

“Nem-paraméteres” eljárások

Adatfájlok

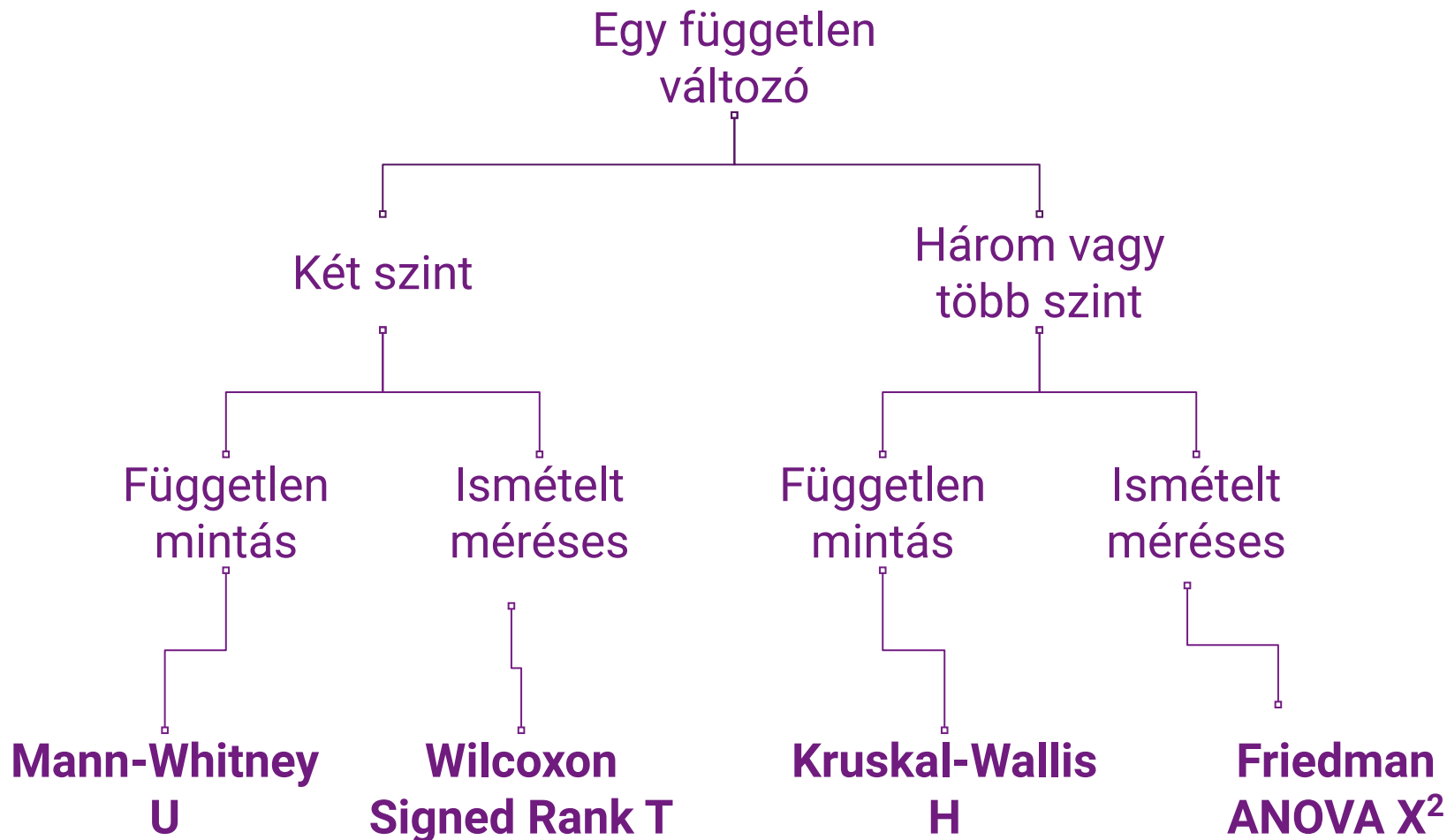
- stereotyping.csv
- coffeeDriving.csv
- discomfortLight.csv
- creativityReward.csv
- DownloadFestival.csv
- Coulrophobia.csv

A “nem-paraméteres” statisztikák

A függő változó valós értékei helyett az értékek rangsorban elfoglalt helyére (legkisebb, második legkisebb, ... legnagyobb) épülnek.

- Az adatok nem normál eloszlásúak
 - vagy nem is folytonosak (tehát lehet ordinális)
- A statisztikák nem ismert paraméterű (átlagú és szórású) eloszlásra épülnek
- Átlag és szórás helyett az adatokat gyakran a mediánnal vagy a rangok átlagával jellemezzük
 - folytonos skála adatok esetén néha még nem-normál eloszlás esetén is az átlag a leghasznosabb leíróstatisztika
- Szemléltető ábra lehet a dobozdiagram

Melyik teszt?



Mann-Whitney U teszt

A független mintás t-próba nem-paraméteres megfelelője

Lépések:

1. U statisztika kiszámítása a pontszámok rangsorban elfoglalt helyéből
2. a kapott U -hoz tartozó p érték megbecsülése

Annak a valószínűségét mutatja, hogy két különböző minta ugyanabból a populációból jön:

Pontosabban a két minta rangátlagainak a valószínűségét abban az esetben, ha a két minta ugyanabból a populációból jön

Az U képlete (Mann-Whitney próba)

1. A két csoport pontszámait összevonjuk.
2. A pontszámokat sorba rendezzük a legkisebبتől a legnagyobbig.
3. A legkisebb pontszám értéke 1 lesz, a második legkisebbé 2, stb.
4. Azonos pontszámok a rangszámok átlagát kapják (tehát ha a 3., 4. és 5. pontszám azonos, mindhárom értéke 4 lesz.
5. A kapott rangszámokat szétválasztjuk a két csoportra.
6. Összeadjuk a csoport rangszámait, és:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

R: a rangszámok
összege

Példa: Stereotyping (stereotyping.csv)

Gyerekeknek egy történetet kellett értelmezniük, és a válaszaikat pontozták: 0 pont nem jelenek meg nemi sztereotípiák, 100: erős nemi sztereotípiák jellemzik.

Két csoport: dolgozó anya és háztartásbeli anya

Dolgozó anya	Háztartásbeli anya
17	19
32	63
39	78
27	29
58	39
25	59
31	77
	81
	68

Rangsorolás

Dolgozó anya		Háztartásbeli anya	
Pontszám	Rangszám	Pontszám	Rangszám
17	1	19	2
32	7	63	12
39	8.5	78	15
27	4	29	5
58	10	39	8.5
25	3	59	11
31	6	77	14
		81	16
		68	13
	39.5		96.5

ezeket a rangszámösszegeket
illesztjük a képletbe

Másik módszer:
 minden pontszámra
 megszámloljuk, hogy
 a másik csoportban
 hány pontszám
 magasabb. Ha
 azonos pontszám van
 a másik csoportban,
 az 0,5 pontot ér. Az
 U az értékek összege.

Dolgozó anya		Háztartásbeli anya	
Pontszám	Érték	Pontszám	Érték
17	9	19	6
32	7	63	0
39	6.5	78	0
27	8	30	4
58	6	39	1.5
25	8	59	0
30	7	77	0
		81	0
		68	
	51.5		11.5

a kettő közül valamelyik az U

A Mann-Whitney próba p értéke

- Bootstrap módszer: pontos p érték (JASP)
 - Szoftver az adatokat minden lehetséges módon két mintára osztja (az eredeti elemszámokat megtartva), és megadja annak a valószínűségét, hogy a két csoport átlagos rangszáma között legalább akkora a különbség, mint az eredeti mintáink között.
- Becsléses módszer
 - A szoftver a rangszámösszegeket z értékekre konvertálja, és a kapott z érték valószínűségét adja meg
 - A JASP ezt nem csinálja, de más szoftver igen, és ezért cikkekben gyakran z értékeket jelentenek U helyett.

Leíróstatisztikák

- Két csoport mediánja
 - előny: könnyen elérhető (JASP Descriptive Statistics-ból)
 - hátrány: ezek lehetnek egyformák, miközben szignifikáns különbség van a két csoport között
- Két csoport rangszámainak átlaga
 - előny: jól érthető
 - hátrány: JASP nem adja meg
- Hodges-Lehmann estimator: két csoport pontszámai közötti különbség mediánja
 - előny: JASP megadja (Location parameter)
 - hátrány: önmagában nem sokat mond, de hatásméretnek jó

Leíróstatisztikák

Dolgozó anya		Háztartásbeli anya		Különbségek
Pont szám	Rang szám	Pont szám	Rang szám	
17	1	19	2	2, 46, 61, 12, stb
32	7	63	12	-13, 31, 45, stb
39	8.5	78	15	-20, 24, 39, -10, stb
27	4	29	5	
58	10	39	8.5	
25	3	59	11	
31	6	77	14	
		81	16	
		68	13	
	39.5		96.5	

Hodges-Lehmann: összes különbséget sorrendbe rakjuk, ezek mediánja az érték



rangsorszámösszegek.



JASP > T-tests > Independent samples > Tests

Independent Samples T-Test

Independent Samples T-Test

	W	p	Hodges-Lehmann Estimate	95% CI for Hodges-Lehmann Estimate	
				Lower	Upper
Stereotyping	11.500	0.039	-29.000	-47.000	-1.000

Note: Mann-Whitney U test.

Mann-Whitney U
(Néha W-nek hívják,
mert Wilcoxon nevű
matematikus tesztjének
a változata)

Különbségek
mediánja

Eredmények jelentése

Children were told a story and asked questions about it. Their answers were rated on a scale of 0 - 100 based on the extent of gender stereotyping they showed (0: not at all stereotyping, 100: extremely stereotyping). The children were divided into two groups: 7 children had mothers working in full-time jobs and 9 children had mothers not working outside the home. The results showed that children whose mothers work full-time are less likely to show stereotypical behaviour, with the mean rank of this group being 5.57, while the children whose mothers do not work outside the home had a mean rank of 10.78. A Mann-Whitney test revealed that this difference was statistically significant, $U = 11.5$, $p = .03$. The median of the difference between the two groups as estimated by the Hodges-Lehmann method was 29 points. That is, whether the mother works or not seems to make a substantial difference to the gender stereotyping behaviour of their children.

Eredmények jelentése

A gyerekek meghallgattak egy történetet, majd kérdésekre kellett válaszolniuk. A válaszaikat a nemi sztereotípiák megjelenése mentén értékelték egy 101-pontos skálán: a 0 pont azt jelenti, hogy semmilyen nemi sztereotípiát nem jelent meg a válaszokban, 100 pont extrém mértékű nemi sztereotipizálást jelölt. A gyerekek két csoportra oszlottak: 7 gyermek édesanyja dolgozott, míg 9 gyermek édesanyja háztartásbeli volt. Az eredmények szerint a dolgozó anyák gyermekeire kevésbé jellemző a nemi sztereotipizálás (átlagos rang: 5,57) mint a háztartásbeli anyák gyermekeire (átlagos rang: 10,78). Egy Mann-Whitney próba szerint a két csoport közötti különbség szignifikáns, $U = 11,5$; $p = ,03$. A Hodges-Lehmann módszerrel megbecsülve a két csoport közötti különbség mediánja 29 pont. Tehát az hogy az anya dolgozik-e vagy sem igen jelentős hatással van a gyerekek viselkedésére a nemi sztereotípiák használatát illetően.

A Kruskal-Wallis H próba

Az egyszempontos független-mintás ANOVA nem-paraméteres megfelelője.

- folytonos, de nem-normál eloszlású függő változó
- vagy ordinális függő változó
- egy független változó 3 vagy több szinttel

p érték annak a valószínűsége, hogy az adott rangátlagokat kapjuk, ha a minták ugyanabból a populációból jönnek. (p értéket befolyásolja a szabadságfok: $k - 1$)

H számítása

1. A csoportok pontszámait összevonjuk.
2. Rendezzük őket a legkisebبتől a legnagyobbig.
3. A legkisebb pontszám értéke 1 lesz, a második legkisebbé 2, stb.
4. Azonos pontszámok a rangszámok átlagát kapják (tehát ha a 3., 4. és 5. pontszám azonos, mindhárom értéke 4 lesz.
5. A kapott rangszámokat szétválasztjuk a két csoportra.
6. Összeadjuk a csoport rangszámait, és:

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$$

R: rangok összege

N: összes pontszám száma

n: egy adott csoport pontszámainak a száma

Utólagos összehasonlítások

- Páros összehasonlítások
 - Páronkénti Mann-Whitney próbák
 - A hibák elkerülése érdekében minél kevesebb összehasonlításra érdemes törekedni
 - Lehet kézzel Bonferroni korrekciót alkalmazni:
korigált p érték: az eredeti p értéket megszorozzuk az összehasonlítások számával

Példa: a koffein hatása a figyelemre (coffeeDriving.csv)

- Javítja-e a koffein a figyelemirányítást
- Három csoport:
 - kávé
 - koffeinmentes kávé
 - víz
- autóvezetés szimulátor (magasabb pontszám = jobb vezetés)

JASP > ANOVA > ANOVA > Nonparametrics


Kruskal-Wallis Test

Factor	Statistic	df	p
Drink	12.788	2	0.002

Kruskal-Wallis H



Csoportok száma - 1
(k - 1), szabadságfok



JASP: szűrővel kell kizárni az összehasonlításból kihagyandó csoportokat, és azután Mann-Whitney próba

Kávé és víz ▼

Independent Samples T-Test

	W	p	Hodges-Lehmann Estimate	95% CI for Hodges-Lehmann Estimate	
				Lower	Upper
DrivingScore	36.000	0.005	7.000	4.000	8.000

Note. Mann-Whitney U test.

Mentes és víz ▼

Independent Samples T-Test

	W	p	Hodges-Lehmann Estimate	95% CI for Hodges-Lehmann Estimate	
				Lower	Upper
DrivingScore	35.500	0.006	5.000	2.000	6.000

Note. Mann-Whitney U test.

Kávé és mentes ▼

Independent Samples T-Test

	W	p	Hodges-Lehmann Estimate	95% CI for Hodges-Lehmann Estimate	
				Lower	Upper
DrivingScore	29.500	0.075	2.000	-2.491e-5	5.000

Note. Mann-Whitney U test.

Eredmények jelentése

The study looks at the effects of caffeine on attentional control and driving performance. Participants' driving performance was measured using a driving simulator after drinking coffee (N = 6), decaffeinated coffee (N = 6) or water (N = 6). The results of a Kruskal-Wallis test showed that participants' driving performance was significantly affected by the type of drink given to them, $H(2) = 12.79$, $p = .002$. Follow-up Mann-Whitney tests showed that people who received either coffee (Median = 10.5) or a decaffeinated drink (Median = 8.5) drove significantly better than the group who drank water (Median = 3.5) ($p < .01$ in both cases). There was no significant difference between the coffee and the decaf conditions ($p = .07$).

Eredmények jelentése

A kutatás a koffein figyelemre és autóvezetésre gyakorolt hatását vizsgálja. A kísérleti személyek kávé, koffeinmentes kávé vagy vizet ittak (N = csoportonként 6 fő), majd egy szimulátor segítségével lemérték a vezetési készségeiket. Egy Kruskal-Wallis próba eredménye szerint a vezetés minőségét szignifikánsan befolyásolta az ital, $H(2) = 12,79$; $p = ,002$. Utólagos Mann-Whitney próbák azt jelezték, hogy a kávé fogyasztó csoport (Medián = 10,5) és a koffeinmentes italt fogyasztó csoport (Medián = 8,5) szignifikánsan jobban vezetett, mint a víz fogyasztó csoport (Medián = 3,5) ($p < ,01$ mindkét esetben). A kávé és a koffeinmentes ital hatása között nem volt szignifikáns különbség ($p = ,07$).

Wilcoxon signed rank T próba: (néha W -t használnak T helyett)

A páros mintás t -próba nem-paraméteres megfelelője

Lépések:

1. T statisztika kiszámítása a pontszámok rangsorban elfoglalt helyéből
2. a kapott T -hez tartozó p érték megbecsülése

Annak a valószínűségét mutatja, hogy két különböző kondíció ugyanabból a populációból jön:

Pontosabban a két kondíció rangátlagainak a valószínűségét abban az esetben, ha a két kondíció ugyanabból a populációból jön

T kiszámítása

1. Kiszámoljuk minden pár pontszám közötti különbséget.
2. A különbségek abszolút értékét sorba rendezzük.
(Az azonos méretű különbségek a rangjaik átlagát kapják)
3. Összeadjuk a pozitív előjelű különbségekhez tartozó rangszámokat, és külön a negatív előjelű különbségekhez tartozó rangszámokat:

Az egyik a T, nincs képlet :-)

Példa:

discomfortLight.csv

A ksz-eket beültették egy szobába, és 1-50 skálán megbecsülték, hogy mennyire érzik kényelmetlenül magukat.

Majd leoltották a szobában a villanyt, és ismét megbecsülték, hogy mennyire érzik kényelmetlenül magukat.

Világos szoba	Sötét szoba
23	33
14	22
35	38
26	30
28	31
19	17
42	42
30	25
26	34
31	24
18	21
25	46
23	29
31	40
30	41

Wilcoxon *signed-rank* próba:

Világos	Sötét	Különbség	Különbség rangszáma
23	33	10	12
14	22	8	9.5
35	38	3	3
26	30	4	5
28	31	3	3
19	17	-2	1
42	42	0	
30	25	-5	6
26	34	8	9.5
31	24	-7	8
18	21	3	3
25	46	21	14
23	29	6	7
31	40	9	11
30	41	11	13

$$T = 1 + 6 + 8 = 15$$

vagy

$$T = 12 + 9,5 + 3 + 5 + 3 + 9,5 + 3 + 14 + 7 + 11 + 13 = 90$$

A Wilcoxon signed rank próba p értéke

- Bootstrap módszer: pontos p érték (JASP)
 - Szoftver az adatokat minden lehetséges módon kétfelé osztja és megadja annak a valószínűségét, hogy a pozitív és negatív előjelű különbségek között legalább akkora a különbség, mint az eredeti adatokban.
- Becsléses módszer
 - A szoftver a rangszámösszegeket z értékekre konvertálja, és a kapott z érték valószínűségét adja meg
 - A JASP ezt nem csinálja, de más szoftver igen, és ezért cikkekben gyakran z értékeket jelentenek T helyett.

Leíróstatisztikák

- Két kondíció mediánja
 - előny: könnyen elérhető (JASP Descriptive Statistics-ból)
 - hátrány: ezek lehetnek egyformák, miközben szignifikáns különbség van a két kondíció között
- Két kondíció rangszámainak átlaga
 - előny: jól érthető
 - hátrány: JASP nem adja meg
- Hodges-Lehmann estimator: két kondíció pontszámai közötti különbség mediánja
 - előny: JASP megadja (Location parameter)
 - hátrány: önmagában nem sokat mond, de hatásméretnek jó

JASP > T-tests > Paired samples T-tests > Tests

Paired Samples T-Test ▼

Paired Samples T-Test

						95% CI for Hodges-Lehmann Estimate	
		W	p	Hodges-Lehmann Estimate	Lower	Upper	
BrightRoom	-	DarkRoom	15.000	0.020	-5.500	-9.000	-0.500

Note. Wilcoxon signed-rank test.

Wilcoxon T
(Néha ezt is W-nek hívják, hogy érdekesebb legyen)

Különbségek mediánja

Eredmények jelentése

Thirty participants' discomfort was rated on a scale of 1 (no discomfort) to 15 (extreme discomfort) in a brightly lit room and in a dark room. A Wilcoxon Signed Ranks test revealed that people experienced significantly more discomfort in a dark room (Mdn = 31) than in a brightly lit room (Mdn = 26), $T = 15$, $p = .02$. Hodges-Lehmann median of the differences was 5.5 points, indicating that darkness had a substantial negative effect on participants' well-being.

Eredmények jelentése

Harminc ember kényelmetlenség érzetét becsülték meg 1-től 15-ig terjedő skálán (nagyobb érték: több kényelmetlenség) amint egy jól megvilágított, majd egy sötét szobában várakoztak. A Wilcoxon Signed Ranks próba alapján elmondhatjuk, hogy a kísérleti személyek szignifikánsan kényelmetlenebbül érezték magukat a sötét szobában (Mdn = 31) mint a világos szobában (Mdn = 26), $T = 15$, $p = ,02$. A különbségek Hodges-Lehmann mediánja 5,5 pont volt, ami arra utal, hogy a sötétség jelentős hatással volt a kísérleti személyek jólétére.

Friedman ANOVA (X_F^2)

Az egyszempontos ismételt méréses ANOVA nem-paraméteres megfelelője.

- folytonos, de nem-normál eloszlású függő változó
- vagy ordinális függő változó
- egy független változó 3 vagy több szinttel

p érték annak a valószínűsége, hogy az adott rangátlagokat kapjuk, ha a kondíciók ugyanabból a populációból jönnek. (p értéket befolyásolja a szabadságfok: $k - 1$)

Kiszámítása

- Vegyük a kísérleti személyeket egyenként
- Rendezzük egy adott személy pontszámait a legkisebبتől a legnagyobbig
- Adjuk össze a rangpontszámokat
- Ezt ismételjük meg minden személyre
- Majd:

$$X_F^2 = \left[\frac{12}{Nk(k-1)} \sum R_i^2 \right] - 3N(k+1)$$

Példa: Jutalom és kreativitás

(creativityReward.csv)

A jutalom kreativitásra gyakorolt hatását vizsgáljuk. A feladat: kollázs készítése adott anyagokból. Művészek értékelik az alkotás kreativitását 15 pontos skálán (15: nagyon kreatív). A kísérleti személyek háromszor végzik el a feladatot: egyszer szórakozásból, egyszer 10 dollárért és egyszer 100 dollárért.

JASP > ANOVA > Repeated Measures ANOVA > Nonparametrics

Ez most minket nem érdekel. Értékelők közötti egyetértést mutatná (0 és 1 között)

Friedman Test

Factor	Chi-Squared	df	p	Kendall's W
Jutalom	17.897	2	< .001	0.258

Friedman Khi négyzet

Kondíciók száma - 1
(k - 1), szabadságfok

JASP > ANOVA > Repeated Measures ANOVA > Nonparametrics: Conover's Post Hoc Tests

Conover's Post Hoc Tests

Conover's Post Hoc Comparisons - Jutalom

		T-Stat	df	W_i	W_j	p	P_{bonf}	P_{holm}
Semmi	10 dollár	1.587	18	18.500	11.500	0.130	0.389	0.130
	100 dollár	2.608	18	18.500	30.000	0.018	0.053	0.036
10 dollár	100 dollár	4.195	18	11.500	30.000	< .001	0.002	0.002

Párok

A
statisztika

kondíciók
rangszámainak az
összegei.
Ha ezt osztjuk az
elemszámmal,
megvan az átlagos
rang
leíróstatisztika.

p korrekció nélkül, és
Bonferroni és Holm
korrekcióval

Eredmények jelentése

[...] The results of a Friedman's ANOVA showed that rewards significantly affected creativity, $X_F^2(2) = 17.89, p < .001$. Post-hoc pairwise comparisons with Holm correction revealed that people were significantly more creative when given a large reward (M Rank = 3.0) than when either given no reward (M Rank = 1.85) ($p = .04$) or a small reward (M Rank = 1.15) ($p = .02$). However, there was no difference in creativity between the small reward and the no reward conditions.

Eredmények jelentése

[...] A Friedman ANOVA kimutatta, hogy a jutalom szignifikáns hatással volt a kreativitásra, $X_F^2(2) = 17,89$; $p < ,001$. Holm módszerrel korigált páros összehasonlítások szerint amikor a kísérleti személyek nagyobb jutalmat kaptak, szignifikánsan kreatívabb alkotásokat hoztak létre (M Rang = 3,0), mint akkor, amikor nem kaptak jutalmat (M Rang = 1,85) ($p = ,04$) vagy csak kisebb jutalmat kaptak (M Rang = 1,15) ($p = ,02$). A 10 dolláros jutalom azonban nem növelte a kreativitást a jutalom nélküli kondícióhoz képest.