,5 pontot.

7. Az itteni 4. pont szerint előállítottam az illető biztostudás-kódhoz rendelt *súlyokat*:

- az 111_ _ _ _0 szakasz súlya: (3 alatt a 3)·(5 alatt az 1) = 1·5=5;
- az 11_ _ _ _00 szakasz súlya: (3 alatt a 2)·(5 alatt a 2) = 3·10 = 30;
- az 1_ _ _ _000 szakasz súlya: (3 alatt az 1)·(5 alatt a 3) = 3·10 = 30;
- a _ _ _ _0000 szakasz súlya: (3 alatt a 0)·(5 alatt a 4) = 1·5 = 5.

8. A várható értékeket a súlyokkal rendre megszorozva, e szorzatokat összeadva, és az összeget a súlyok összegével elosztva elkészítettem *a várható értékek súlyozott átlagát*. Ez az 50%-os tudásszintre nézve a következő eredményeket hozta:

- a levonás nélküli rendszer 2,196 pontot;
- a levonásos, negatívba is menni tudó rendszer 1,393 pontot;
- a levonásos, negatívba menni nem tudó rendszer 1,464 pontot;
- a „csak helyes” rendszer 0,804 pontot.

9. Ezen várható pontértékeket az elérhető 3 pont százalékában kifejezve megkaptam azt az *eredményt*, amelyre egy 50%-os tudásszintű tanuló a négy pontozási rendszerben *számíthat*, ha az általa nem tudott válaszlehetőségeket merész tippeléssel tölti ki:

- a levonás nélküli rendszerben 73,21%-ra;
- a levonásos, negatívba is menni tudó rendszerben 46,43%-ra;
- a levonásos, negatívba menni nem tudó rendszerben 48,81%-ra;
- a „csak helyes” rendszerben 26,79%-ra.

10. Az így kapott értékek azt mutatják, hogy az 50%-os tudásszinten a levonás nélküli rendszer *23 százalékponttal felülpontozza a tanulót*, míg a két levonásos rendszer sokkal közelebb áll a valós tudáshoz (szaknyelven kifejezve: nagyobb validitású értékelést ad). A „csak helyes” rendszer 23 százalékponttal alulpontoz e tudásszinten.

11. A fent részletezett elemzést sorra elvégeztem a 8-3-as tesztfeladathoz tartozó tudásszintekre 0%-tól 87,5%-ig. A kapott eredményeket a legalsó táblázatban összesítettem, és grafikonon ábrázoltam. Ebből a következők látszanak:

- *A levonás nélküli rendszer minden tudásszinten felülpontoz*, és minél gyengébb a tanuló tudása, arányaiban annál inkább.
- A levonásos, negatív rendszer alacsony tudásszinten erősen alulpontoz, de 50%-os tudás felett nem;
- *A levonásos, nemnegatív rendszer a négy közül a legközelebb áll a maximális validitást jelző szaggatott* piros *egyeneshez;*
- A „csak helyes” rendszer 75%-os tudásszint alatt erősen alulpontoz.

12. Ezzel voltaképpen megválaszoltam a tanulmány bevezetőjében föltett kérdést. A Redmentának *nagyon jó oka volt a levonásos pontozás alapértelmezetté tételére*, és

azok a hangok, amelyek ezt a pontozást bármely meggondolásból pedagógushoz méltatlannak vagy igazságtalannak próbálták bélyegezni, nem állják ki a tüzetes vizsgálat próbáját. A felvetett kérdést ugyanis nem a mínuszpontokról szóló akadémikus, dogmatikus kijelentések vagy érzelemtől fűtött minősítések dönthetik el, hanem csakis a *plauzibilis feltevéseken nyugvó, szigorú elméleti számítást vagy valósághű modellezést alkalmazó bizonyítás*.

13. Egyetlen ellenvetés tehető e következtetés ellen, ti. az, hogy az eredményeket az általam egyszerűsítésképpen felállított „egyenértékűségi hipotézis" alapján kaptam. Erre azonban könnyű a felelet az egyik speciális eset említésével. *Nulla százalékos tudásszinten* nincsenek semmiféle tudáseloszlási szakaszok, és ezek várható értékei és a súlyozott átlagaik sem képezendők. Tehát *az egyenértékűségi hipotézis itt elhagyható, mert nincs rá szükség*. Itt elég az összes lehetséges kitöltést (vagyis tippet) felsorolni, és rájuk egy közönséges átlagolást alkalmazni – mind a négy pontozási rendszerben. A grafikon görbéinek (ezen esetet leíró) *függőleges tengelymetszetei* akkor is kényszerítő erővel mutatják a négy rendszer validitása közti eltérést, ha elejtjük az egyenértékűségi hipotézist.

14. A kérdés tehát eldőlni látszik. A Redmenta korábbi, **levonásos alapértelmezése volt a valósághoz hűbb pontozás**, **míg a levonás nélküli rendszer, amit sokan hangosan követeltek, éppenséggel (a leg)távolabb áll attól**. Ez a következtetés annyira meggyőző, hogy a 4-2-es tesztfeladatokra nézve az ellenvetést tevő tanulók kedvéért *a táblán is le lehet vezetni*. Valahányszor diákjaim az óráimon követelni kezdték, hogy bíráljam felül a Redmenta alapértelmezését, csak eléjük kellett tárnom a következő táblázatot, amit *a Redmenta levonásos pontozásáról szóló ábécéskönyvnek* nevezhetek. Ebben a tudásszint 0%, így az összes kitöltés vaktában és merészen történik. Az A és B válaszlehetőség helyes, a C és D helytelen. Az elérhető pontszám 2.

A	B	C	D	Levonás nélkül	Levonással, nemneg.
0	0	1	1	0	0
0	1	0	1	1	0
0	1	1	0	1	0
1	0	0	1	1	0
1	0	1	0	1	0
1	1	0	0	2	2
				Átlag: 1	Átlag: 0,33
				Eredmény: 50%	Eredmény: 16,7%

15. A következtetés annyira lehengerlő, hogy annak láttán a tanulók rögtön elálltak a meggyőzésemtől. És valóban: hogy is lehetne igazságos az a pontozási rendszer, amely egy csupa 4-2-es tesztkérdésből álló redmentás dolgozatnál *a semmit nem tudó, vaktában tippelgető tanulóknak statisztikailag 50%-os eredményt valószínűsít?* És hogyan lehetne ez „az egyetlen korrekt pontozás?" A választ rábízom az értő olvasóra, és tanulmányomat néhány megjegyzéssel folytatom.

16. A tippelgetéssel kitöltött kétválasztásos teszt voltaképpen a teljes válaszlistán végrehajtott kételemű, visszatevés nélküli mintavétel. Ismeretes, hogy ilyenkor a mintában a helyes válaszok várhatóan ugyanolyan arányban jönnek elő, mint a teljes sokaságban. Tehát vaktában tippelgetve, majd *levonás nélkül értékelve* a várható pontszám a maximálisnak ugyanakkora hányada lesz, mint a helyes válaszlehetőségek aránya az összeshez. Ez esetben a feladatsor készítőjének olyan többválasztásos tesztfeladatokat érdemes készítenie, amelyekben *a „jó per összes" arány az elégtelen érdemjegy tartományában marad*, iskolánkban például 30% alatt. De nem várható el a tanártól, hogy már a kétválasztásos tesztfeladatokban következetesen még *legalább öt hibás válaszlehetőséget* izzadjon ki magából – ezek jelentős része nagy valószínűséggel komolytalan, azaz *ordítóan helytelen lesz,* és a feladat önmaga paródiájává válik.

17. Itt térek ki Visy Zoltánnak a tanulmány elején idézett nyilatkozatára, mely szerint a Redmenta csapata a vaktában tippelgetés visszaszorítása végett vezette be a levonásos pontozást. Ennek ellenőrzése végett elég lesz összevetni pl. a 6-3-as tesztre vonatkozó grafikonokban a levonás nélküli és a levonásos nemnegatív pontozást. Látható, hogy 0%-os tudású, merészen tippelő tanuló esetében az eredetileg 50%-os várható eredményt 20%-ra szorítja le a levonásos értékelés. Ennek alkalmazása esetén viszont a tanulónak *nem lesz érdeke, hogy óvatosan tippelgessen*, mert úgy csak 14,3%-ra számíthat. Vagyis az idézett indoklás nem helyénvaló. A Redmentán alkalmazott levonásos pontozásnak nem tartozik az erényei közé, hogy visszafogja a vaktában totózást. (Ehhez a hibás válaszra egynél több mínuszpontot kellene adni.) Helyesebb volna így fogalmazni: a levonásos pontozás célja az volt, hogy *korlátok közé szorítsa a tesztfeladatokra eredendően jellemző felülpontozási tendenciát*, vagyis csökkentse a tippelgetéssel megszerezhető ingyenpontok számát.

VI.

1. Érdemes még megvilágítani az első négy Excel-munkalap jobb oldalán lévő, véletlenszám-generátoros ellenőrző táblázatok felépítését. Itt olyan nyolcelemű ikszelési sorokat kellett generálnom, melyekben (legfeljebb) három bejelölés van (itt ezeket végig egyesek jelzik). Mivel az Excelben nincs olyan függvény, amely véletlenszerűen előállítaná 8 elem 3;5-ödosztályú ismétléses permutációit (vagy 8 elem 3-adosztályú kombinációit), ezt kézzel kellett megoldanom.

2. Az első oszlopokban véletlen módon létrehoztam három sorszámot, pl. a 8-3-as tesztfeladat 0%-os tudásszintjén először nyolcfélét (0…7), aztán hétfélét (0…6), majd hatfélét (0…5). Itt a nulla mint kezdőelem a kódolás miatt volt szükséges. Ha a sorsolt számok pl. rendre 4/2/5-nek adódnak, akkor az első beikszelt válaszlehetőség a 4-es, a második a 2-es, de a harmadik nem az 5-ös, mert ez utóbbi szám azt jelenti, hogy az első két ikszelés után *megmaradt* válaszlehetőségek közül a hatodikat kell bejelölni, ami ez esetben a 7-es, vagyis az utolsó. Ezért a harmadik sorsolt számot szoftveresen kétszer át kellett ugratnom a már beiktatott sorszámokon. Így jöttek létre a kék hátterű oszlopokban található sorszámok, melyek azt jelölik, hogy a 0…7 kódú nyolc

válaszlehetőségből *mely hármat ikszeljük be*. Hasonlóan az 0/0/0 sorsolt számokból 0;1;2 ikszelés lesz, ami az első három pozíciót jelenti.

3. Érdemes megjegyezni, hogy minden egyes kitöltéshez *hat különböző sorsolás* tartozik, amelyek emezt állítják elő. Például a 2;3;5 kitöltést a következők: 5/3/2, 5/2/2, 3/4/2, 3/2/3, 2/2/3 és 2/5/2 – aszerint, hogy a 2;3;5 válaszlehetőségeket milyen sorrendben ikszeljük be. (Ezt könnyű ellenőrizni, mert akkor van szükség sorszámugratás(ok)ra, ha az itt szereplő számok valamelyike nagyobb az előtte lévő(k)nél.) Tehát belátható, hogy a sorsolás egyenlő valószínűséggel állítja elő az összes lehetséges bejelölést.

4. Ezután a zöld mezőkbe, a véglegesített sorszámokon el kellett helyeznem egy-egy 1-est, és a többi helyre 0-t. Ezt úgy végeztem el, hogy a sorszámokból 2 hatványai segítségével készítettem egy kódot, majd ebből maradékos osztásokkal generáltam a szükséges 0 és 1 helyiértékes jegyeket, mintegy a 2-es számrendszerbe való átírással.

5. Ha valamely tudásszinten bizonyos válaszlehetőségek már rögzítetten ki voltak töltve, illetve ki volt zárva a kitöltésük, akkor ezeket sárga háttérrel írtam be az „ikszelések” közé, és a kék oszlopokban látható, értelemszerűen csökkentett hatókörű sorsolással hoztam létre a fennmaradó, zöld hátterű 0-kat illetve 1-eseket.

6. Ugyanígy csökkentettem a véletlenszám-oszlopok számát, ha a feladatban *óvatos* tippelés szerepelt.

7. Ha valamely tudásszinten és szakaszban a tippelésre már nem volt mód, hanem minden mezőbe *kötelezően meghatározott módon* kellett 0/1 számjegyeket beírnom, akkor az adott sorba ezeket kézzel, véletlengenerálás nélkül írtam be – akár a zöld mezőkbe is.

8. Egy adott szakasz (véletlen vagy meghatározott) sorsoló-sorát *annyiszor szerepeltettem* a táblázatban, amennyi az illető tudáseloszlás *súlya* volt, a bal oldali elméleti számításban meghatározott módon. Ezután a jobb oldali táblázaton alkalmaztam a négy tárgyalt pontozási rendszert, majd a pontszámokon egy egyszerű átlagolást végeztem. Az eredményeket százalékban is megadtam.

VII.

1. Az egyenértékűségi hipotézis egy komoly megkötése az eddig használt, determinisztikus számítási és sorsolási modellnek. Ezért kerestem más utakat is a kívánt eredmény előállításához. Szigorú kombinatorikai eszközökkel nem tudtam olyan összeszámolást készíteni, amely figyelembe vette volna az egyes válaszlehetőségek könnyebb vagy nehezebb tudhatóságát (megtippelhetőségét). Ezért több helyen *tapogatódzó képletekkel kifejezhető összefüggéseket* kellett feltételeznem az egyes válaszlehetőségekre vonatkozó (tömegszintű vagy egyéni) *tudás* és ezek *igaznak tartása* között. Emellett a nem tudás bizonytalanságot okozó hatásának modellezése végett erősen támaszkodtam a *véletlenszám-generátorra*. Az itt ismertetett két modell tehát nem-determinisztikusnak nevezhető. Természetesen statisztikai értelemben továbbra is alkalmat adnak érvényes következtetések levonására.

2. Az első, *szorzatsúlyosnak* nevezhető módszer (1. modellezés 8-3) lényege az volt, hogy felsoroltam a 8-3-as munkalap első táblázatának lehetséges 56 sorát, majd a nyolc válaszlehetőséghez (fix, vagy adott érték körül szimmetrikus véletlen, vagy 0…1 közti, sűríthető véletlen módon előállított) *nehézségi számokat* rendeltem. A 0 jelenti azt, hogy az állítás a lehető legnehezebb, vagyis róla semmit nem tud a kitöltő. Az 1 pedig azt, hogy biztos tudással rendelkezik róla, vagyis számára ez az állítás a lehető legkönnyebben eldönthető. Az így értelmezett nehézségek egyben a tanuló tudását is jellemzik.

3. Ezután a tudásszintből egy ad hoc, determinisztikus képletpárral előállítottam az illető válaszlehetőség *igaznak tartási valószínűségét*:

   - *Helyes* válaszlehetőségeknél feltettem, hogy a tanuló 0 tudás esetén valamely „egyedi alapvalószínűséggel" (pl. 0,5-del) választaná ezt (ha csak ez az egyetlen kérdés állna előtte), illetve biztos tudás esetén 1 valószínűséggel. E két véglet közé lineáris (növekvő) összefüggést interpoláltam.
   - Ha a válaszlehetőség *helytelen* volt, akkor 0 tudás esetén szintén a 0,5-es alapvalószínűséget feltételeztem az igaznak tartásra, illetve 1 tudásszint esetén a 0 értéket. E két végpont közé szintén lineáris (csökkenő) függvénykapcsolatot vázoltam.

4. Ezután képeztem minden válaszlehetőség igaznak tartási valószínűsége és az adott tippsorbeli ikszelése (0 vagy 1) közti nemnegatív *különbséget*. E különbségeket 1-ből kivonva egy durva becslést kaptam arra, hogy mennyire lesz *népszerű* az illető sorban beikszelt válaszlehetőség a róla adott mértékű tudással rendelkező tanuló számára.

5. Az ikszelési *sorok* relatív gyakoriságával arányosnak veendő súlyokat úgy állítottam elő, hogy képeztem az egyes sorbeli nyolc népszerűség-becslés *szorzatát*. Ez a súlyozás teljesítette a következő észszerű követelményeket:

   - teljesen lehetetlen ikszelésekre 0 súlyt szolgáltatott;
   - monoton növő módon függött az egyes népszerűségektől;
   - valami módon a tanulónak mindegyik válaszlehetőségről birtokolt tudása szerepet kapott benne.

6. Ezután képeztem az egyes ikszelési soroknak a négy választott pontozási rendszerben szolgáltatott pontszámát, és ezeket az imént képzett súlyokkal átlagoltam.

7. A tapasztalat az volt, hogy az utolsó sorok, vagyis a szinte hibátlan kitöltések súlyai *alaposan felülmúlják a többit*, és közülük is kimagaslik az utolsó sor. A súlyok nagyságrendileg is erősen különböztek egymástól. Ez önmagában nem elfogadhatatlan, hiszen az adott sorhoz rendelt súlyt (vagyis az adott sor szerint kitöltők relatív gyakoriságát) az egyes valószínűtlen ikszelések mindegyike tovább apasztja, és ez az apasztás osztások sorozataként modellezhető.

8. Az alapvalószínűséget először *0,5-en rögzítettem* (vagyis mintha a tanuló önmagában

tekintené az összes válaszlehetőséget), és a szűken, lineárisan szóró nehézséggenerátorral apró lépésenként végigpásztáztam a 0…1 tudásszinttartományt. Ez úgy történt, hogy a pirossal keretezett átlagolósorból az Excel program Adatok – Lehetőségelemzés menüpontjainak segítségével *adattáblát* hoztam létre, melyben *a tudásszintet szabályozó nehézséggenerátor-középértéket egyszerű ciklusban végigfuttattam 0 és 1 között.* Mind a négy pontozási rendszerre elkészítettem a „B” jelű tudás-eredmény grafikont. Ezekre alkalmas módon választott trendvonalat is illesztettem.

9. Egy másik eljárás a *sűrítőkitevős* nevet viseli, és azon a jelenségen alapul, hogy az 0 és 1 közti kitevőjű hatványfüggvények konkávok, az 1-nél nagyobb kitevőjűek konvexek, így az előbbiek a véletlenszám-generátor szolgáltatta értékeket *fölfelé*, az utóbbiak *lefelé tolják el*. Az így létrehozott válaszlehetőség-nehézségek esetében is adattáblát készítettem, és a sűrítőkitevőt végigfuttattam néhány alkalmasan választott értéken. Ezután az átlagolásból származó százalékos eredményértékeket is ábrázoltam a tudás függvényében („A” grafikon).

10. A két grafikontípus között egy *főkapcsoló* nevű mezőbe való számbeírással lehet lépkedni. Ilyen esetben csak a választott grafikon mutat jól (ezt jelzi a főkapcsoló által felvett szín), míg a másik grafikon az illesztett görbék típusa és a tengelyek skálázása miatt eléggé esetlenül fest.

11. A modellezési helyzet másfelől elgondolható úgy is, hogy a tanuló *nem önmagában* tekinti az illető válaszlehetőséget, hanem mint egy 8-ból 3-as választásos feladat egy elemét, és ilyenkor az igaznak tartás elemi alapvalószínűsége egyenlőnek vehető a helyes/összes aránnyal, vagyis 3/8=0,375-del. A szorzatsúlyos modell tudja kezelni ezt a változtatást is: egy kék mezőbe kell beírni az alapvalószínűség megváltozott értékét.

12. A szorzatsúly-modell viszonylag elfogadható tudás-eredmény összefüggést szolgáltatott (mert a görbék erősen érzékenyek voltak a tudásszintre). E modell a tudásszintek egész spektrumát képes kezelni, és plauzibilis előfeltevésekkel dolgozik. Eredményei a következők voltak:

    - A *levonás nélküli pontozás minden tudásszinten erősen föléhord a tényleges tudásnak* (tehát méréselméleti értelemben nem valid), *a nemnegatív levonásos pontozás viszont lényegében együtt halad vele* (tehát valid).
    - Az alapvalószínűség mégoly merészen változtatott értéke (amely az eljárás legfőbb paramétere) *csak enyhén tolja el függőlegesen a görbéket,* és a jellegüket nem befolyásolja.
    - A nehézséggenerátor szórása befolyásolja ugyan a függvénypontokat, de a trendvonalat alig.
    - A másik két pontozási rendszer szolgáltatta eredmény a legtöbb tudásszinten alaposan elmarad a valós tudástól.

13. Érdemes összevetni a szorzatsúlyos modell fő grafikonját a *determinisztikus modell* 8-3-as grafikonjával. A négy vizsgált pontozási rendszer görbéi igen hasonlóan futnak e

két modellben. Ilyen módon a szorzatsúly-modell a determinisztikus modell finomításának, *általánosításának tekinthető*. Mivel a két modell alapvetései közt található egy fontos különbség (az utóbbiból hiányzik az egyenértékűségi hipotézis, viszont változatos szórási paraméterrel kezeli a válaszlehetőségek különböző nehézségét), erősen kínálja magát az a következtetés, hogy *mindkettő megbízhatóan modellezi a tesztbeli ikszelés pszichológiáját mint vizsgált valóságot*.

14. Erre támaszkodva újra kimondom a V/14. pontbeli következtetést: A Redmenta korábbi, **levonásos alapértelmezése volt a valósághoz hű pontozás**, **míg a levonás nélküli rendszer, amit sokan hangosan követeltek, éppenséggel a legtávolabb áll attól**.

15. Érdemes lehet még *ellenpróbákat* végezni annak kiderítésére, hogy mennyire volt *szükségszerű* a szorzatsúly alkalmazása. Egy (itt nem közölt) munkalapon ötféleképp megváltoztatott súlyozással előállítottam egy-egy alternatív („elrontott”) tudás-eredmény grafikont. Ezekben az a közös, hogy az adott sor szerinti ikszelés valószínűsége teljesen független volt a kitöltő tudásától, vagy olykor ellentétes jellegű volt vele. Ezen ellenpróbákban a főváltozattól lényegileg eltérő, és általában látványosan rossz grafikonok adódtak. Itt csak címszavakban sorolom fel őket, de a tanulmányhoz csatolva, a honlapon közlöm a grafikonokat képformában (ELLENPR1.GIF stb.), és a címükben megadom a súlyképzés képletét is (a táblázat súlyoszlopának felső mezőjére nézve):

    a) a szorzatsúlyban egy mínusz előjel pluszra változott,
    b) a szorzatsúly a reciprokára változott,
    c) szorzat helyett minimummal képeztem a súlyt,
    d) véletlen súlyokat alkalmaztam,
    e) a súlyokat szorzás helyett összegeztem.

16. Az ellenpróbák azt igazolták, hogy a modellezés érzékeny a súlyok megválasztására, vagyis joggal beszélhetünk jó és rossz súlyozásokról. A szorzatsúly az előbbiek közé tartozik. A determinisztikus modell egy adott (pl. 50%-os) tudásszintjén minden szerepeltetett sor súlya egyenlő (és ezeket az a modell csokrokba szedi a tudáseloszlásoknak megfelelően), míg az ott nem szereplő sorok súlya nulla.

VIII.

1. Most rátérek a *sorbarendezéses modell* ismertetésére, amely a „2. modellezés 8-3” nevű munkalapon található. A modell nem ikszelési sorokhoz rendel (tudásfüggő módon) átlagolási súlyokat, hanem az egyes válaszlehetőségekhez rendel nagyságsorrendbe állítható jelzőszámokat. Definíció szerint az első három helyezett válaszlehetőség nyer, a többinek az értékelésben semmi szerep nem jut. E modell azt a kitöltési stratégiát próbálja megragadni, mellyel a tanuló minden másra való tekintet nélkül három minél valószínűbb válaszlehetőséget keres, a többire nem fordít figyelmet.

2. Ennek megfelelően a jelzőszámok *az összeadásra lesznek érzékenyek*, a konkrét értékeik aránya nemigen jut szerephez. Ez azt jelenti, hogy a kérdések (pl. tömeges mértékben érvényes) *szándékolt nehézsége e modellben eleve meghatározhatná a végső sorrendet,* és ezt mindenképpen ki kellett küszöbölnöm. Ezért mindenekelőtt véletlengenerátorral megvariáltam a nehézségjelző számokat, hogy ezt a determináltságot megszüntessem. (Mivel egy másik helyen is beépítettem variálási lehetőséget, ezt az első variálást kikapcsolhatóvá tettem.)

3. A variálás miatt az *egyedi kitöltések eredményei* nem voltak többé külön-külön vizsgálhatók (ábrázolhatók), így az ábrázolás előtt átlagolást kellett alkalmaznom. Ehhez viszont *egymáshoz igazított nagyságú tudásegyüttesek* kellettek, amit úgy igyekeztem megvalósítani, hogy a variálással kapott eredményeket úgy normáltam, hogy a legnagyobb érték 1 legyen. (Ez nem jelent valószínűséget, hanem első vagy utolsó helyezést a sorban, attól függően, hogy helyes vagy helytelen válaszlehetőségről van szó.)

4. Ezután egy, a fentiekben (VIII./3) ismertetett ad hoc képletpárral előállítottam az illető válaszlehetőség *igaznak tartási értékét.* 0 tudáshoz 0,5-et rendeltem (ez az additív értékelési mód miatt bármi más is lehetett volna), biztos tudáshoz pedig helyes válaszlehetőségnél 1-et, helytelennél 0-t. A két végpont közé lineáris függvényt illesztettem.

5. Mivel ilyen módon az igaznak tartási értékekben fennálló *legkisebb különbség* is eleve meghatározta volna két állítás erősorrendjét, a képletbe beillesztettem egy olyan hozzáadott tagot, amely *a tudáshiánnyal arányos módon két irányban lineárisan megvariálta* az előző pontbeli lineáris függvény értékét. Ennek az volt a jelentése, hogy a tanuló tippje a tudásával arányos módon egyre közelít a válaszlehetőség valóságos igazságértékéhez, de a tanuló a tudáshiányával arányos módon eltérhet e tendenciától. A paraméterezést úgy állítottam be, hogy 50%-os tudás esetén éppen elváljon egymástól egy helyes válaszlehetőség lehetséges legrosszabb, illetve egy helytelen válaszlehetőség lehetséges legjobb igaznak tartási jelzőszáma.

6. Az így kapott igaznak tartási számok közti egyszerű sorbarendezéssel megállapítottam a modellben bejelölendő három válaszlehetőséget, és képeztem az így elért tesztpontszámot mind a négy vizsgált pontozási rendszerben. Mivel a bemeneti adat (a nyolc válaszlehetőségre számított tudásösszeg) a többszöri variálás miatt erősen szórt, *nem volt érdemes soronként ábrázolni*, mert az így kapott grafikon vízszintes csíkokra esett volna szét. Először mozgóátlaggal próbáltam simítani a tudás-eredmény grafikonokat, de így is nagyon durva képet (pl. csökkenő tudás-eredmény trendvonalakat) mutattak. Ezt a mozgóátlagos simítást tanulságul meghagyom e tanulmányban, és további kísérletezés végett az Excel-munkafüzetben is.

7. Második próbálkozásként a kapott sorokat simítás végett *globálisan átlagoltam,* persze tudván azt is, hogy így az egyedi tudásszintek és a hozzájuk tartozó, igencsak rakoncátlan egyedi eredmények kapcsolata szinte teljesen el fog mosódni. A pirossal keretezett átlagolósorból a már leírt módon *adattáblát* hoztam létre, melyben *a*

tudásszintet szabályozó sűrítőkitevő értékeit szisztematikusan végigpásztáztam. Így négy erősen simított görbét kaptam – tulajdonképpen az előző ábrázolás afféle trendvonalaként. A levonás nélküli és a „csak hibátlan" pontozási rendszer görbéjében föl lehetett ismerni a simítatlan ábrázolás görbéit, a másik kettő esetében pedig kiátlagolódott a variálás okozta sok ingadozás, és a „csak hibátlan" rendszerhez hasonló alakú grafikonok adódtak. Viszont az előző modellhez képest láthatólag már 70%-os tudásszinten is beállt az eredmények 100%-ra telítődése.

8. Ezen modellt jobbára érdekességként közlöm, mert egyáltalán nem vagyok meggyőződve arról, hogy a kétszeri variálás és azok paraméterei (melyek a modell alapeszméje miatt szükségesnek látszottak) nem visznek-e be a modellbe olyan jelenségeket, amelyeket a valóság (vagyis a kitöltés pszichológiai tényezői) nem indokolnak.

IX.

1. Nem vagyok méréselméleti szakember, csak egy gyakorlatias vénával rendelkező középiskolai tanár. Ezt a tanulmányt és a benne elemzett Excel-munkafüzetet nem szánom tudományos publikációnak, és semmiféle ellenőrzésre, lektorálásra nem is bocsátottam. Ennélfogva *maradhatott számítástechnikai vagy matematikai hiba úgy egyikben, mint másikban*, és az általam levont következtetések is szabadon vitathatók maradnak.

2. Ha valamelyik olvasóm jónak látja, hogy a tanulmányra *hivatkozzék*, a nevem és az állományok föllelési helye alapján, esetleg a szakaszok és a számozott pontok alapján könnyen megteheti.

3. A tanulmányból (ill. az Excel-munkafüzetből) származó, az idézet mértékét meghaladó mennyiségű *szöveg- ill. kódfelhasználás engedélyezése* céljából bármikor elérhető vagyok a nemo44@hotmail.com címen.

4. Az általam alkalmazott *matematikai eljárásokat* a szerzői jogi törvény értelmében szabadon (vagyis engedélyem nélkül is) fel lehet használni. Erre én minden olvasómat bátorítom is – annak a közmondásnak a jegyében, hogy több szem többet lát.

5. Minden *építő visszajelzést (kritikát, méltatást)* szívesen fogadok, és megpróbálok rá személyesen is válaszolni. Ezek segíthetik a tanulmány továbbfejlesztését.

6. Ezt a munkát feleségemnek, Jakab Júliának, foglalkozására nézve biztosítási matematikusnak ajánlom.

Pécel, 2022. június 25.

Németh Ferenc