

Sisällysluettelo

ALKUSANAT	4
ALKUSANAT E-KIRJA –VERSIOON	5
SISÄLLYSLUETTELO	6
LYHYT SANASTO VASTA-ALKAJILLE	7
1. MONIMUUTTUJAMENETELMÄT IHMISTIEETEISSÄ	9
1.1 MONIMUUTTUJA-AINEISTON ERITYISPIIRTEITÄ	10
1.2 AINEISTON ALUSTAVA TARKASTELU	13
1.2.1 Korrelaatio ja käyräviivainen yhteys	13
1.2.2 Outlierit	14
1.2.3 Normaalisuus	16
1.2.4 Multikollinearisuus ja singulaarisuus	17
1.3 KIRJAN RAKENTEESTA	18
2. FAKTORIANALYYSI	19
2.1 PÄÄKOMPONENTTIANALYYSI (PCA)	21
2.1.1 Missä tilanteessa toimii parhaiten	21
2.1.2 Rajoitukset ja oletukset	21
2.1.3 Lyhyesti teoriasta ja käsitteistä	21
2.1.4 Lisätestit ja jatkoanalyysit	23
2.1.5 Tekninen suoritus SPSS-ohjelmistolla ja tulkinta	24
2.2 PERINTEINEN EKSPLORATIIVINEN FAKTORIANALYYSI (EFA)	32
2.2.1 Missä tilanteessa toimii parhaiten	32
2.2.2 Rajoitukset ja oletukset	32
2.2.3 Lyhyesti teoriasta ja käsitteistä	32
2.2.4 Lisätestit ja jatkoanalyysit	34
2.2.5 Tekninen suoritus SPSS-ohjelmistolla ja tulkinta	35
2.3 KONFIRMATORINEN FAKTORIANALYYSI (CFA)	46
2.3.1 Missä tilanteessa toimii parhaiten	47
2.3.2 Rajoitukset ja oletukset	47
2.3.3 Lyhyesti teoriasta ja käsitteistä	48
2.3.4 Lisätestit ja jatkoanalyysit	56
2.3.5 Tekninen suoritus SPSS-ohjelmistolla ja tulkinta	57
3. LOPUKSI	64
LIITE A. AINEISTOSSA KÄYTETYT ALKUPERÄISET MUUTTUJAT	65
LIITE B. MATRIISILASKENNASTA KEVYESTI	68
LÄHTEET	71
ASIA- JA HENKILÖHAKEMISTO	73

Leskinen ja Kuusinen totesivat (1991), että yleensä ilmiöt eivät ole sellaisia, että pitäisi olettaa dimensioiden olevan toisistaan riippumattomia. Pääkomponenttien suorakulmaisuus nimittäin tarkoittaa sitä, että komponentit eivät korreloi keskenään ja ovat näin ollen riippumattomia toisistaan. Jos tutkittavana ilmiönä on esimerkiksi asenne ja asenteen eri komponentit, mikään teoreettinen tai käytännöllinen syy ei edellytä, että esimerkiksi asenne palkkaan ei voisi korreloida asenteeseen työtovereita kohtaan. Mikäli sallitaan ajatus siitä, että pääkomponentit voivat korreloida keskenään, on suositeltavaa tehdä rotaatio vinokulmaisesti (kuva 2.2). Tunnetuin vinorotaatio on OBLIMIN-rotatio¹⁰, joka voidaan tehdä vain *maximum likelihood* -tyyppisen ratkaisun etsimisen yhteydessä. Mikäli SPSS-ohjelmistossa haluaa tehdä pääkomponenteille vinokulmaisen rotaation, ohjelmisto valitsee automaattisesti oikean käsittelytavan.

Pääkomponentin hyvyttä voidaan arvioida joko sisällöllisesti tai muuttujien latausten perusteella. Latausten perusteella laskettua komponenttien hyvyyden mittaa kutsutaan **ominaisarvoiksi**. Ominaisarvoa laskettaessa huomioidaan kunkin muuttujan lataus kyseisellä pääkomponentilla. **Perinteinen nyrkkisääntö¹¹ on ollut, että pääkomponentin ominaisarvon olisi hyvä olla vähintään yksi.** Tämä ei tietenkään ole tarkka raja; jos pääkomponentti on helposti tulkittavissa, myös ykköstä pienempi ominaisarvo on sallittu. Myös **muuttujan hyvyttä** voidaan arvioida sen latausten perusteella. Eri pääkomponenteille tulevien yksittäisten muuttujien latausten neliöiden summa on nimeltään **kommunaliteetti**. Mitä voimakkaammin muuttuja latautuu jollekulle pääkomponenteista, sitä lähemmäksi arvoa 1 kommunaliteetti tulee. Tähän ideaan perustuu myös Lauri Tarkkosen kehittämä tarkka reliabiliteettikerroin (Tarkkonen 1987), jossa faktorianalyysin perusteella löydettyjen kommunaliteettien perusteella voidaan laskea tarkasti kyseisen faktorin/pääkomponentin reliabiliteetti. Matalien kommunaliteettien perusteella voidaan osa muuttujista poistaa analyysistä. Mikäli joku tai jotkut muuttujista eivät juurikaan lataudu yhdellekään pääkomponenteista, on sitä tai niitä turha pitää mukana analyyseissa. Kyseiset muuttujat sekoittavat turhaan tulkintoja. Nyrkkisääntönä voi olla, että mikäli muuttujan lataukset eivät yhdelläkään pääkomponentilla ylitä arvoa 0.30, on muuttuja ”sakkaa”; se voidaan poistaa muuttujien joukosta.

Pääkomponentit yleensä tulkitaan ja nimetään rotatoinnin jälkeen. Tulkinnan perusteena käytetään muuttujien latautumista kyseiselle pääkomponentille; pääkomponentti nimetään voimakkaammin latautuneiden muuttujien mukaan. Menettely muistuttaa suuresti laadullisen tutkimuksen sisällön analyysissä tehtävää luokkien nimeämistä. Välttämättä nimeäminen ei ole yksikäsitteistä, mutta ymmärrettävää on, että pääkomponentin nimeksi halutaan tietenkin jotain, mikä mahdollisimman hyvin ja kattavasti kuvaisi parhaiten latautuneiden muuttujien sisältöä.

2.1.4 Lisätestit ja jatkoanalyysit

Ennen pääkomponenttianalyysia saattaa olla hyödyllistä tutkia korrelaatiomatriisin rakennetta, sitä **onko käytettävä korrelaatiomatriisi sovelias pääkomponenttianalyysiin**. Eräs käytetyimmistä menetelmistä on **Bartlettin testi**, joka tutkii hypoteesia, ovatko korre-

¹⁰ OBLIMIN (tai tarkemmin DIRECT OBLIMIN)-rotaatiossa minimoidaan latausten väliset ristitulot.

¹¹ Useiden monimuuttujamenetelmien tulkinta perustuu ns. nyrkkisääntöihin (engl. *Thumb Rules* tai *Quick-and-dirty* -säännöt). Tämä viittaa siihen, että todellisuudessa ei välttämättä ole ”oikeaa” arvoa välille ja huonolle mallille tai testisuurelle. Nyrkkisäännöt saattavat hieman poiketa toisistaan riippuen auktoriteetista.

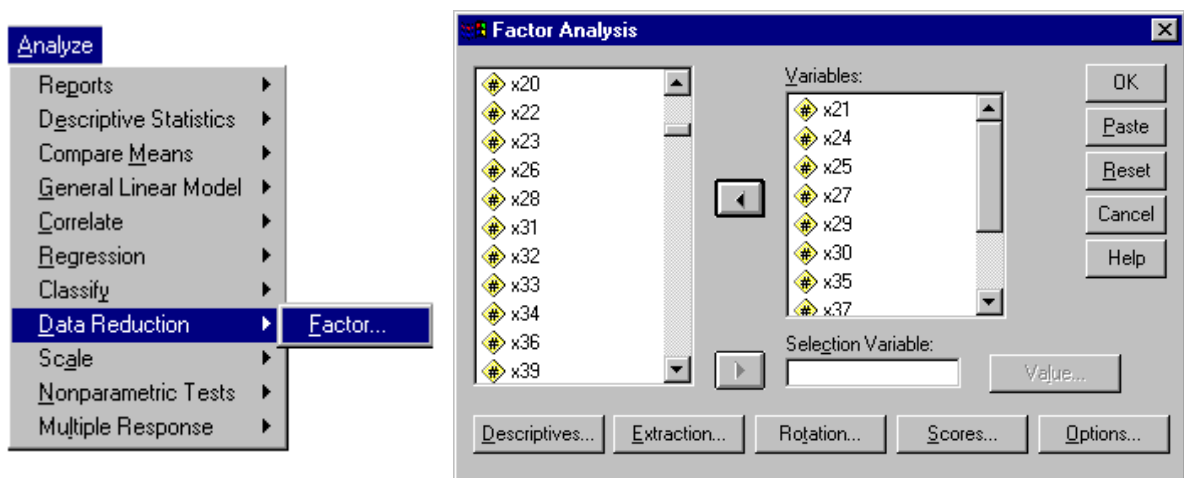
laatiot nolliä. Suurilla otosko'oilla testi kuitenkin liian herkästi antaa signaalin, että korrelaatiot eroavat nolllasta. Muita tapoja ovat korrelaatioiden merkitsevyytestaus, anti-image korrelaatiomatriisi sekä Kaiserin testi otoksen osuvuudelle. **Korrelaatioiden merkitsevyytestauksessa** huomattavan määrän kahden muuttujan välisiä korrelaatioita pitää olla nolllasta poikkeavia. Ongelma tietenkin on sama kuin Bartlettin testissä; suuret otoskoot vaikuttavat tulokseen niin, että mitättömätkin korrelaatiot saattavat olla tilastollisesti merkitseviä. **Anti-image korrelaatiomatriisissa** lasketaan osittaiskorrelaatiomatriisi, jonka tuloksena ns. anti-image -matriisissa lävistäjän ulkopuoliset arvojen tulisi olla pieniä. **Kaiserin testissä** (tai Kaiser-Meyer-Olkin -testi) lasketaan suhde korrelaation ja korrelaatio+osittaiskorrelaation välillä. Mikäli osittaiskorrelaation arvo on pieni, suhde lähestyy arvoa 1. Nyrkkisääntö voi olla esimerkiksi se, että jos Kaiserin testi antaa arvon 0.6 tai suurempi, korrelaatiomatriisi on sovelias pääkomponenttianalyysiin. Osa muuttujista voidaan jo alunperinkin sulkea pois analyysistä, mikäli niillä ei ole riittävästi varianssia tai ne eivät korreloi muiden muuttujien kanssa.

Pääkomponenttianalyysin yhteydessä tai sen jälkeen voidaan kullekin vastaajalle laskea **pistemäärä, joka kuvaa sitä, kuinka voimakkaasti kukin henkilö edustaa kutakin pääkomponenttia**. Itse lasku suoritetaan regressiomallituksen tapaan ennustamalla, milläisen arvon kukin yksilö saisi, jos kutakin pääkomponenttia painotettaisiin. Näitä ns. pääkomponenttipisteitä (*Principal Component Scores*) voidaan sittemmin käyttää jatkoanalyysissä ja tarkastella esimerkiksi eri ryhmien välisiä eroja pääkomponenttien suhteen. Saman asian voi tietenkin tehdä perinteisesti muodostamalla summamuuttujan pääkomponentille voimakkaimmin latautuneista muuttujista. Tällöin kuitenkin kadotetaan se hienojakoinen tieto, jonka saavutimme pääkomponenttianalyysin perusteella muuttujien ja yksilöiden tarkasta latautumisesta ko. pääkomponenteille.

2.1.5 Tekninen suoritus SPSS-ohjelmistolla ja tulkinta

Perusnäky

Pääkomponenttianalyysi SPSS-ympäristössä alkaa valinnoilla *Analyze – Data Reduction – Factor*. Pääkomponenttianalyysi (ja faktorianalyysin) päävalikko näyttää seuraavalta:



Valinnat

Valitsemme pääkomponenttianalyysiin motiivimittarin osioita. Tarkoituksemme on selvittää, löytyykö motiivimittarista erilaisia motiividimensioita kuvaavia muuttujajoukkoja. Tilan ja ajan säästämiseksi valitsemme vain osan 54 motiivimuuttujasta (x20 – x73 ks. liite A). Mukaan tulevat 18 muuttujaa ovat taulukon 2.1 mukaiset. Samoja muuttujia käsitellään myös faktorianalyysin yhteydessä; nyt tehdään suorakulmainen rotaatio ja faktorianalyysin yhteydessä vinokulmainen rotaatio.

Taulukko 2.1 Analyysiin mukaan tulevat muuttujat

Mja	selite	mja	selite
X21	Voin kehittää mielikuvitusta	X40	Siinä saan uusia ystäviä
X24	Siinä voin purkaa energiaa	X41	Siinä saan uusia ideoita
X25	Siitä on hyötyä	X42	Siinä voin voittaa
X27	Voi tulla vahvaksi	X50	Kaverit ovat tuttuja
X29	Siinä voin oppia uutta	X51	On hyvä yhteishenki
X30	Opin hyödyllisiä taitoja	X59	Tulen fyysisesti hyvään kuntoon
X35	On hauskaa kilpailla	X64	Saan hyvää seuraa
X37	Siinä tapaan kavereita	X66	On hyödyllistä osata
X39	Siinä voin luoda	X73	Siinä voin liikkua

Päävalikossa on muuttujien lisäksi valittavana viisi erilaista seikkaa, joilla voidaan vaikuttaa analyysin sisältöön ja laajuuteen: *Descriptives*, *Extraction*, *Rotation*, *Factor Scores* ja *Options*. Oletuksena on, että teemme pääkomponenttianalyysin. Tehdään oletusten lisäksi seuraavat valinnat:

- **Descriptives**

Statistics: Initial solution [Näemme alkuperäiset lataukset.]

Correlation matrix: Coefficient, Significant levels, KMO and Bartlett's test of sphericity

[Näemme korrelaatiomatriisin sekä -kertoimien tilastollisen merkitsevyyden ja saamme käsityksen siitä, soveltuuko korrelaatiomatriisimme pääkomponenttianalyysiin.]

- **Extraction**

Display: Unrotated factor solution, Scree plot

[Näemme rotatoimattoman ratkaisun sekä Cattellin Scree-testin sopivan pääkomponenttien määrän arvioimiseksi. Huomattakoon, että oletuksena on se, että tehdään pääkomponenttianalyysi – faktorianalyysin yhteydessä käsittelemme hieman muita vaihtoehtoja.]

- **Rotation**

Method: Varimax

Display: Loading Plots

[Suorakulmainen rotaatio – oletamme, että motiividimensiot eivät korreloi toistensa kanssa. Tulemme tekemään myös vinokulmaisen rotaation OBLIMIN-valinnalla faktorianalyysin yhteydessä – tulkinnat ovat identtisiä sekä PCA:ssa että EFA:ssa. Saamme analyysin lopuksi graafisen esityksen siitä, miten havainnot sijoittuvat pääkomponenttien muodostamaan avaruuteen.]

- **Scores**

Save as variables

Metsämuuronen: Monimuuttujamenetelmien perusteet SPSS-ympäristössä

[Muodostetaan pääkomponenttien perusteella pääkomponenttipisteet, joiden voimme myöhemmin verrata ja ryhmitellä vastaajia.]

Options

Coefficient Display Format: Sorted by size

Suppress absolute values less than **.30**

[Lopullinen tulos esitetään muodossa, jossa kullakin faktorilla latautuneet muuttujat ovat suuruusjärjestyksessä eikä pienempiä latausarvoja kuin 0.30 näytetä.]

Tulokset ja niiden tulkinta

Pyyntömme mukaisesti tulostuksessa esitetään ensimmäisenä korrelaatiomatriisi:

Correlation Matrix																			
Correlation	X21	X24	X25	X27	X29	X30	X35	X37	X39	X40	X41	X42	X50	X51	X59	X64	X66	X73	
X21	1,000	,150	,108	,031	,297	,203	-,120	-,003	,500	,107	,561	-,106	,002	-,039	-,159	,054	,097	-,168	
X24	,150	1,000	,213	,453	,087	,076	,325	,304	,146	,264	,147	,331	,253	,241	,373	,298	,098	,364	
X25	,108	,213	1,000	,144	,300	,530	,032	,140	,120	,156	,193	,072	,058	,119	,055	,137	,498	,032	
X27	,031	,453	,144	1,000	-,027	,040	,516	,290	-,013	,259	,004	,494	,350	,282	,561	,298	,038	,505	
X29	,297	,087	,300	-,027	1,000	,492	-,132	,062	,248	,114	,335	-,108	-,055	,032	-,158	,028	,331	-,152	
X30	,203	,076	,530	,040	,492	1,000	-,039	,104	,260	,197	,323	-,026	-,024	,108	-,084	,112	,547	-,111	
X35	-,120	,325	,032	,516	-,132	-,039	1,000	,402	-,125	,324	-,078	,788	,446	,422	,659	,401	,010	,655	
X37	-,003	,304	,140	,290	,062	,104	,402	1,000	,079	,715	,113	,400	,592	,622	,403	,654	,158	,408	
X39	,500	,146	,120	-,013	,248	,260	-,125	,079	1,000	,172	,534	-,065	-,018	,022	-,202	,116	,219	-,225	
X40	,107	,264	,156	,259	,114	,197	,324	,715	,172	1,000	,211	,333	,470	,610	,310	,644	,255	,306	
X41	,561	,147	,193	,004	,335	,323	-,078	,113	,534	,211	1,000	-,050	,041	,085	-,121	,174	,299	-,121	
X42	-,105	,331	,072	,494	-,108	-,026	,788	,400	-,065	,333	-,050	1,000	,450	,405	,615	,367	,006	,590	
X50	,002	,253	,058	,350	-,055	-,024	,446	,692	-,018	,470	,041	,450	1,000	,618	,468	,542	,089	,449	
X51	-,039	,241	,119	,282	,032	,108	,422	,622	,022	,610	,085	,405	,618	1,000	,399	,639	,187	,390	
X59	-,159	,373	,055	,561	-,158	-,084	,659	,403	-,202	,310	-,121	,615	,468	,399	1,000	,384	-,038	,766	
X64	,054	,298	,137	,298	,028	,112	,401	,654	,116	,644	,174	,367	,542	,639	,384	1,000	,217	,377	
X66	,097	,098	,498	,038	,351	,547	,010	,158	,219	,255	,299	,006	,089	,187	-,038	,217	1,000	-,010	
X73	-,168	,364	,032	,505	-,152	-,111	,655	,408	-,225	,306	-,121	,590	,449	,390	,766	,377	-,010	1,000	
Sig. (1-tailed)																			
X21		,000	,002	,205	,000	,000	,001	,470	,000	,002	,000	,002	,474	,146	,000	,074	,005	,000	
X24	,000		,000	,000	,010	,021	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,004	,000	
X25	,002	,000		,000	,000	,000	,195	,000	,001	,000	,000	,027	,060	,001	,071	,000	,000	,198	
X27	,205	,000	,000		,237	,140	,000	,000	,366	,000	,452	,000	,000	,000	,000	,000	,156	,000	
X29	,000	,010	,000	,237		,000	,000	,047	,000	,001	,000	,002	,071	,193	,000	,229	,000	,000	
X30	,000	,021	,000	,140	,000		,150	,003	,000	,000	,000	,244	,259	,002	,012	,001	,000	,001	
X35	,001	,000	,195	,000	,000	,150		,000	,000	,000	,018	,000	,000	,000	,000	,000	,397	,000	
X37	,470	,000	,000	,000	,047	,003	,000		,017	,000	,001	,000	,000	,000	,000	,000	,000	,000	
X39	,000	,000	,001	,366	,000	,000	,000	,017		,000	,000	,040	,318	,279	,000	,001	,000	,000	
X40	,002	,000	,000	,000	,001	,000	,000	,000	,000		,000	,000	,000	,000	,000	,000	,000	,000	
X41	,000	,000	,000	,452	,000	,000	,018	,001	,000	,000		,089	,133	,011	,001	,000	,000	,001	
X42	,002	,000	,027	,000	,002	,244	,000	,000	,040	,000	,000	,089		,000	,000	,000	,441	,000	
X50	,474	,000	,060	,000	,071	,259	,000	,000	,318	,000	,133	,000	,133		,000	,000	,008	,000	
X51	,146	,000	,001	,000	,193	,002	,000	,000	,279	,000	,011	,000	,000	,000		,000	,000	,000	
X59	,000	,000	,071	,000	,000	,012	,000	,000	,000	,000	,001	,000	,000	,000	,000		,152	,000	
X64	,074	,000	,000	,000	,000	,229	,001	,000	,001	,000	,000	,000	,000	,000	,000	,000		,000	
X66	,005	,004	,000	,156	,000	,000	,397	,000	,000	,000	,000	,441	,008	,000	,152	,000		,398	
X73	,000	,000	,198	,000	,000	,001	,000	,000	,000	,000	,001	,000	,000	,000	,000	,000		,398	

Tulostus on valitettavasti tässä kovin pientä, mutta matriisista huomaamme, että valtaosin korrelaatiokertoimet poikkeavat nolasta. [Lukijalle: Acrobat Readerissa voi tietenkin valita suuremman kuvakoon]

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,871
Bartlett's Test of Sphericity	Approx. Chi-Square	6452,114
	df	153
	Sig.	,000

Seuraavaksi raportoidaan korrelaatiomatriisin kuntoisuuteen liittyviä testisuureita (*KMO and Bartlett's Test*). Sekä **Kaiserin testi** (Kaiser-Meyer-Olkin -testi) – jonka arvo on suurempi kuin 0.6 – että **Bartlettin sväärisyystesti** ($p < 0.0001$) osoittavat, että korrelaatiomatriisi on sovelias pääkomponenttianalysiin.