

Norsk kortversjon av Big Five Inventory (BFI-20)

PSY
KOL
OGI

Det er ikke mulig å måle noe så omfattende og sammensatt som personlighet på en optimal måte med kun 20 ledd. Men der optimal måling ikke er påkrevd, kan denne versjonen av Big Five Inventory likevel være praktisk nyttig.

TEKST

Harald Engvik

Sten-Erik Clausen

PUBLISERT 5. september 2011

ABSTRACT:

A Norwegian short version of the Big Five Inventory

The aim of this study is to develop a short version of the Big Five Inventory. This may be very useful in situations where time is limited. We used a Norwegian version of BFI-44 to develop a 20 items version of the test. This short version reached adequate levels of psychometric properties, which was measured by structural validity, factor divergens, maximal representation, test-retest reliability and criterion validity. This short version of BFI is limited, but it may be useful in several settings that do not require optimal measurement of personality. An example may be large scale survey studies in the general population.

Keywords: personality, Big Five Inventory, short version

EMNER

personlighet

Big Five Inventory

Kortversjon

Big Five Inventory (BFI) er blant de mest brukte personlighetstester basert på femfaktormodellen (FFM) for personlighet. Den utmerker seg ved å være kortere enn de fleste alternativer (44 testledd), og ved å være eksplisitt forankret i den leksikalske personlighetstradisjonen (John, Angleitner & Ostendorf, 1988). I korthet betyr dette at individuelle egenskaper som er språklig framtrædende eller viktige, kan beskrives med noen få enkle og fundamentale begreper. Datareduksjon og faktoranalyse står sentralt i denne tradisjonen.

«Forkortelse av tester er en avveining mellom praktiske effektiviseringsbehov og hva som er psykometrisk forsvarlig»

Testleddene i BFI ble valgt ut fra «prototypikalitet» for «the Big Five» (John & Srivastava, 1999). Sammen med det korte formatet gjør dette at BFI operasjonaliserer de 5 hovedfaktorene uten målsetting om å dekke alle nyanser. En norsk versjon av BFI-44 har vist akseptable psykometriske egenskaper (Engvik & Føllesdal, 2005).

Det korte formatet har utvilsomt bidratt til testens anvendbarhet og popularitet. Men for noen formål anses selv denne testen som for omfattende. Versjoner helt ned til 10 testledd er foreslått, som «Ten Items Personality Inventory» (Gosling, Rentfrow & Swann, 2003) og BFI-10 (Rammstedt & John, 2006).

Det har lenge vært kjent at testledd som måler samme faktor, ofte ikke er likeverdige (Burisch, 1997), og at man derfor kan velge ut gode testledd på en måte som minimaliserer tap av psykometrisk kvalitet. Men det er utvilsomt en nedre grense for hvor langt en slik forkortingsprosess kan gå før det blir vanskelig å påberope seg at testen måler «de fem store». Og selv under forutsetning av likeverdige testledd er det slik at forkortelse av tester alltid fører til svekkelse av reliabilitet og validitet.

Forkortelse av tester er derfor en avveining mellom praktiske effektiviseringsbehov og hva som er psykometrisk forsvarlig. Avveilingen avhenger av hvilken definisjon av FFM man starter med. Vårt utgangspunkt er BFI-44, som teoretisk ikke er satt sammen av underfaktorer. Den inneholder likevel et spektrum av egenskaper for hver faktor. Dette spekteret kan beskrives i en modell som først ble lansert på 90-tallet under betegnelsen Abridged Big Five Dimensional Circumplex (AB5C) (Hofstee, De Raad, & Goldberg, 1992). AB5C systematiserer underfaktorer for «de fem store» på grunnlag av faktoranalytiske undersøkelser av personlighetsbeskrivende adjektiver i språket. Mens femfaktormodellen klassifiserer trekk kun etter høyeste faktorladning, tar AB5C-modellen hensyn til den største av eventuelle bifaktorer.

Norsk tilpasning

Formålet med denne undersøkelsen er å komme fra 44 til 20 testledd, med minimalt tap av psykometrisk kvalitet. Vi vil betegne faktorene med tall fra 1 til 5 og angi retning med fortegn. Her er tre eksempler: «Pratsom» er vanligvis ren ekstraversjon og kan betegnes som 1+. «Sosial», som har hovedladning på faktor 1 og positiv biladning på faktor 2, betegnes som 1+2+, mens «dominerende» betegnes som 1+2-, dvs. hovedfaktor 1+ og bifaktor 2-. De tre eksemplene har alle samme hovedfaktor (1), men ulike bifaktorer.

Metoden som ble benyttet, kan beskrives som utprøving av ulike versjoner under overvåking av psykometrisk kvalitet. Den psykometriske kvaliteten ble målt med *strukturell validitet* (AB5C-klassifisering, indre konsistens og faktordivergens, dvs. minimal korrelasjon mellom ulike faktorer), *maksimal representativitet* (multippel korrelasjon mellom faktorene i BFI-20 og tilsvarende faktorer i BFI-44), *maksimal stabilitet* over tid (test–retest-reliabilitet) og *maksimal kriterievaliditet* (multippel korrelasjon mellom egenvurderinger og medstudentvurderinger). Vi forsøkte dessuten å oppnå balansert skåring, som innebærer like mange positive og negative ladninger for

hver faktor. Dette siste kriteriet ble oppnådd for BFI-20 som helhet, men ikke for hovedfaktor IV og V, som har henholdsvis 1 og 3 positive ladninger.

Utvalg og analyser

Strukturell validitet, indre konsistens, representativitet og faktordivergens ble testet i et bekvemmelighetsutvalg (n=630, 54 % menn, aldersvarisjon 18–72 år). Retestreliabilitet og samsvar med medstudentvurderinger ble undersøkt i to studentutvalg fra BI, med henholdsvis n=133 og n=150, begge utvalg med 57 % kvinner.

Egenvurderinger ble foretatt på 7-punkts skalaer forankret med «Passer ikke – Passer helt», og det var kun ytterpunktene som hadde verbale beskrivelser.

Medstudentvurderinger ble foretatt med 20 7-punkts bipolare adjektivskalaer: *ekstraversjon* (aktiv, pratsom, dominerende, utadvendt), *vennlig* (varm, ikke kynisk, omsorgsfull, vennlig), *kontroll* (velorganisert, strukturert, planmessig, disiplinert), *emosjonell stabilitet* (ubekymret, ikke hårsår, ikke nervøs, ikke anspent) og *fantasi* (original, fantasifull, kreativ, oppfinnsom) (Engvik, 1993). Medstudentvurderingene ble samlet inn i 20 seminargrupper med 6–10 deltagere i hver gruppe etter nesten to semestres deltagelse i samme gruppe. Skårer for den enkelte ble beregnet som gjennomsnitt av medstudentvurderingene i gruppa.

For BFI-44 ble det laget tre parseller for hver av de fem hypotetiske faktorene ved å beregne gjennomsnitt for 3 (4) fortløpende testledd. Parseller består av kombinasjoner av enkelttestledd som utgjør såkalte parsellskårer. Dette gjøres bl.a. for å redusere innvirkning av eventuelt avvikende testledd på resultatet. 15 parseller ble analysert med PCA- og VARIMAX- rotasjon, og resultatet viste svært god overensstemmelse med FFM. AB5C- klassifisering av testledd ble foretatt på grunnlag av testleddenes korrelasjoner med ortogonale faktorskårer (noe som er det samme som testleddenes faktorladninger i parsellstrukturen).

Indre konsistens ble definert på grunnlag av testleddenes korrelasjon med øvrige testledd for hver faktor. Vurdering av representativitet ble foretatt på grunnlag av trinnvis regresjonsanalyse med skåre på BFI-44 som avhengig variabel. Kriterierelatert validitet ble beregnet på tilsvarende måte med medstudentvurderinger som avhengig variabel.

Resultater

Analyse av test–retest-reliabilitet (2 måneders intervall) viste tilfredsstillende stabilitet for samtlige testledd i BFI-44. Reliabilitet varierte mellom 0,38 og 0,78 for de enkelte testledd. Dette viser at testleddene i BFI-44 faktisk måler relativt stabile individuelle forskjeller. To testledd hadde reliabilitet lavere enn 0,40: «er reservert» og «er avslappet, takler stress godt», og kunne derfor med fordel utelates fra BFI-20. Vi inkluderte imidlertid det siste leddet på grunn av ønsket om maksimal balansert skåring.

Samtlige testledd i BFI-44 har akseptable strukturelle egenskaper, men tre testledd er likevel marginale og kan utelates på dette grunnlag («er selvhevdende», «er full av

energi» og «foretrekker rutinearbeid»).

Tabell 1 viser testleddene som ble valgt ut. Kolonnene for alfa, R BFI-44 og R Medstudentvurdering viser hvilke kriterier som bidro til utvelgelsen. Som ventet konvergerer ikke kriteriene. Maksimum alfa har en tendens til å snevre inn definisjonen av faktorene, mens regresjon mot BFI-44 skårer tenderer til å ta vare på ulike nyanser.

Tabell 1. Testledd i BFI 20¹

Testledd		Strukturell validitet AB5C	Alfa	R BFI-44	R medstudentvurdering
Ekstraversjon					
1	Er pratsom	1+	X		X
- 21	Har en tendens til å være stille av meg	1-	X	X	X
- 31	Kan være sky og hemmet	1-4-		X	
36	Er utadvendt og sosial	1+2+	X	X	
Vennlig					
7	Er hjelpsom og uegoistisk i forhold til andre	2+			
- 27	Kan være kald og fjern	2-1-	X	X	X
32	Er hensynsfull og vennlig overfor de fleste mennesker	2+	X		X
- 37	Kan noen ganger være uhøflig mot andre	2-3-	X	X	X
Kontroll					
3	Gjør en grundig jobb	3+	X	X	X
-8	Kan være uforsiktig	3-1+		X	X

- 18	Har en tendens til å ha lite orden på tilværelsen	3-	X	X	X
38	Legger planer og følger dem opp	3+			X
Emosjonell stabilitet					
-4	Er deprimert, nedtrykt	4-1-	X	X	X
9	Er avslappet, takler stress godt	4+	X	X	X
- 19	Bekymrer seg mye	4-1-	X	X	X
- 39	Blir lett nervøs	4-		X	
Fantasi					
5	Er original, kommer med nye ideer	5+1+	X		X
20	Har livlig fantasi	5+			X
40	Liker å spekulere, lek med ideer	5+	X	X	
- 41	Har få kunstneriske interesser	5-1-			

1 X markerer antall innfridde kriterier for hver faktor. X betyr altså at testleddet er blant de fire beste for en gitt faktor og et gitt kriterium.

Tabell 2 oppsummerer psykometriske egenskaper for BFI-20. På bakgrunn av det korte formatet har alle faktorer akseptable verdier. Spesielt imponerende er verdiene for faktor 1 med reliabilitet på rundt 0,80, representativitet rundt 0,90 og validitet = 0,73. Spesielt høy intersubjektiv validitet for faktor 1 er i overensstemmelse med andre undersøkelser (Ulveseth, 2007).

TABELL 2. INDRE KONSISTENS, RETEST-RELIABILITET, REPRESENTATIVITET OG PREDIKTIV VALIDITET FOR BFI-20

<i>Ekstraversjon</i>	<i>0.78</i>	<i>0.79</i>	<i>0.91</i>	<i>0.73</i>
<i>Vennlig</i>	<i>0.63</i>	<i>0.67</i>	<i>0.85</i>	<i>0.41</i>
<i>Kontroll</i>	<i>0.57</i>	<i>0.67</i>	<i>0.88</i>	<i>0.44</i>

TABELL 2. INDRE KONSISTENS, RETEST-RELIABILITET, REPRESENTATIVITET OG PREDIKTIV VALIDITET FOR BFI-20

<i>Emosjonell stabilitet</i>	0.73	0.70	0.89	0.32
<i>Fantasi</i>	0.63	0.76	0.89	0.42

Begrensninger

At BFI-20 kun består av 20 testledd, begrenser den som et mål på «de fem store». Det er rett og slett ikke mulig å måle noe så omfattende og sammensatt som personlighet på en optimal måte med kun 20 testledd. BFI-20 kan likevel være nyttig og praktisk i situasjoner som ikke krever optimal måling av personlighet.

BFI-20 begrenses også av de testleddene som finnes i BFI-44. Som nevnt over er dette utvalget foretatt på grunnlag av prototypikalitet. Dette fremmer konvergens heller enn diversitet, og mange av nyansene som man finner i AB5C, er rett og slett ikke dekket i BFI-44. Noen testledd i BFI-44 bryter dessuten med vanlige konvensjoner for hva som er gode testledd, blant annet ved at de kan inneholde ulike betydninger i samme testledd, eksempelvis «kan være kald og fjern», «kan være sky og hemmet». På denne bakgrunn er det ikke urimelig å tro på muligheten av å utvikle tester med 20 testledd som både har bedre konvergens og divergens enn BFI-20 (Se Saucier (2002) for et forslag med god divergens). Testleddene i International Personality Item Pool (2001) er et naturlig sted å starte søket etter gode og dekkende testledd. Imidlertid er det ikke trolig at den generelle kvaliteten for en slik test ville være vesentlig bedre enn for BFI-20.

En siste begrensning ved undersøkelsen er at alle empiriske analyser er foretatt på bekvemmelighetsutvalg av informanter. Dette gjør det umulig å foreta statistiske generaliseringer. Det er likevel gode grunner til å tro at hovedtrekkene i de empiriske analysene er generaliserbare.

Tilgjengelighet

Denne versjonen av BFI-20 er gjengitt i tabell 3, og er fritt tilgjengelig for alle som ønsker å benytte den. Det er en forutsetning at denne artikkelen siteres i forbindelse med publisering av resultater. 1

TABELL 3. BFI-20 TESTLEDD

NEDENFOR FINNER DU EN REKKE PÅSTANDER SOM PASSER MER ELLER MINDRE GODT FOR ULIKE MENNESKER. KRYSS AV I DEN RUTEN SOM PASSER BEST FOR DEG SLIK DU VANLIGVIS ER. IKKE TENK FOR MYE PÅ HVER OPPGAVE, MEN SETT ET KRYSS I RUTEN DU UMIDDELBART SYNES STEMMER BEST. SETT ETT KRYSS PER LINJE.

Er pratsom

Kan være kald og fjern

Gjør en grundig jobb

TABELL 3. BFI-20 TESTLEDD

NEDENFOR FINNER DU EN REKKE PÅSTANDER SOM PASSER MER ELLER MINDRE GODT FOR ULIKE MENNESKER. KRYSS AV I DEN RUTEN SOM PASSER BEST FOR DEG SLIK DU VANLIGVIS ER. IKKE TENK FOR MYE PÅ HVER OPPGAVE, MEN SETT ET KRYSS I RUTEN DU UMIDDELBART SYNES STEMMER BEST. SETT ETT KRYSS PER LINJE.

Er deprimentert, nedtrykt

Er original, kommer med nye ideer

Har en tendens til å være stille av seg

Er hjelpsom og uegoistisk i forhold til andre

Har en tendens til å ha lite orden på tilværelsen

Er avslappet, takler stress godt

Har livlig fantasi

Er utadvendt og sosial

Kan noen ganger være uhøflig

Legger planer og følger dem opp

Bekymrer seg mye

Liker å spekulere, leke med ideer

Kan være sky og hemmet

Er hensynsfull og vennlig overfor de fleste mennesker

Kan være uforsiktig

Blir lett nervøs

Har få kunstneriske interesser

Teksten sto på trykk første gang i Tidsskrift for Norsk psykologforening, Vol 48, nummer 9, 2011, side 869-872

TEKST

Harald Engvik

KONTAKT: harald.engvik@psykologi.uio.no

Sten-Erik Clausen, Norsk institutt for by- og regionforskning (NIBR) Norsk institutt for forskning om oppvekst, velferd og aldring (NOVA)

Referanser

- Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, 11, 303-315.
- Engvik, H. (1993). «Big Five» på norsk. *Tidsskrift for Norsk Psykologforening*, 30, 884-896.
- Engvik, H. & Føllesdal, H. (2005). Big Five på norsk, *Tidsskrift for Norsk Psykologforening*, 42, 128-129.
- Gosling, S. D., Rentford, P. J., Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-528.
- Hofstee, W. K. B., De Raad, B. & Goldberg, L.R. (1992). Integration of the Big Five and Circumplex Approaches to Trait Structure, *Journal of Personality and Social Psychology*, 63, 146-163.
- International Personality Item Pool (2001). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences. Web site: <http://ipip.ori.org/>
- John, O. P., Angleitner, A. & Ostendorf, F. (1988). The Lexical Approach to Personality: A Historical Review of Trait Taxonomic Research. *European Journal of Personality*, 2, 171-203.
- John, O. P. & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives In L. A. Pervin & O. P. John (red.), *Handbook of Personality. Theory and Research*, (2nd ed., s. 102-138). New York: The Guilford press.
- Rammstedt, B. & John, O. P. (2006). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203-212.
- Saucier, G. (2002). Orthogonal Markers for Orthogonal Factors: The Case of the Big Five. *Journal of Personality*, 36, 1-31.
- Ulveseth, E. (2007) Personpersepsjon. Realitet eller sosiale stereotypier, Hovedoppgave, Psykologisk institutt, Universitetet i Oslo.