



Társadalompolitikai Programok Értékelésének Támogatása Európában

A randomizált kísérletek szerepe a társadalompolitikai programok tesztelésében

Scharle Ágota/Váradi Balázs

Vezető kutató

Budapest Szakpolitikai Elemző Intézet

Hélène Giacobino

Ügyvezető igazgató

J-PAL Europe

SPARK – Budapest, 2014. február 19.

- Hogyan lehet költséghatékonyan csökkenteni a tartós munkanélküliséget?
- Mi a korai iskolaelhagyás kezelésének legmegfelelőbb módja?
- Hogyan ösztönözzük a vállalkozási kedvet?
- Hogyan csökkenthetjük leghatékonyabban az idős emberek baleseti kockázatát?

Megbízható értékelésekre van szükség

- Gazdasági visszaesés esetén különösen fontos a források jó elosztása
- Ennek módszere a szakszerű értékelési technikák alkalmazása az új szakpolitikai programok hatásainak vizsgálatára
- A tényeken alapuló szakpolitika-alkotás megvalósítható és nagyon hatékony

Tartalom

1. Mi a társadalompolitikai kísérlet?
2. Különböző hatásvizsgálati módszerek
3. Randomizált kísérletek
4. Következtetések

Mi az a társadalompolitikai kísérlet?

- A társadalompolitikai kísérletek úgy tesztelik a szakpolitikák hatásosságát, hogy a beavatkozások tényleges, az érintettekre kifejtett hatásairól gyűjtenek bizonyítékot (evidence)
- A cél:
 - társadalmi szükségletekre innovatív válaszokat adni
 - kicsiben tesztelni a beavatkozások hatását (pilot)
 - azt vezetni be széles körben, ami sikeresnek bizonyult

Miért értékeljük a társadalmi programokat?

- **Megalapozott döntés a program elvetéséről vagy módosításáról**
 - Elérte a program a célját?
 - Pozitív hatása volt az emberek életére?
 - Melyek a siker vagy a kudarc okai?
- **A társadalompolitikai eszközök minőségének javítása**
 - Hogyan lehet a kívánt eredményt a leghatékonyabban elérni?
 - Melyek azok az általános stratégiák, amelyek más területeken is sikeresek lehetnek?
- **Legfőbb cél:**
 - Hatékonyabb programokkal sikeresebben kezelni a társadalmi problémákat
 - Több forrást biztosítani a bizonyítottan sikeres programoknak

Nem minden beavatkozást lehet értékelni

- Egy jól megtervezett kísérlethez szükséges:
 - explicit és releváns szakpolitikai kérdés
 - megalapozott identifikációs stratégia
 - kellően nagy elemzési minta
 - jó minőségű adatok
- Nem célszerű értékelni, ha:
 - túl kicsi a minta
 - a mérni kívánt hatás mindenkit érint
 - a pilot program kiterjesztése nagymértékben módosítja a hatást
 - a beavatkozást csak a kísérleti méretben/körülmények között lehet reálisan megvalósítani (pl. mert túlságosan költséges)
 - sürgős beavatkozásra van szükség

Nem minden beavatkozást kell értékelni

- A szakpolitikai döntések megalapozásához és a közpénzek felelős felhasználásához elengedhetetlen **néhány kulcsfontosságú, kiterjedt hatásértékelésbe** befektetni
- A kulcsfontosságú, **nagy hatású, stratégiai szempontból fontos vagy újszerű beavatkozásokat** szakszerűen kell értékelni
- Nincs hozzáadott értékük az olyan programok vagy szakpolitikák értékeléseinek, amelyek csak nagyon kevés embert érnek el, vagy amelyeket szakszerűen értékeltek korábban hasonló kontextusban

Finanszírozási és etikai kérdések

- A szakszerű értékeléshez **jó minőségű adatok** kellenek: ezért lehet költséges egy kísérlet, ha nagy a minta.
 - Az is pazarlásához vezethet, ha nagyméretű beavatkozásokat hatásvizsgálat nélkül finanszírozunk
 - Átváltás a tesztelés alapos előkészítése és a hatásvizsgálat bonyolultsága között
- A társadalompolitikai kísérleteket, mint minden embereket érintő kísérletet, az **etikai elvek** figyelembevételével kell megtervezni (akkor is, ha nincsen erre vonatkozó szabályozás)
- Mindig van lehetőség a kísérletet úgy **megtervezni**, hogy az igazságos, transzparens és költség-hatékony legyen

Folyamatértékelés vs. hatásértékelés

- **Folyamatértékelés (monitoring):** A terv szerint implementálták a programot?
 - Kiosztott könyvek száma, képzést elvégző tanárok száma, meglátogatott iskolák száma, elköltött források: hogyan és mikor, kvalitatív terepmunka...
- **Hatásértékelés:** Hogyan hatott a beavatkozás a résztvevőkre?
 - A gyerekek jobb tanulmányi eredményt értek el? Ha igen, mennyivel?
 - Kinek javult inkább a teljesítménye (lányoknak? fiúknak?)
 - Hosszú távú eredmények: magasabb lett a jövőbeni keresetük?

Hatásértékelés

- **A hatásértékelés** összeveti a résztvevők (egyének, háztartások, iskolák...) eredményeit azzal, hogy mi történt volna ugyanezekkel a résztvevőkkel a program hiányában
- **A tényellentétes állapot** egy olyan hipotetikus helyzetet jelent, amelyben a programrésztvevők akkor lennének, ha a program nem valósult volna meg (vagyis nem vettek volna részt a programban)

Probléma: a tényellentétes állapotot nem tudjuk megfigyelni
→ létre kell hoznunk egy viszonyítási csoportot

Hogyan végezzünk társadalompolitikai kísérletet?

- A társadalompolitikai programokat különböző módszerekkel lehet értékelni:
 - A randomizált-kontrollált kísérlet (Randomized Controlled Trial, RKK) az egyik legjobb módszer a szakpolitikai hatások értékelésére
 - Más módszerek is léteznek, erről szól a következő rész

Tartalom

1. Mi a társadalompolitikai kísérlet?
2. Különböző hatásvizsgálati módszerek
3. Randomizált kísérletek
4. Következtetések

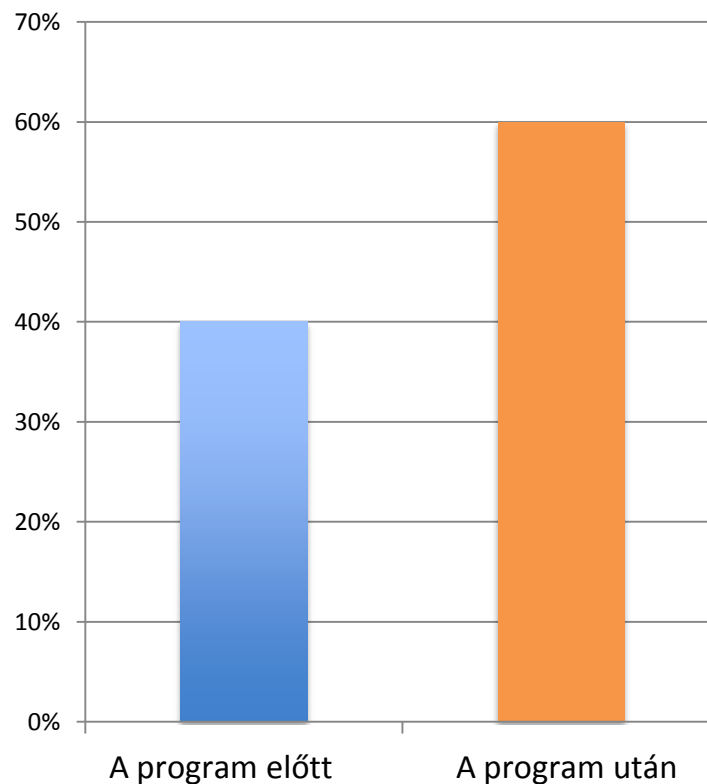
Nem-randomizált módszerek

- A nem-randomizált módszerek mindig erős feltevéseken alapulnak
 - A leggyakoribb módszerek közül néhány:
 - *Előtte-utána összevetés*
 - *Különbségek különbsége*
 - *Szakadósos vizsgálat*
 - *Statisztikai párosítás*
 - A kutatóknak és a döntéshozóknak óvatosan kell az eredményeket értelmezni, a feltevések hihetőségének kritikus átgondolása szükséges.
-

“Előtte-utána” összevetés

Egy felzárkóztató oktatási program hatásainak vizsgálata
vagy: baleseti statisztika sebességkorlátozás bevezetése előtt és után

- A kimenet változásának vizsgálata egy adott mintában, a program előtt és után mért kimenetek összevetésével
- Ennél a módszernél a viszonyítási csoportot maguk a programrésztvevők jelentik a program kezdete előtt



“Előtte-utána” összevetés

Feltevések

- Bármely változás a kimenetben közvetlenül a programnak tudható be

Gyengeségek

- Lehetetlen teljes biztonsággal meghatározni, hogy a változás mekkora részét okozta a program és mennyiben köszönhető egyéb tényezőknek, mint pl.
 - az osztálylétszámot csökkentő intézkedések,
 - az iskolai tanterv változása,
 - jobb tanárok,
 - az idő múlásával együtt járó változások, stb.

Különbségek különbsége

- A kimenetek változását veti össze időben résztvevők és nem-résztvevők között.
 - a kimenet változása a beavatkozási csoportban
(természetes változás + a programnak köszönhető változás)
 - a kimenet változása a viszonyítási csoportban
(természetes változás)
- = a programnak köszönhető változás

Egy felzárkóztató oktatási program hatásainak vizsgálata

	Előtte	Utána	Különbség
A programban résztvevő gyerekek átlagos eredménye	24.80	51.22	26.42
A programban nem résztvevő gyerekek átlagos eredménye	36.67	56.27	19.60
Különbség			6.82

Különbségek különbsége

Feltevések

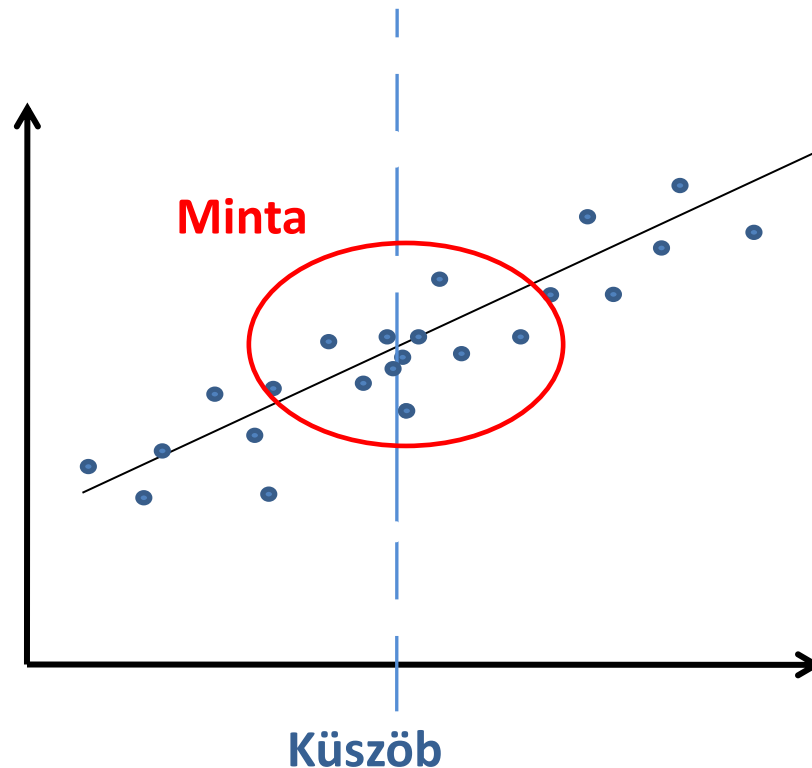
- Párhuzamos trend feltétel: a program hiányában a résztvevők és a nem-résztvevők kimenetei nagyjából ugyanúgy alakultak volna

Gyengeségek

- A program előtti trendek különbözősége torzíthatja az eredményt
- A program előtti trendek hasonlóságának ellenőrzése sok adatot igényel a beavatkozás előtti időszakra vonatkozóan, mind a résztvevőkről, mind a nem-résztvevőkről
- A programon kívüli helyi változások módosíthatják a trendeket és torzíthatják az eredményeket

Szakadásos vizsgálat

- Akkor lehetséges, ha létezik egy világos, számszerű jogosultsági kritérium vagy küszöb
- Az éppen csak a küszöb felettieket veti össze azokkal, akik éppen csak alatta vannak



Feltevések

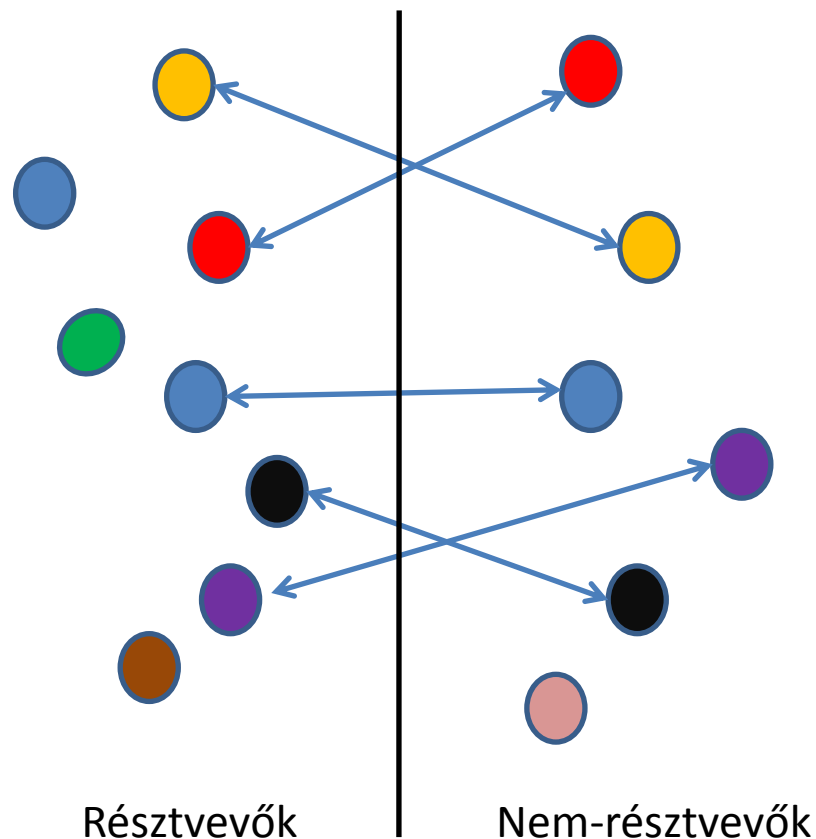
- A beavatkozás során szigorúan veszik a küszöbértéket
- A küszöb közelében lévők nem képesek kijátszani a jogosultsági szabályt és manipulálni azt a tulajdonságot, amin a szabály alapul
- A közvetlenül a küszöb alatt és felett lévők nem különböznek szignifikánsan egymástól

Gyengeségek

- Csak a jogosultsági küszöbérték közelében lévő egyénekre (településekre, stb.) kifejtett hatást képes mérni

Statisztikai párosítás

- A viszonyítási csoportot úgy hozzuk létre, hogy minden résztvevőt egy-egy hozzá hasonló nem-résztvevőhöz párosítjuk a megfigyelhető tulajdonságaik alapján



Feltevések

- Minden lényeges tulajdonság hatását ki tudjuk szűrni

Gyengeségek

- A gyakorlatban nem tudjuk a nem/nehezen mérhető tulajdonságok (pl. motiváció, önbizalom stb.) okozta torzítást kiszűrni
- Emiatt a csoportok valójában nem összehasonlíthatók

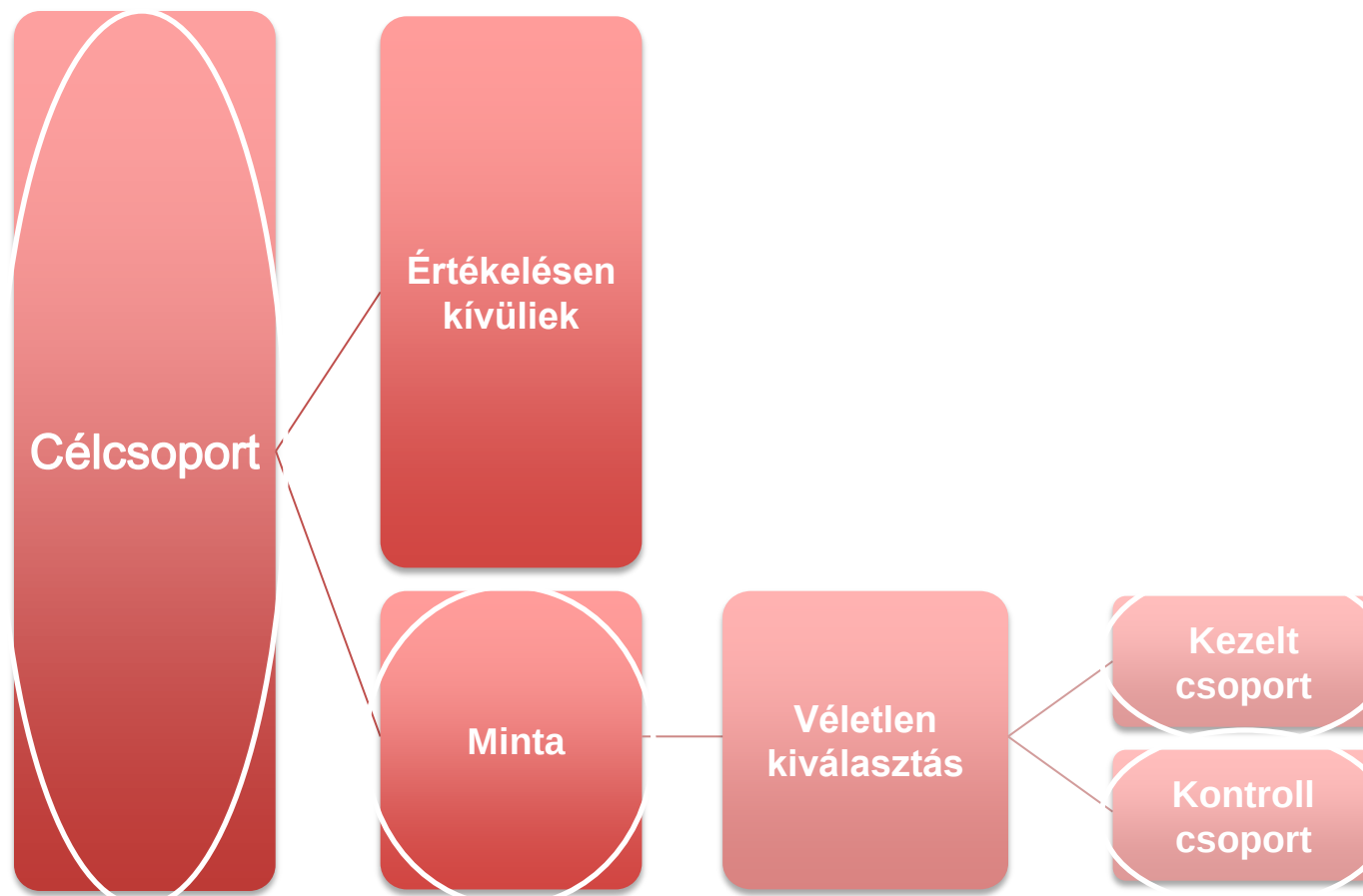
Tartalom

1. Mi a társadalompolitikai kísérlet?
2. Különböző hatásvizsgálati módszerek
3. Randomizált kísérletek (RKK)
4. Következtetések

Hosszú történeti előzmények

- Kísérleti pszichológia: 19. század vége
- Oktatás: 20. század eleje
- Kísérleti szociológia, 20. század eleje:
 - falusi egészség programok, önkormányzati lakásprogramok társadalmi hatásai
- Nagyméretű randomizált klinikai vizsgálatok: az 1962-es amerikai gyógyszerbiztonsági törvények óta általános eljárásnak számítanak
- Jelentős viták övezték, de **ma már széles körben elfogadott**
- USA: fellendülés a `60-as években (250 RKK)
- 1990-es évek: a J-PAL bevezeti az RKK-kat a fejlődésgazdaságban (ma: 450 RKK)
- 2000-es évek: AIDS megelőzés (Bg), álláskeresés ellenőrzése, (Hu), Mikrohitel (Bosznia), gyereknevelési készségek (Sk)

A randomizált értékelések szerkezete



Orr (1999) alapján

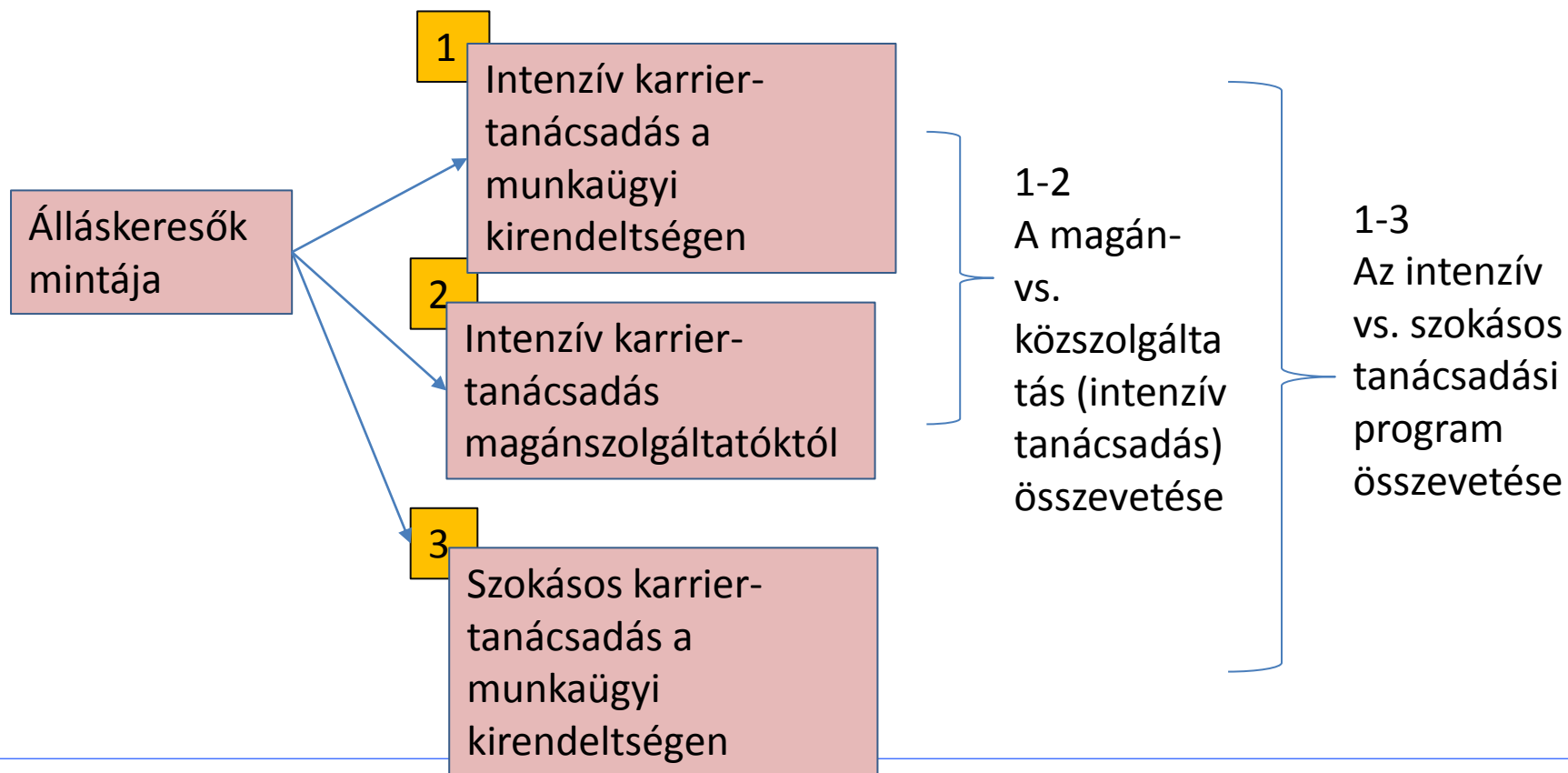
- **A hatásértékelés hitelessége közvetlenül függ** attól, mennyire hasonlít egymásra a program és a kontroll csoport mind a megfigyelhető, mind a nem megfigyelhető tulajdonságokban
- Egy kellően nagy mintából vett megfigyelési egységek **véletlenszerű beosztása** program és kontroll csoportba
 - biztosítja a csoportok statisztikai hasonlóságát
 - mind a megfigyelhető, mind a nem megfigyelhető tulajdonságok szempontjából.
- Ezért ennyire megbízható az RKK

Az RKK előnyei

- Számos szakpolitikai területen lehet beavatkozások hatását tesztelni
- Robosztus és transzparens eredményeket nyújt, amelyeket egyszerű értelmezni és kommunikálni
- Segít megérteni az adott szakpolitika mögött húzódó közgazdasági mechanizmusokat
- Korlátozott erőforrások esetén igazságosan osztja el a hozzáférést
- Meggyőző bizonyítékot ad a hatékony beavatkozások hatásáról

Javítja a szakpolitika-alkotást

Egy RKK többféle kezelést is tartalmazhat, hogy minden egyes programvariáció hatását mérni tudjuk



Tartalom

1. Mi a társadalompolitikai kísérlet?
2. Különböző hatásvizsgálati módszerek
3. Randomizált kísérletek
4. Következtetések

Következtetések

- A társadalompolitikai kísérletek a társadalompolitikai programok javításának **fontos eszközei**
- Számos módszerrel értékelhetjük a szakpolitikákat:
 - Nem-randomizált értékelési módszerek esetén óvatosan kell értelmezni az eredményeket, **figyelembe véve a kiinduló feltevéseket**
 - Ha megvalósítható, az RKK a **legmegbízhatóbb hatásvizsgálati módszer**, és segít a beavatkozás mögött húzódó **hatásmechanizmusok megértésében is**, megkönnyítve a hatékony programok és szakpolitikák kiterjesztését és a **tévhitek megcáfolását**

Következtetések

- Az **értékelésre fordítható erőforrásokat** a nagy hatású, stratégiai szempontból fontos vagy innovatív beavatkozások mérésére érdemes felhasználni:
 - elősegítse a hatékony programok országos kiterjesztését
 - támogassa a jó programok megismétlését más kontextusban is
 - szolgáljon tanulsággul a későbbi szakpolitika-alkotás számára
- De: nem minden beavatkozást lehet értékelni
- És nem is minden beavatkozásnál szükséges hatásértékelés
- Az **értékelések minősége** a társadalomtudományi programok esetében is érje el azt a szintet, amit az orvostudomány megkövetel