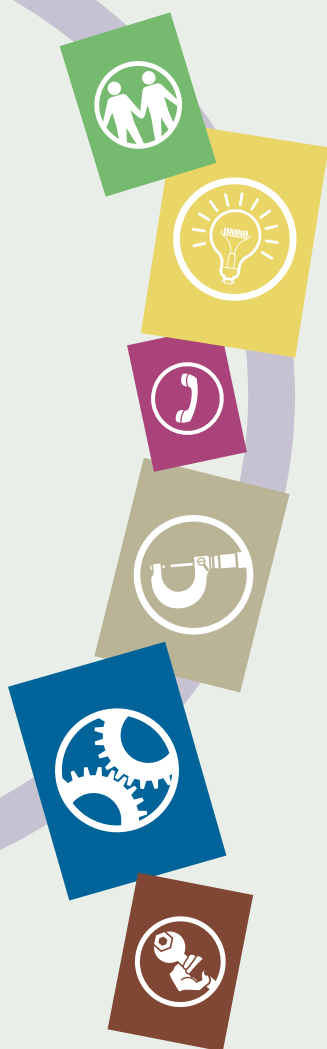











En gjennomgang av forskning og kvalitet



Teamrollene i kortversjon

I følge Belbins forskning beskrives alle roller med noen typiske positive bidrag, men også ved at rollene har det vi kan kalle "tillatte svakheter". Disse må tolereres for å kunne gjøre seg nytte av de styrker som henger sammen med dem. Rollene kan grupperes i tre hovedgrupper: 1) Reflekerende roller, 2) Sosiale roller og 3) Handlende roller.

ROLLENE	STYRKER	TILLATTE SVAKHETER SOM KAN FOREKOMME
Reflekerende roller		
PL 	Ideskaper (Plant) Kreativ, fantasirik, ukonvensjonell. Løser vanskelige problemer.	Overser detaljer. For opptatt av egne idéer til å kunne kommunisere effektivt.
ME 	Analytiker (Monitor Evaluator) Sindig, strategisk og skarpsindig. Ser alle valgmuligheter. Bedømmer nøyaktig.	Mangler drivkraft og evne til å inspirere andre.
SP 	Spesialist (Specialist) Fagorientert, egenmotivert og opptatt av arbeidet sitt. Representerer fagstolthet og profesjonalitet på eget felt.	Bidrar på et begrenset fagområde. Dveler ved detaljer.
Sosiale roller		
CO 	Koordinator (Co-ordinator) Moden, trygg, en god møteleder. Fokuserer på mål og beslutninger, er god til å delegere.	Kan oppfattes som manipulerende. Dytter oppgaver over på andre.
RI 	Ressursinnhenter (Resource Investigator) Utadvendt, entusiastisk, kommuniserende, undersøker muligheter. Skaper og utvikler kontakter.	Overoptimistisk, mister interessen så snart den første entusiasmen har lagt seg.
TW 	Lagspiller (Teamworker) Samarbeidsvillig, mild, empatisk og diplomatisk. Lytter, bygger opp, avverger friksjoner i gruppen.	Ubesluttosom når avgjørelser skal tas.
Handlende roller		
SH 	Pådriver (Shaper) Utfordrende, dynamisk, trives under press. Har drivkraft og mot til å overvinne vanskeligheter.	Tilbøyelig til å provosere. Sårer andre.
IMP 	Iverksetter (Implementer) Disiplinert, pålitelig, konservativ og effektiv. Omsetter idéer til praktisk handling.	Lite fleksibel. Reagerer tregt på nye muligheter.
CF 	Avslutter (Completer Finisher) Samvittighetsfull, nøyaktig. Finner feil og mangler. Overholder tidsfrister.	Tilbøyelig til å bekymre seg. Motvillig til å delegere.

Belbin

– en gjennomgang av forskning og kvalitet

Forord

Helt siden Meredith Belbin presenterte sin teamrollemodell på 80 tallet, har modellen vært gjenstand for omfattende etterprøving og kontrollerende forskning. Modellen var ny, kreativ og ikke minst praktisk orientert. Den vant hurtig praktikernes tilslutning og ble et av verdens mest brukte konsepter for utvikling og evaluering av team. Dette var ikke minst med på å skjerpe akademias kritiske blikk. Særlig tidlig på 90-tallet ble det publisert artikler som var meget kritiske til Belbins forskningsmetodikk og modellens holdbarhet. Etter den tid er det publisert nær 50 vitenskapelige artikler som på ulik måte har ettergått Belbins teori og modell.

I 2007 publiserte et forskerteam bestående av Aitor Aritzeta, Stephen Swailes and Barbara Senior en dyptpløyende gjennomgang av 43 vitenskapelige studier i *Journal of Management Studies* 44: Belbin's Team Role Model: Development, Validity and Applications for Team Building.

Denne artikkelen hviler blant annet på deres gjennomgang.

ATFERD ELLER PERSONLIGHET

Meredith Belbin er opptatt av å kartlegge en persons faktiske atferd i samspill med andre. Hans skjema for egenvurdering (SPI – Self Perception Inventory) er derfor designet for å måle trekk ved den atferden enkeltpersoner viser i slikt samspill og ikke personens personlighetstrekk.

Med et menneskes personlighet tenker vi på personens rimelig faste og stabile trekk som særpreger vedkommendes måter å tenke, føle og reagere på. Formålet med en personlighetstest er altså å kunne si noe om personens typiske og regelmessige mønster. Kartleggingen skjer ved en form for egenvurdering hvor kandidaten skal vise sin typiske måte å tenke, føle og reagere på.

Kartleggingen i Belbins tilnærming skjer ved at personen selv skal vurdere hvordan han eller hun vil oppføre seg i bestemte beskrevne situasjoner, i tillegg skal et antall observatører si hvilken atferd de har sett at vedkommende faktisk har vist.

Hvis vi nå antar at en persons atferd i en konkret setting alltid utelukkende påvirkes av hvem personen er (personlighet) vil Belbin og personlighetstesten vise det samme – om enn kanskje i litt ulike begreper. Men er det slik at det bare er vår personlighet som helt ut bestemmer hva vi gjør i en bestemt situasjon? De lærde strides. Noen hevder at det i all hovedsak er personlighet det handler om. Andre mener at personlighet er uinteressant, det handler primært om lært atferd og påvirkning fra andre. Vi tilhører de som mener at de fleste mennesker har noen klart gjenkjennelige trekk og typiske måter å være på, men at vi samtidig har et repertoar av ulike væremåter. Hva vi velger henger nær sammen med hva vi har lært, hvilke andre personer vi er sammen med og ikke minst hvilken reell posisjon vi har i øyeblikket. Vi vil framstå på en annen måte som trainee enn som administrerende.

Det å bruke personlighetstest til å forutsi atferd er derfor ikke uten videre enkelt. Mange psykologer uttrykker da også skepsis til denne bruken. Hvis formålet er å vite noe om sannsynlighet for visse handlingsmønstre, kan det være svært nyttig, men hvis tanken er å forutsi konkret atferd i enkeltsituasjoner, er sjansen for overraskelse stor. Belbin legger til grunn at en persons atferd i en bestemt samhandling er avhengig av en rekke faktorer, hvorav personlighet er én. Han bruker en modell som er vist til høyre. Atferd henger sammen med alle disse faktorene.



Det vi har sagt har konsekvenser når vi bruker Belbin og personlighetstester i praksis. Nettopp fordi personlighet handler om å finne en persons regelmessige mønster, vil vi forvente at ny personlighetstest etter en tid vil vise om lag samme resultat. For Belbin vil vi imidlertid forvente at en ny måling kan få helt andre resultater. Spesielt hvis kandidaten har en annen formell rolle, samhandler med andre eller i perioden har gjennomgått en læringsprosess. I praksis viser det seg at Belbin har en

relativt høy stabilitet over tid. Noe som taler for at den stabile personligheten (og andre rimelig stabile faktorer) spiller en rolle, men det er ikke bare den som avgjør. Fornyet kartlegging viser gjerne endringer innenfor et hovedmønster. Man vil typisk finne noen av de samme rollene høyt, ikke i samme rekkefølge og ikke alltid like høyt, men hovedmønsteret er ofte gjenkjennelig.

“The dynamic configuration of team roles measured by the SPI and the relative stability of traits measured by personality questionnaires leads to the conclusion that traits measured by the latter are different from those measured by the SPI. Thus, both instruments may be tapping different but complementary constructs.”
~ Aritzeta, Swailes & Senior, 2007

I mange miljøer er man skeptiske til tester fordi det gjerne leder til en type stemping som virker lite konstruktivt når det gjelder videre personlig utvikling. Tolkningen av resultatet blir at du er slik eller slik. Belbins tilnærming gir et annet resultat, ved å peke på atferdsmessig repertoar og mangfoldet i en persons potensial som kan videreutvikles.

Belbins måleinstrument - SPI

Egenvurderingsskjemaet – SPI, er laget for måling av atferd i en organisatorisk sammenheng. Skjemaet måler 9 dimensjoner av atferd i team og har en skala kalt DR (Dropped Points) som er en kvalitetssjekk på utfylling ved at det fanger påståtte egenskaper snarere enn valide bidrag i en teamrolle. Overskriftene i skjemaet gir ulike arbeidslivsbaserte situasjoner som kandidaten kan identifisere seg med. Hensikten er å sette den type atferd som måles inn i en arbeidsmessig sammenheng og slik bidra til at kandidaten tenker konkret og operasjonelt under besvarelsen.

Når kandidaten fyller ut skjemaet skal han eller hun fordele 10 poeng på like mange utsagn. Mener kandidaten å kunne kjenne seg like sterkt igjen i to utsagn, fordeles det 5 poeng på hver. Hvis man mener at flere er relevante fordeles de ti poengene på flere utsagn. Dette gjentas for i alt 8 seksjoner. Kandidatene oppmuntres til å unngå begge ekstremer: Alle poengene på ett utsagn eller ett poeng på hver.

SPI er delvis ipsativt

Skjemaet er designet for å sikre så mye informasjon som mulig, samtidig som det skal være håndterbart i oppsett og lengde sett fra brukers side. Mange psykometriske tester som for eksempel 16PF, krever at du må ta stilling til et stort antall spørsmål, ikke sjelden mellom 100 og 200. Typisk brukes en Likert skala, dvs at svaret skal markeres på en skala fra for eksempel "Helt enig" til "Helt uenig", gjerne med et nøytralt midtpunkt. SPI på sin side ber kandidaten vurdere ulike utsagn opp mot hverandre og prioritere noen foran andre.

I SPI er den samlede skåren alltid 10 innen en seksjon og alltid 70 totalt. Det vil si at de enkelte utsagnene og den skår de får er avhengige av hverandre innen seksjonen, men ikke på tvers av seksjonene, noe som betyr at SPI er ipsativ innen seksjonene, men ikke på tvers. Belbins begrunnelse for dette valget er at dette samlet gir tydeligere informasjon om preferanser enn en egenvurdering på helt uavhengige variable.

I 1998 konverterte Sommerville & Dalziel¹ SPI til en Likert skala. De fant at 73 % av respondentene hadde samme teamroller på tvers av målingsform noe som indikerer at de to metodiske alternativer ikke gir helt ulike svar. Selv om det er anerkjent at ipsative målinger representerer noen statistiske og metodiske utfordringer er heller ikke normative data uten utfordringer, se tekstboks².

"There is a strong controversy in the literature about the ipsative scoring of constructs, and such discussion while very interesting is still far from clear. Ipsative data is not free from criticism as neither are normative data [sic]. Likert type scales are not free from important threats as central tendency bias, acquiescence, social desirability and falsification of responses which may be much larger from these types of scales than for ipsative scales" ~ Aritzeta, Swailes & McIntyre-Bhatty, 2004

¹ *Project teambuilding – the applicability of Belbin's team role self-perception inventory* (Sommerville & Dalziel, 1998), pp. 166-167.

² *Further Evidence on the validity of the Belbin Team Role Self Perception Inventory and the Observer's Assessment Sheet* (Aritzeta, Swailes & McIntyre-Bhatty, 2004), p.8.

Etter å ha gjennomført en grundig statistisk analyse av data fra mer enn 5000 kandidater konkluderte Swailes og Aritzeta med at metoden gir ingen negative effekter i forhold til konstruksjonsvaliditeten av instrumentet og at den gjensidige avhengigheten er forholdsvis lav. Vi skal komme tilbake til dette under kapitlet om validitet. Belbin har derfor valgt å beholde den delvis ipsative tilnærmingen siden den gir viktige fordeler for brukere.

Bruken av SPI

Verktøyet er i Norge kun tilgjengelig gjennom Belbin Norge AS. Selv om man kan finne versjoner av skjemaene på for eksempel internett kan det kun skåres og beregnes profiler ved bruk av Belbins verktøy Belbin Online (Tidligere e-Interplace). I sin bok *Management Teams: Why They Succeed or Fail*, først publisert i 1981, inkluderte Meredith Belbin en versjon av et tidlig skjema med en skåringstabell som kunne brukes til å beregne en egenutfylling. Hensikten var å gi leseren en kjapp måte å få en indikasjon på egne teamroller. Videre forskning viste imidlertid at denne versjonen var uegnet som måleinstrument. Derfor ble dataprogrammet e-Interplace utviklet som eneste pålitelige måte å håndtere de dataanalyser, beregningsmodeller og normeringskrav som må til for å få pålitelige rapporter.

I den tidlige versjonen av skjemaet var heller ikke den niende rollen – Spesialisten – inkludert. Den innledende delen av forskningen var først og fremst basert på forretningsspill og simuleringer, da ble ikke rollen identifisert. Her var et av poengene å gjøre spillene generelle og bransjeuavhengige. Når Belbin flyttet arbeidet fra eksperimentering i laboratoriet til den virkelige verden, ble kompetanse og faglig profesjonalitet en vesentlig del av samspillet i teamene.

Belbin® er et registrert varemerke også i Norge og all bruk av Belbins spørreskjemaer på norsk som ikke er autorisert av Belbin Norge AS er et brudd på copyright rettigheter.

Observatører

Mye av den metodiske diskusjonen om Belbins teammodell har vært knyttet til holdbarheten av skjemaet for egenvurdering - SPI. Det er dette som i hovedsak diskuteres i kapitlet om validitet. Det er samtidig viktig å understreke betydningen av at observatører er en integrert del av Belbins evalueringsmetode som det anbefales sterkt at brukes³. De fleste personlighetstester baserer seg på kandidatens egenvurdering. Når fokus er atferd, er det vesentlig å være klar over at atferd både kan oppfattes forskjellig og at egen vurdering av atferd kan ha skjevheter i retning av ønsket atferd, ikke bare faktisk atferd. Bruk av observatører gir en mer robust vurdering og

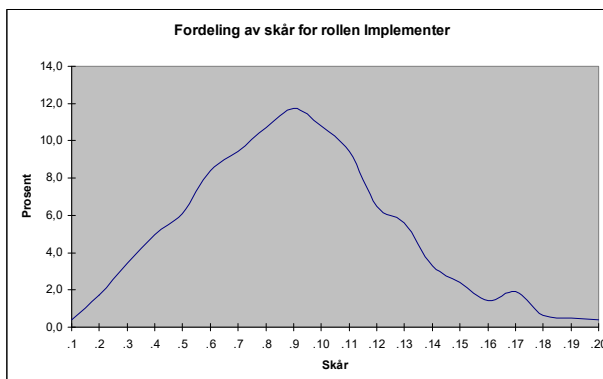
“Line managers were usually wary of using self-reporting measures when reaching crucial decisions about people. That reservation is seldom connected with technical issues of test construction but more with the recognition that people are subject to illusions about the self and are also tempted to distort their responses once they believe that their answers affect job and career prospects. Line managers place a greater emphasis on observations of others, believing that such material has greater validity, in terms of effective decision making, providing it is properly gathered.” ~ Meredith Belbin, 1993

³ A reply to the Belbin Team-Role Self-Perception Inventory by Furnham, Steele and Pendleton (Belbin, 1993), *Journal of Occupational and Organizational Psychology* (1993), 66, p.259.

representerer en kvalitetssikret 360° tilbakemelding på kandidatens atferd i samspill med andre. Dermed vises kandidatens evne til å veksle atferd – bruke ulike teamroller – i samspillet med ulike typer andre personer.

Normering

Når en kandidat fyller ut egenvurderingen tildeles som nevnt poeng til ulike utsagn. For å kunne tolke en bestemt kandidats poengfordeling trenger vi noe å sammenlikne med. Hva som er mange poeng og få poeng må avgjøres gjennom en slik sammenlikning. Sammenlikningsgrunnlaget er det vi kaller normeringen.⁴⁾ Til grunn for normering antas det at Belbins teamroller har en fordeling som er tilnærmet normalfordelt. Når Belbin første gang ble tilrettelagt for norsk på siste del av 90 tallet, ble det gjort statistiske kvalitetssjekker på norsk tilrettelegging med utgangspunkt med utgangspunkt i et begrenset antall personer på noen hundre. Fordeling av rå-skår for en tilfeldig rolle så likevel ut som i figuren til høyre. I praksis består normeringen i å beregne prosentilene som gir oss anledning til å si hvilken skår har en bestemt andel av utvalget under seg. 50-prosentilen er lik medianen og har halvparten over og halvparten under skåren. Den norske databasen i dag består av over 8000 profiler.



“Without norming, it would be impossible to compare scores between roles. Without that comparison, the selection of natural, managed and avoided roles would be impossible. Secondly, norming filters out cultural impacts. This makes the role selection relative to the norms used.”

Chris Morison, 2008

Etter en grundig analyse av landforskjeller i Nord-Europa og intra-nasjonale spredninger, har vi valgt å benytte en felles normeringsdatabase både i den engelske og norske versjonen. Det ble beregnet effektstørrelser (Effect sizes – ES) for gjennomsnittsverdien for hvert land for hver rolle som igjen ble sammenholdt med det største enkeltlandet (UK). Dette, sammenholdt spredningen innen hvert land, som var til dels større en spredningsmålet på tvers, gjorde at man valgte en slik løsning fordi det gir betydelige praktiske fordeler. Eksempelvis vil i en internasjonalisert verden team på tvers av landegrensler produsert på både norsk og engelsk gir samme resultat. Normbasen inkluderer 9500 kandidater for egenvurderingsnormen og 14000 observatører for observatørnormen.

“If within-country variance is high, we think this offsets (or at least reduces) the need for different country norms because variations within a country seem much greater than variations between countries.”

Stephen Swailes, 2013

⁴ *An Investigation of Belbin Team Roles as a Measure of Business Culture* (Chris Morison, 2008), p.32.

Reliabilitet

Reliabilitet handler om pålitelighet. I testsammenheng er det en vurdering av hvor stabil og pålitelig det instrumentet som brukes i målingen er. Det er flere tilnærminger til en reliabilitetsanalyse. Det mest vanlige kravet knyttes til at hvis testen brukes flere ganger på samme fenomen skal den gi et konsistent resultat. Det er også vanlig å se på indre konsistens, det vil si at ulike deler av instrumentet som er ment å måle samme fenomen faktisk gjør det. Det er naturlig å drøfte reliabilitet før vi ser nærmere på validiteten fordi hvis instrumentet er svakt spiller det liten rolle om validiteten (at man faktisk måler det man påstår å måle) er tilfredsstillende.

Test-retest reliabilitet

Hensikten med Belbin er å måle atferd i en organisatorisk sammenheng. Det gjør en vurdering av test-retest svært krevende. Måling av personlighet vil man forvente er konsistent over tid. En kandidats atferd derimot, vil vi forvente endrer seg over tid med forandringer i stilling, oppgave og samspillspartnere.

Det å tilpasse seg og finne praktiske arbeidsformer og væremåter som passer med situasjonen ser vi vanligvis på som en fordel. Det betyr at det er utfordrende å gjennomføre holdbar test-retest analyser fordi trekk ved situasjonen kan ha endret seg selv om ikke forskeren er klar over det. I praksis er det sjelden at en persons profil endres totalt. Vi ser gjerne at de samme rollene går igjen som foretrukket, men gjerne i noe ulik rekkefølge.

Intern konsistens

En bedre tilnærming til reliabilitet for et instrument hvis hensikt det er å fange fenomener som per definisjon forventes å forandre seg, er indre konsistens. Dette handler om i hvilken grad de enkelte leddene i testen faktisk måler det samme. Belbin måler atferd opp mot 9 teamroller. En teamrolle defineres som et sett av måter å bidra på i et samspill med andre mennesker. Det vil si at beslektede væremåter eller kjennetegn på atferd er slått sammen og gitt et rollenavn. Rollenavnet er ikke et personlig trekk, men et navn på en måte å bete seg på.

Siden SPI er delvis ipsativt og ikke alle leddene i testen faktisk tilordnes verdi av kandidaten har SPI vært en statistisk utfordring å gjennomføre en holdbar reliabilitetsanalyse på. Dette har resultert i at det er publisert undersøkelser som viser svake resultater. I 2000 ble denne utfordringen grundig

adressert av to forskere: Stephan Swailes and Tim McIntyre-Bhatty⁵. De gjennomførte en grundig diskusjon av styrker og svakheter knyttet til Cronbach α som er det anerkjente målet for indre konsistens opp mot de utfordringer som følger av SPIs konstruksjon. Tidligere har forskere tilordnet påstander i SPI som kandidaten ikke har gitt poeng verdien 0. Dette påviser de at gir systematisk feil resultater. Det er statistisk stor forskjell på 0 og ingen verdi overhodet. Siden kandidaten ikke scorer på alle leddene gir det liten mening å regne ren korrelasjon mellom alle ledd. De velger derfor i tråd med Cronbach å bruke veide gjennomsnitt av korrelasjonen mellom leddene i måleskjema.

"This paper argues that, in overcoming the problem of unequal scale length between respondents, researchers have tampered with the true error variance contained within item response sets, leading to a misapplication of Cronbach's alpha in this case. Using weighted average inter-item correlation coefficients in a formula unrelated to scale length, alternative estimates of internal consistency reliability of the SPI scales are derived. For a large data set (n=5003), results show that the internal consistency of item responses is better than previous research suggests and tend towards the boundary at which internal consistency is considered acceptable in social and psychological research." ~ Swailes/McIntyre-Bhatty 2003

Disse forskerne hadde i motsetning til tidligere studier et meget stort datasett. Derfor kunne de for testhensyn også omgå problemet med ledd som ikke ble skåret, ved å kjøre en separat analyse for respondenter som hadde skår på alle rolle på tvers av seksjonene i skjemaet.

Denne første studien ga følgende resultat. Rollene har jevnt over gode reliabilitetskår, men vi ser at et par som IMP og ME ligger faretruende lavt.

Team roller	CF	CO	IMP	ME	PL	RI	SH	SP	TW
Cronbach α (Med 0)*	.46	.41	.23	.40	.54	.44	.66	.48	.53
Cronbach α (Skår i alle seksjoner)	.67	.79	.56	.60	.72	.78	.68	.65	.76
N for α med skår i alle seksjoner	161	133	358	124	86	126	482	84	197

n=5003⁶

Med utgangspunkt i denne studien iverksatt derfor Belbin Associates et utviklingsarbeid for å styrke presisjonen på instrumentet. Spørsmål ble både reformulert og det ble lagt til nye. En fornyet reliabilitetsstudie i 2012 viste følgende:

Team roller	CF	CO	IMP	ME	PL	RI	SH	SP	TW
Cronbach α (2002)	.67	.79	.56	.60	.72	.78	.68	.65	.76
Cronbach α (ny 2012)	.77	.75	.80	.77	.84	.71	.77	.67	.72

⁵ *The Reliability of the (Belbin) Team Role Self-Perception Inventory: Cronbach's alpha and ipsative scales* (McIntyre-Bhatty & Swailes, 2000).

Uses and Abuses of Reliability Estimates: The Case of the Belbin TRSPI (Swailes & McIntyre-Bhatty, 2001).

⁶ Stephan Swailes, McIntyre-Bhatty, "The Belbin team role inventory: reinterpreting reliability estimates", *Journal of Managerial Psychology* 2002, "Scale structure of the Team Role Self Perception Inventory", *Journal of Occupational and Organizational Psychology* 2003

Validitet

En undersøkelse har god validitet om den måler det den påstår å måle. Utfordringen er å sannsynliggjøre at så er tilfelle. Det finnes en rekke tilnærminger til validitet. Vanlig, men ikke tilstrekkelig alene, er "face-validity", dvs hvordan testen fremstår for bruker. Viktigere er konstruksjonsvaliditeten og det som kalles kriterievaliditet.

Face-validity

Det at testen fremstår som korrekt og pålitelig for bruker er vesentlig fordi det vil ha en tendens til å påvirke måten brukeren besvarer på. Ser det seriøst og forståelig ut øker sjansen for at man tar det på alvor og responderer ærlig og etter beste forståelse. Det er et stort antall rapporter som peker på at her skårer Belbins instrument høyt, et eksempel er Balderson & Broderick⁷ som i en studie fra 1996 slår fast at metoden har "svært høy face validity" og belegger det blant annet med måten kandidater mottar, aksepterer og omtaler resultatene av undersøkelsen. En vesentlig grunn til at modellen har fått den spredningen blant praktikere den har verden over, ligger her. De færreste leser validitetsrapporter, at undersøkelse og resultater fremstår pålitelig, forståelig og gjennomarbeidet er for mange viktigere. Forskningsmessig er det imidlertid ikke tilstrekkelig.

Konstruksjonsvaliditet

Validiteten har med tolkningen av data å gjøre, ikke selve dataene. I forhold til Belbin blir nøkkelspørsmålet om hans teamroller finnes i den virkelige verden som et fenomen og om hans begrepssett og modell gjenspeiler dette på en god nok måte. Det er flere måter å analysere dette på. Det som kalles konvergerende og diskriminerende validitet er viktig for å kunne ha tillit til konstruksjonen eller modellen. Hvis den konvergerende validiteten er høy vil vi forvente at resultatene fra Belbin vil være tilsvarende som andre undersøkelser som foregir å måle tilsvarende fenomener. Vi vil også forvente at et spørsmål som for eksempel måler én rolle, har høyere korrelasjon seg i, enn det settet av spørsmål som måler en annen rolle. Det vil si at det er en demonstrert diskriminering mellom elementene i modellen eller den samlede konstruksjonen.

Aritzeta⁸ al har i en grundig studie av 43 forskningsarbeider analysert konstruksjonsvaliditeten. Ved en gjennomgang av både teoretiske analyser og empiriske studier har de summert opp i hvilken grad det er mulig å fastslå om Belbins teamroller i tilstrekkelig grad nærmer seg eller gjenfinnes i andre konstruksjoner som arbeider med samme fenomener.

Taking the empirical studies together, there is sufficient evidence that definitions of team roles are valid and that independently of the measurement used to measure team roles, results are consistent with other theoretical models. The team role model shows evidence for validity that cannot be disregarded as presented in the table ~ Aritzeta, Swailes and Senior 2007

⁷ *Behaviour in Teams: Exploring Occupational and Gender Differences* (Balderson & Broderick, 1996), p.33.

⁸ *Belbin's Team Role Model: Development, Validity and Applications for Team Building* (Aritzeta, Swailes & Senior, 2007)

I sin oppsummering av studiene viser de bl.a til fellestrekk mellom teamrollene hentet fra andre teoretiske konstruksjoner om mennesker og atferd. (Her gjengitt på originalspråket for å holde begrepspresisjon).

Theoretical commonalities among teamroles ⁹	
Team role	Theoretical association ¹⁰
Completer Finisher	Adaptive cognitive style: tries to avoid conflict; Low emotional intelligence; High Moral values
Implementer	Adaptive cognitive style; Tries to avoid conflict; Low Machiavellianism; Cohesion; Low intellectual orientation; High Moral values
Team Worker	Bridge/Adaptive cognitive style; Tries to avoid conflict, Low Machiavellianism; Cohesion
Specialist	Adaptive cognitive style; In conflicts will try to dominate or use avoiding behaviour
Monitor Evaluator	Bridge/Adaptive cognitive style
Co-ordinator	Bridge; Attempts to control; In conflicts tries to find a compromise; Low Machiavellianism; High emotional intelligence; Cohesion
Resource Investigator	Innovative Cognitive style; Attempts to control; In conflicts tries to find a compromise; High emotional intelligence; Low conflict
Plant	Innovative cognitive style; In conflict tries to dominate; Shows Machiavellian behavior; Low cohesion, Intellectual orientation
Shaper	Innovative cognitive style; Attempts to control; In conflicts tries to dominate; Shows Machiavellian behavior; Achievement orientation; Low emotional orientation; Conflict

De foretar en tilsvarende gjennomgang av faktorstrukturen i teamrollemodellen og summerer opp med: “[This] leads us to conclude that the team role model has acceptable convergent validity. Factor structures for the TRSPI are coherent in its ipsative and normative forms as well as with personality measures.” ~ Aritzeta, Swailes & McIntyre-Bhatty, 2004¹¹

Et annet forskerteam (Fisher, Hunter & Macrosson) leverte tilsvarende støtte for diskriminerende validitet ved å dokumentere evne til å predikere forekomst av de ulike teamrollene basert på Belbins modell og konkluderte ”this adds weight to claims of construct validity for the SPI”¹².

Kriterievaliditet

Kriterievaliditet handler om med hvilken sikkerhet vi kan trekke holdbare slutninger fra de resultater vi får, for eksempel forutsi en fremtidig tilstand (basert på et definert kriterium) eller får parallelle resultater med et annet samtidig kriterium.

⁹, *Belbin’s Team Role Model: Development, validity and Applications for team Building*. Aritzeta, Swailes and Senior, 2007, pp 108)

¹⁰ Adaption-innovation cognitive styles, Kirton 1989, Aritzeta et al 2005b
Machiavellianism; W.D.K. Macrosson and D.J. Hemphill, 2001
Emotional intelligence; Dulewicz & Higgs, 1999
Cohesion subscale of the family environment survey, Fisher & Macrosson, 1995

¹¹ *Belbin’s Team Role Model: Development, Validity and Applications for Team Building* (Aritzeta, Swailes & Senior, 2007), p.111.

¹² *The distribution of Belbin team roles among UK managers* (Fisher, Hunter & Macrosson, 1998), p.132.

For Belbins modell vil det naturlige samtidige kriterium være observatørvurdering (OA – Observer Assessment). I sin artikkel fra 2004 fant Aritzeta et al støtte for samtidig kriterievaliditet basert på at av 3351 observasjoner, viste 66% signifikant overensstemmelse mellom observatørene når det gjelder teamrollefordeling. Ved en parallell analyse av SPI, fant de at 75% av korrelasjonene lå mellom moderat og sterk. Dette understøtter validiteten samtidig som det er viktig å ha i mente at vi i noen grad vil forvente ulike observasjoner mellom observatører og mellom observatørene og kandidaten. Belbin snakker om fire ulike profiler avhengig av samsvar mellom kandidat og observatørene.:

- Konsekvente; det vil si god sammenheng mellom det kandidaten mener er sterke roller og det som sees som de mest markerte rollene fra observatørene. Observatørene er i hovedsak enig seg i mellom. Tyder på at kandidaten kjenner sin atferd, er rimelig tydelig og at observatørene har oppfattet atferden på samme måte.
- Samsvarende; det er viktige fellestrekk mellom kandidat og observatører, selv om observatørene har tydelige ulikheter i sin observasjon. Forklaringen er typisk at observatører ser kandidaten i forskjellige situasjoner eller oppfatter vedkommende ulikt ut fra egen stilling og situasjon.
- Uoverensstemmende; det er tydelige ulikheter mellom kandidat og observatørene, men observatørene er i all hovedsak enig seg i mellom. Kan tyde på at kandidaten ikke helt kjenner seg selv eller søker bevisst å fremstille seg noe annerledes enn vedkommende faktisk oppfattes.
- Usammenhengende; intet klart bilde fremtrer, observatørene ser og vektlegger ulikt og det stemmer i begrenset grad med kandidatens eget bilde. Kan skyldes en forvirrende situasjon, gjerne kombinert med at kandidaten selv ikke helt vet verken hva som forventes eller hvilke ressurser vedkommende selv måtte ha. Ikke sjeldent forekommende når man er ung og eller helt ny i for eksempel en jobbsituasjon.

Det å kunne forutsi en fremtidig tilstand på individnivå er mer aktuelt når man måler ulike typer personlige trekk. En interessant studie tar det motsatte utgangspunkt. I følge teammodellen til Belbin er det ingen nødvendig sammenheng mellom teamrolleprofil og karriere. Det er for eksempel ikke slik at lederrollen forutsetter en bestemt teamrolle. Hvis dette er holdbart skal vi ikke forvente å finne spesielle sammenhenger mellom teamprofil på ett tidspunkt og andre statuselementer knyttet til karriereutvikling på tvers av miljøer. Dulewicz¹³ undersøkte sammenheng mellom teamrolle og indikasjoner på status/ansvar. Han så bl.a på forhold som lønn, antall underordnede og størrelse på det budsjett man hadde ansvar for. Han fant ingen sammenheng, hvilket må kunne tolkes som et argument for kriterievaliditet.

¹³ *A Validation of Belbin's Team Roles from 16PF and OPQ using Bosses' Ratings of Competence* (Dulewicz, 1995)

Litteratur

En betydelig del av tidlig publiserte studier basert på Belbin er utført med utgangspunkt i hans noe mangelfulle skjema fra tidlig 80-tall. Disse viser helt andre resultater enn de som baserer seg på data fra e-Interplace og den senere delen av hans forskning. Følgende studier har benyttet data fra e-Interplace:

- Aritzeta, Ayestaran & Swailes, *Team Role Preference and Conflict Management Styles* (2005)
- Aritzeta, Senior & Swailes, *Belbin Team Role Preference and Cognitive Styles: A Convergent Validity Study* (2004)
- Aritzeta, Senior & Swailes, *Belbin's Team Role Model: Development, Validity and Applications for Team Building* (2007)
- Aritzeta, Swailes & McIntyre-Bhatty, *Further Evidence on the Validity of the Belbin Team Role Self Perception Inventory and the Observer's Assessment Sheet* (2004)
- Aritzeta, Swailes & Senior, *Team Roles: Psychometric Evidence, Construct Validity and Team Building* (2005)
- Beck, Fisch & Bergander, *Functional Roles in Work Groups – An Empirical Approach to the Study of Group Role Diversity* (1999)
- McIntyre-Bhatty & Swailes, *The Reliability of the (Belbin) Team Role Self-Perception Inventory: Cronbach's alpha and ipsative scales* (2000)
- Morison, Chris, *An Investigation of Belbin Team Roles as a Measure of Business Culture* (2008)
- Swailes & Aritzeta, *Scale Properties of the Team Role Self-Perception Inventory* (2006)
- Swailes & McIntyre-Bhatty, *Uses and Abuses of Reliability Estimates: The Case of the Belbin TRSPI* (2001)
- van Dierendonck & Groen, *Belbin Revisited: The Construct Validity of the Interplace II Team Role Instrument* (2008)

Vitenskapelige artikler relatert til Belbins teammodell

Anderson, N. and Sleaf, S. (2004). 'An evaluation of gender differences on the Belbin Team Role Self-Perception Inventory'. *Journal of Occupational and Organizational Psychology*, **77**, 429–37.

Aritzeta, A., Ayestaran, S. and Swailes, S. (2005a). 'Team role preference and conflict management styles'. *International Journal of Conflict Management*, **16**, 157–82.

Aritzeta, A., Senior, B. and Swailes, S. (2005b). 'Team role preference and cognitive styles: a convergent validity study'. *Small Group Research*, **36**, 404–36.

Arroba, T. and Wedgwood-Oppenheim, F. (1994). 'Do senior managers differ in the public and private sectors?'. *Journal of Managerial Psychology*, **9**, 13–16.

Balderson, S. J. and Broderick, A. J. (1996). 'Behaviour in teams: exploring occupational and gender differences'. *Journal of Managerial Psychology*, **11**, 33–42.

Barker, V. L., Patterson, P. W. and Mueller, G. C. (2001). 'Organizational causes and strategic consequences of the extent of top management team replacement during turnaround attempts'. *Journal of Management Studies*, **38**, 235–70.

Beck, D., Fisch, R. and Bergander, W. (1999). 'Functional roles in work groups – an empirical approach to the study of group role diversity'. *Psychologische Beiträge*, **41**, 288–307.

Belbin, M. (1981). *Management Teams, Why They Succeed or Fail*. London: Heinemann.

Belbin, M. (1993a). *Team Roles at Work*. Oxford: Butterworth-Heinemann.

Belbin, M. (1993b). 'A reply to the Belbin Team Role Self-Perception Inventory by Furnham, Steele and Pendleton'. *Journal of Occupational and Organizational Psychology*, **66**, 259–60.

Broucek, W. G. and Randell, G. (1996). 'An assessment of the construct validity of the Belbin Self-Perception Inventory and Observer's Assessment from the perspective of the five-factor model'. *Journal of Occupational and Organizational Psychology*, **69**, 389–405.

Davis, J., Millburn, P., Murphy, T. and Woodhouse, M. (1992). *Successful Team Building: How To Create Teams That Really Work*. London: Kogan Page.

Dulewicz, V. (1995). 'A validation of Belbin's team roles from 16PF and OPQ using bosses' ratings of competence'. *Journal of Occupational and Organizational Psychology*, **68**, 81–99.

Dulewicz, V. and Higgs, M. (1999). 'Can emotional intelligence be measured and developed?'. *Leadership and Organization Development Journal*, **20**, 242–52.

Fisher, S. G. and Macrosson, W. D. K. (1995). 'Early influences on management team roles'. *Journal of Managerial Psychology*, **10**, 8–15.

Fisher, S. G., Macrosson, W. D. and Sharp, G. (1996). 'Further evidence concerning the Belbin Team Role Self-perception Inventory'. *Personnel Review*, **25**, 61–7.

Fisher, S. G., Hunter, T. A. and Macrosson, W. D. K. (1998a). 'The structure of Belbin's team roles'. *Journal of Organizational and Occupational Psychology*, **71**, 283–8.

Fisher, S. G., Macrosson, W. D. K. and Wong, J. (1998b). 'Cognitive style and team role preference'. *Journal of Managerial Psychology*, **13**, 544–57.

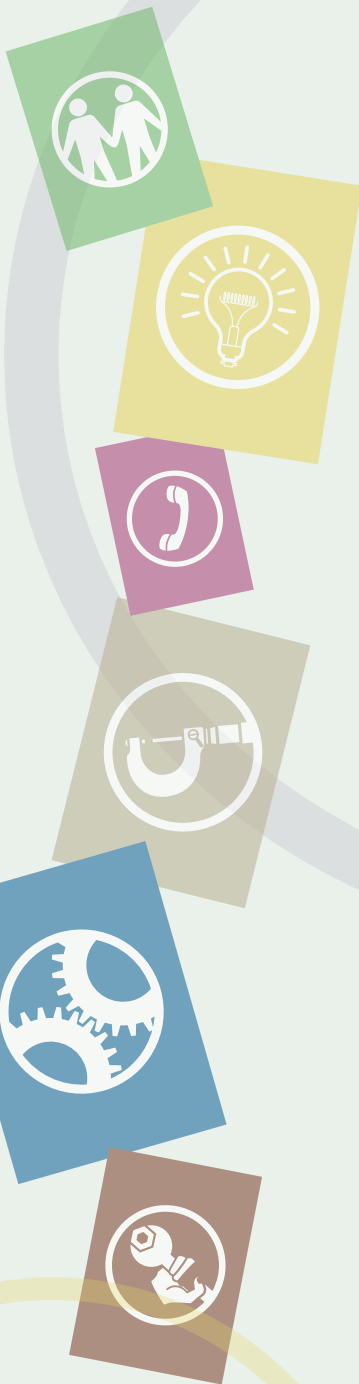
Fisher, S. G., Hunter, T. A. and Macrosson, W. D. K. (2001a). 'A validation study of Belbin's team roles'. *European Journal of Work and Organizational Psychology*, **10**, 121–44.

Fisher, S. G., Macrosson, W. D. and Semple, J. H. (2001b). 'Control and Belbin's team roles'. *Personnel Review*, **30**, 578–88.

Fisher, S. G., Hunter, T. A. and Macrosson, W. D. (2002). 'Belbin's team role theory: for non-managers also?'. *Journal of Managerial Psychology*, **17**, 14–20.

Jackson, C. (2002). 'Predicting team performance from a learning process model'. *Journal of Managerial Psychology*, **17**, 6–13.

- Jarzabkowski, P. and Wilson, D. C. (2002). 'Top teams and strategy in a UK university'. *Journal of Management Studies*, **39**, 355–81.
- Lessem, R. and Baruch, Y. (2000). 'Testing the SMT and Belbin inventories in top management teams'. *Leadership and Organization Development Journal*, **21**, 75–83.
- Macrosson, W. D. and Hemphill, D. J. (2001). 'Machiavellianism in Belbin team roles'. *Journal of Managerial Psychology*, **16**, 355–63.
- O'Doherty, D. M. (2005). 'Working as part of balanced team'. *International Journal of Engineering Education*, **21**, 113–20.
- Park, W.-W. and Bang, H. (2002). 'Team Role Balance and Team Performance'. Paper presented at the Biannual Conference on Belbin's Team Roles, Cambridge, UK.
- Prichard, J. S. and Stanton, N. A. (1999). 'Testing Belbin's team role theory of effective groups'. *Journal of Management Development*, **18**, 652–65.
- Senior, B. (1997). 'Team roles and team performance: is there "really" a link?'. *Journal of Occupational and Organizational Psychology*, **70**, 241–58.
- Senior, B. (1998). 'An empirically-based assessment of Belbin's team roles'. *Human Resource Management Journal*, **8**, 54–60.
- Senior, B. and Swailes, S. (1998). 'A comparison of the Belbin Self Perception Inventory and Observer's Assessment Sheet as measures of an individual's team roles'. *International Journal of Selection and Assessment*, **6**, 1–8.
- Senior, B. and Swailes, S. (2004). 'The dimensions of management team performance: a repertory grid study'. *The International Journal of Productivity and Performance Management*, **53**, 317–33.
- Shi, Y. and Tang, H. K. (1997). 'Team roles behaviour and task environment: an exploratory study of five organizations and their managers'. *Journal of Managerial Psychology*, **12**, 88–94.
- Sommerville, J. and Dalziel, S. (1998). 'Project teambuilding – the applicability of Belbin's team-role self-perception inventory'. *International Journal of Project Management*, **16**, 165–71.
- Swailes, S. and McIntyre-Bhatty, T. (2002). 'The "Belbin" team role inventory: reinterpreting reliability estimates'. *Journal of Managerial Psychology*, **17**, 529–36.
- Swailes, S. and McIntyre-Bhatty, T. (2003). 'Scale structure of the Team Role Self Perception Inventory'. *Journal of Occupational and Organizational Psychology*, **76**, 525–9.
- Watkins, B. and Gibson-Sweet, M. (1997). 'Sailing with Belbin'. *Education + Training*, **39**, 105–10.



BELBIN
NORGE

Belbin Norge AS
+47919 00 601, post@belbin-norge.no
www.belbin-norge.no