

Exploratorní faktorová analýza

Short Dark Triad (SD3)

Dataset získán z: <https://www.kaggle.com/datasets/lucasgreenwell/short-dark-triad-responses?select=data.csv>

Úvod

Koncept Temné triády zahrnující machiavelismus, narcismus a psychopatii, představuje v současné osobnostní a sociální psychologii jeden z hojně zkoumaných konstruktů. Tyto tři sociálně averzivní rysy sdílejí společné jádro v podobě nízké míry přívětivosti i empatie, ale jsou považovány za vzájemně oddělitelné dimenze (Furnham et al., 2013). S rostoucím zájmem o studium subklinických projevů těchto vlastností v běžné populaci vyvstala potřeba jejich efektivního a validního měření, což vedlo k vytvoření metody Short Dark Triad (SD3) (Jones & Paulhus, 2014).

Obecně nám exploratorní faktorová analýza (EFA) pomáhá přenést pozornost od jednotlivých projevů chování k jejich společným jmenovatelům a statisticky vyjádřit míru, do jaké jsou tyto konkrétní projevy syceny latentními proměnnými, které nazýváme faktory. V praxi nám tento postup umožňuje redukovat velké množství dat na několik klíčových dimenzí, aniž bychom ztratili podstatnou informaci o jejich vnitřní struktuře.

Cílem této práce je pomocí exploratorní faktorové analýzy (EFA) ověřit vnitřní strukturu dotazníku SD3. Budeme zkoumat, zda se jednotlivé položky dotazníku skutečně seskupují do tří teoreticky předpokládaných latentních faktorů a jaká je míra jejich vzájemné provázanosti.

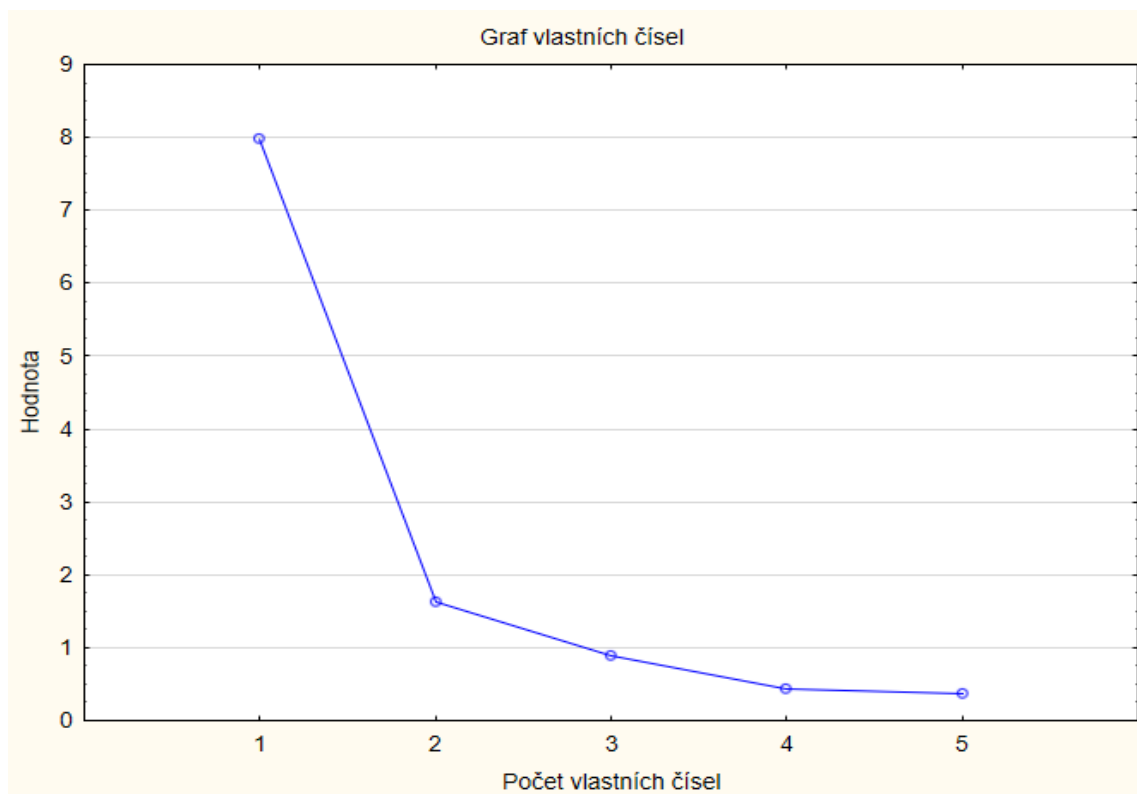
Soubor

Analýza vychází z rozsáhlého datasetu získaného z platformy Kaggle. Soubor obsahuje odpovědi celkem 18 192 respondentů, kteří dobrovolně vyplnili online verzi dotazníku **Short Dark Triad**.

Dotazník SD3 se skládá z **27 položek**, které jsou rozděleny do tří subškál: narcismus (N), machiavelismus (M) a psychopatie (P), z nichž každá obsahuje 9 položek. Respondenti vyjadřovali míru svého souhlasu s jednotlivými tvrzeními na **pětibodové Likertově škále** (1 = nesouhlasím, 5 = souhlasím).

Předpoklady k analýze EFA

Graf 1: Sutinový graf vlastních čísel



Na začátku analýzy byl sestaven **sutinový graf vlastních čísel**, který slouží k určení optimálního počtu extrahovaných faktorů. Dvě používaná kritéria k výběru počtu faktorů jsou:

Kaiserovo kritérium – vybíráme faktory, které mají na svislé ose hodnotu vyšší než **1,0**.

Pravidlo schodu – hledáme bod, kde se strmý pád grafu „zlomí“ a začne být vodorovný.

V našem grafu je patrný strmý pokles mezi prvním a druhým faktorem. Další patrný „schodek“ se v datech dle mého názoru nevyskytuje. Hodnoty vlastního čísla vyššího než 1 dosahují pouze první dva faktory.

Z tohoto pohledu je zajímavé popřemýšlet, proč tedy původní autoři zvolili třífaktorový model namísto jedno či dvoufaktorového, na základě výše zmíněných pravidel. To se možná dozvíme sami v následující analýze.

Faktorová zátěž

Zatímco předchozí graf vlastních čísel nám napověděl, *kolik* skupin (faktorů) se v datech skrývá, analýza faktorových zátěží nám říká, co tyto skupiny obsahují.

Faktorová zátěž je v podstatě korelace mezi konkrétní otázkou v dotazníku a daným latentním faktorem. Nabývá hodnot od -1 do 1. Pro snadnou interpretaci hledáme tzv. **jednoduchou strukturu**:

Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psycholegie.cz/stat4?i=705>

- Každá otázka by měla mít vysokou zátěž (ideálně > 0,40) pouze v jednom faktoru.
- V ostatních faktorech by měla mít zátěž nízkou (blízkou nule).

Pokud má otázka podobně vysokou **zátěž ve více faktorech**, není dostatečně specifická a měří více věcí naráz. To je pro čistotu psychologického testu problematické.

Proč provádíme rotaci faktorových zátěží? Nerotované řešení bývá nepřehledné, protože první faktor se snaží vysvětlit všechno. Normalizovaná rotace matematicky „pootočí“ osy grafu tak, aby se rozdíl mezi faktory maximálně vyostřily. Představte si to jako zaostření objektivu u fotoaparátu – struktura dat zůstává stejná, ale my ji vidíme mnohem jasněji.

U třífaktorového modelu analýza faktorových zátěží po normalizované rotaci ukázala, proč graf vlastních čísel primárně doporučoval dvoufaktorové řešení. Zatímco **machiavelismus** (faktor 1) a **narcismus** (faktor 2) vykazují jasnou a silnou strukturu, položky **psychopatie** (faktor 3) jsou v datech slabší a vykazují značný překryv s prvním faktorem. Přesto vynucené třífaktorové řešení dovoluje identifikovat psychopatii jako samostatný, byť méně robustní konstrukt.

Pro názornost jsem se rozhodla přiložit i ukázkou dvoufaktorového řešení, které je podle Kaiserova kritéria optimálnější. Můžeme vidět, že se se machiavelismus pohltit faktor psychopatie pod jeden antisociální faktor, kdežto narcismus se udržel jako samostatný egocentrický faktor.

Tabulka 1: faktorové zátěže třífaktorového a dvoufaktorového modelu

Proměnné	Třífaktorový model			Proměnné	Dvoufaktorový model			
	Machiavelismus	Narcismus	Psychopatie		Antisociální faktor	Egocentrický faktor		
M1	0,5201	0,0207	0,0344	M1	0,5176	0,0157		
M2	0,6604	0,2839	0,2077	M2	0,6767	0,3206		
M3	0,5968	0,3749	-0,0163	M3	0,5644	0,3242		
M4	0,4742	0,1362	-0,2262	M4	0,4083	0,0394		
M5	0,7599	0,1617	0,1846	M5	0,7758	0,1944		
M6	0,7209	0,0993	0,1952	M6	0,7417	0,1412		
M7	0,5525	0,0349	0,0155	M7	0,5451	0,0215		
M8	0,6005	0,2199	0,1827	M8	0,6155	0,2539		
M9	0,5675	0,2323	0,1623	M9	0,5793	0,2584		
N1	0,1163	0,5446	0,1012	N1	0,1132	0,5385		
N2	0,0739	-0,5594	-0,1257	N2	0,0718	-0,5704		
N3	0,2667	0,5551	0,2079	N3	0,2777	0,5869		
N4	0,2272	0,5433	0,0918	N4	0,2198	0,5308		
N5	0,4227	0,5259	-0,0582	N5	0,3849	0,4445		
N6	-0,0447	-0,4901	-0,1518	N6	-0,0517	-0,5128		
N7	0,2003	0,5049	0,1437	N7	0,2035	0,5172		
N8	-0,1616	-0,4579	-0,2149	N8	-0,1793	-0,5021		
N9	0,4060	0,4048	0,1922	N9	0,4182	0,4361		
P1	0,5092	0,1464	0,4397	P1	0,5666	0,2818		
P2	-0,0735	-0,2755	-0,4708	P2	-0,1515	-0,4091		
P3	0,4845	0,1082	0,3184	P3	0,5262	0,2034		
P4	0,2336	0,1705	0,4689	P4	0,3051	0,3147		
P5	0,5059	0,1052	0,3336	P5	0,5502	0,2049		
P6	0,5428	0,2683	0,3730	P6	0,5885	0,3695		
P7	-0,0087	-0,1227	-0,4189	P7	-0,0825	-0,2549		
P8	0,2600	0,2911	0,3215	P8	0,3024	0,3783		
P9	0,6037	0,3163	0,2819	P9	0,6326	0,3797		
Vysvětleno podíl	5,4629	3,1587	1,7807	rozptylu.	Vysvětleno podíl	5,7790	3,7288	rozptylu.
	0,2023	0,1170	0,0660			0,2140	0,1381	

Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=705>

Vysvětlený rozptyl

Vysvětlený rozptyl udává, jak velkou část celkové informace obsažené v dotazníku (všech odpovědí respondentů) se nám podařilo úspěšně zařadit a popsat pomocí zvolených faktorů. V praxi nám jeho procentuální hodnota říká, do jaké míry naše faktory věrně reprezentují realitu.

Při srovnání obou řešení vidíme, že **dvoufaktorový model** vysvětluje celkem **35,2 %** rozptylu dat. Antisociální faktor (směs machiavelismu a psychopatie) je s 21,4 % velmi dominantní. Egocentrismus (narcismus) vysvětluje zbylých 13,8 % rozptylu.

Pokud vynutíme **třífaktorové řešení**, celkový vysvětlený rozptyl vzroste na **38,5 %**. Vidíme tedy, že přidáním třetího faktoru (Psychopatie) získáme pouze **3,3 %** nových informací (nárůst z 35,2 na 38,5). Tento malý statistický fakt je v souladu s naší hypotézou z grafu vlastních čísel, že třetí faktor je v tomto souboru slabý. Jeho položky mají tendenci splývat s prvním faktorem.

Závěr

Exploratorní faktorová analýza SD3 jasně ukázala, že ačkoliv autoři deklarují 3 oddělené rysy, v reálných datech dochází k výraznému překryvu mezi **machiavelismem** a **psychopatií**. Tyto dvě dimenze se v našem souboru projevily jako jeden společný „antisociální“ faktor, který tvoří dominantní složku dotazníku. Naproti tomu **narcismus** se ukázal jako samostatný a vnitřně konzistentní rys, který si udržuje svou identitu i při snížení počtu faktorů.

Autoři dotazníku SD3 zvolili třífaktorový model především proto, aby zachovali tradiční teoretické rozdělení Temné triády na machiavelismus, narcismus a psychopatii. Zároveň uvádějí, že **Velicerova MAP analýza** v jejich původních studiích jednoznačně naznačila existenci tří faktorů, což podpořilo požadavek na zachování samostatnosti každého rysu. (Jones & Paulhus, 2014).

Volba počtu faktorů není pouze mechanickým procesem založeným na Kaiserově kritériu, ale vyžaduje citlivou interpretaci. Přestože třífaktorové řešení vysvětluje o 3,3 % více rozptylu, statistická slabost třetího faktoru naznačuje, že hranice mezi psychopatií a machiavelismem je v běžné populaci velmi tenká. Pro výzkumnou praxi z toho vyplývá, že musíme brát v úvahu vysokou vzájemnou provázanost složek Temné triády, která může ovlivnit přesnost diagnostiky jednotlivých rysů.

Zdroje

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of Personality: A 10 Year Review. *Social and Personality Psychology Compass*, 7(3), 199–216.

<https://doi.org/10.1111/spc3.12018>

Jones, D., & Paulhus, D. (2014). Introducing the Short Dark Triad (SD3). *Assessment*, 21, 28–41.

<https://doi.org/10.1177/1073191113514105>