

Begrepsvalidering av et utvalg indikatorer i
ledergruppeinventoriet *effect*:

En eksplorerende valideringsstudie

Marie Prebensen Juell



Masteroppgave ved Psykologisk institutt

UNIVERSITETET I OSLO

15.05.2018

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

© Marie Prebensen Juell

2018

Begrepsvalidering av et utvalg indikatorer i ledergruppeinventoriet *effect*:
En eksplorerende valideringsstudie

Marie Prebensen Juell

<http://www.duo.uio.no>

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Formålet med denne studien var å undersøke begrepsvaliditeten til et utvalg indikatorer i ledergruppeinventoriet *effect*. Seksten toppledergrupper i statsforvaltningen i Norge (departementer, direktorater og tilsyn) utgjorde datagrunnlaget for begrepsvalideringen. Begrepsvalideringen ble gjennomført ved å sammenligne resultater på indikatorene i *effect* opp mot andre indikatorer som kan fungere som mål på de samme fenomenene som *effect* hevder å måle. Disse eksterne indikatorene ble utarbeidet ut ifra hva ledergruppene kvalitativt beskrev som sine styrker og utviklingsområder. Det ble tatt i bruk en induktiv tematisk analyse for å identifisere styrker og utviklingsområder. Videre ble det tatt i bruk en deduktiv tematisk analyse for å sammenligne kategorier og temaer fra den induktive analysen opp mot indikatorene i *effect*. Den deduktive tematiske analysen viste at 12 kategorier delte felles teoretiske definisjoner med 12 av indikatorene i *effect*, og utgjorde dermed kategorier som kunne benyttes som eksterne kriterier for å validere indikatorene i *effect*. For å undersøke sammenhengen mellom kategoriene og indikatorene, testet jeg om det var en sammenheng mellom hva ledergrupper i sine kvalitative kommentarer beskrev som styrker og utviklingsområder, og hvordan de skåret på de kvantitative indikatorene i *effect*. Det ble anvendt t-tester for uavhengige utvalg for å undersøke disse sammenhengene. Resultatene fra t-testene indikerer at ledergruppenes skåre på 8 av 12 indikatorer i *effect*, kan representere et reelt nivå at det latente begrepet indikatoren er utviklet for å måle.

Forord

Først og fremst vil jeg takke veileder Henning Bang for mange gode diskusjoner og råd i en krevende, men spennende prosess. Du har bidratt med entusiasme, konstruktive tilbakemeldinger og sist, men ikke minst, et spennende datamateriale til disposisjon.

Jeg vil også rette en stor takk til mine kjære medstudenter. Takk for faglige innspill, støtte og oppmuntring underveis dette året. Avslutningsvis vil jeg takke familie, venner og Eskil for støtte, nyttige innspill og for at dere har holdt ut med mitt fravær den siste tiden.

Oslo, mai 2018

Marie Prebensen Juell

Innholdsfortegnelse

Teoretisk rammeverk.....	2
Ledergrupper	2
En modell for effektivitet i ledergrupper.....	3
Ledergruppeinventoriet <i>effect</i>	3
Validitet	4
Begrepsvaliditet.....	5
En prosedyre for begrepsvalidering	6
Studiens oppbygning	7
Del 1	7
Metode.....	7
Utvalg og datainnsamling.....	7
Induktiv tematisk analyse.....	8
Reliabilitet i kodingsprosessen.....	9
Resultater.....	10
Induktiv tematisk analyse.....	10
Beskrivelse av temaer og kategorier	11
Del 2	16
Metode.....	16
Deduktiv tematisk analyse	16
Resultater.....	16
Deduktiv tematisk analyse	16
Del 3	17
Begrunnelse og hypoteser	17
Metode.....	21
Aggregering av data	21
Statistiske analyser	21
Resultater.....	22
Statistiske analyser	22
Diskusjon.....	26
Implikasjoner.....	29
Studiens begrensninger	30
Anbefalinger til fremtidig forskning	33
Konklusjon	34
Litteraturliste	35
Vedlegg	42
VEDLEGG A	42
VEDLEGG B	44
VEDLEGG C	45
VEDLEGG D	49
VEDLEGG E.....	51

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Begrepsvalidering av et utvalg indikatorer i ledergruppeinventoriet *effect*:

En eksplorerende valideringsstudie

Hvorvidt forskning på individer, grupper og organisasjoner bidrar til økt kunnskap om organisasjonspsykologiske fenomener, er avhengig av kvaliteten på forskningen (Hinkin, 1995). I organisasjonspsykologisk forskning er man ofte interessert i holdninger, opplevelser, fenomener eller synspunkter som ikke er direkte observerbare (Taylor, 2013). Spørreskjema er et måleinstrument som er mye brukt i forskning for å innhente informasjon om slike begreper, og er i tillegg en tids- og kostnadseffektiv måte å drive datainnsamling på (Kraut, 1996; Stone, 1978).

Allerede på 1950-tallet diskuterte American Psychological Association (APA) ulike måter å kvalitetssikre psykologiske tester og måleinstrumenter på. APA-komiteen var hovedsakelig opptatt av validitet, altså hvorvidt måleinstrumentet faktisk måler det som det er ment å måle (Cronbach & Meehl, 1955). Fenomener, opplevelser eller holdninger som ikke er direkte observerbare kan operasjonaliseres og måles gjennom indikatorer. En måling har høy begrepsvaliditet dersom det er samsvar mellom begrepet slik det er definert teoretisk og slikt det er operasjonalisert gjennom indikatorene som benyttes (Kleven, 2002).

I 2012 utviklet organisasjonspsykologene Henning Bang og Thomas Nettet Midelfart ledergruppeinventoriet *effect*. *effect* er et web-basert spørreskjema, utviklet for å måle faktorer som er viktige for kvaliteten på de resultatene ledergruppen oppnår. Spørreskjemaet er basert på Bang og Midelfarts (2012; 2017) forskningsbaserte modell for effektivitet i ledergrupper.

Hvorvidt *effect* faktisk måler de begrepene det intenderer å måle, er avhengig av spørreskjemaets kvalitet, og mer spesifikt – de enkeltledd som til sammen utgjør en indikator, og som benyttes for å måle de ulike begrepene. Indikatorer med god begrepsvaliditet gjør at man kan trekke gyldige slutninger om hvorvidt en ledergruppes skåre på en indikator i *effect* korresponderer med en sann skåre på det underliggende begrepet man ønsker å måle (Messick, 1995; Shadish, Cook & Campbell, 2002). I praksis betyr dette at ledergruppens skåre på eksempelvis indikatoren Relasjonskonflikt bør være en reell representasjon av det nivået av mellommenneskelige spenninger og gnisninger som faktisk eksisterer i ledergruppen.

En validitetsstudie av ledergruppeinventoriet *effect* vil være viktig for organisasjonsforskere som bruker spørreskjemaet i forskning på ledergrupper, ettersom man trenger å godtgjøre at fenomenene man fanger opp gjennom spørreskjemaet faktisk blir målt gjennom de indikatorene som er ment å måle disse fenomenene (Cronbach & Meehl, 1955; Shadish et al., 2002).

En validitetsstudie har også betydning for konsulenter som tar i bruk spørreskjemaet i sitt arbeid med utvikling av ledergrupper. Gitt at *effect* faktisk måler de faktorene som påvirker effektivitet i ledergrupper, vil ledergrupper som tar i bruk *effect* kunne få en realistisk oversikt over ledergruppens styrker og utviklingsområder. En slik kartlegging vil kunne gi et godt utgangspunkt for videre utviklingsarbeid i ledergruppen.

I denne studien ønsker jeg å undersøke hvorvidt indikatorer i *effect* måler de latente begrepene som de er utviklet for å måle. Formålet med studien er derfor å gjøre en begrepsvalidering av et utvalg indikatorer i *effect* ved å undersøke problemstillingen:

Hva slags sammenheng er det mellom ledergruppers kvantitative skårer på indikatorene i effect, og de fenomenene medlemmer av ledergrupper kvalitativt beskriver som gruppens styrker og utviklingsområder?

I det følgende presenteres relevant teori, herunder teori knyttet til ledergrupper, validitetsteori og valideringsmetoden som har blitt brukt i denne studien. Jeg har valgt å dele metode- og resultatdelen av oppgaven i tre deler. I første del gjør jeg rede for hvordan det kvalitative datamaterialet ble kodet i kategorier og temaer som sier noe om hva ledergruppene opplever som sine styrker og utviklingsområder. I andre del beskriver jeg hvordan kategorier fra den kvalitative analysen ble parett opp mot indikatorene i *effect* basert på teoretiske fellestrekk. I tredje del undersøker jeg hvorvidt det er en sammenheng mellom de nye kvalitative kategoriene og indikatorene i *effect*. Til slutt følger en diskusjon av funnene, implikasjoner, studiens begrensninger samt anbefalinger til fremtidig forskning.

Teoretisk rammeverk

Ledergrupper

En ledergruppe kan defineres som «[...] en gruppe ledere som rapporterer til en overordnet leder, som møtes regelmessig, og som oppfatter seg selv som en ledergruppe i organisasjonen» (Bang & Midelfart, 2012, s. 22). En ledergruppe består som oftest av ledere for ulike enheter i en virksomhet, som rapporterer til en seniorleder, eksempelvis administrerende direktør, divisjonsdirektør, eller avdelingsleder (Bang & Midelfart, 2012).

Forskning har vist at ledergrupper kan ha stor betydning for suksess og resultater i organisasjonen (Barrick, Bradley, Kristof-Brown & Colbert, 2007; Hambrick & Mason, 1984). Ledergruppen kan bidra til verdiskapingen i organisasjonen ved at man informerer hverandre, diskuterer problemstillinger, koordinerer ulike mål og prosesser, fatter

beslutninger og løser problemer (Bang & Midelfart 2012; Wageman & Hackman, 2010). Hvert av medlemmene i ledergruppen har gjerne ansvar for resultatskapning i sin egen enhet eller avdeling. De er som oftest ikke valgt inn basert på kunnskap som må til for å løse ledergruppens utfordringer, men på bakgrunn av den kompetansen de har til å lede enheten de er ansvarlig for (Bang & Midelfart 2012; McIntyre, 1998).

En modell for effektivitet i ledergrupper

Ifølge Wageman, Nunes, Burruss og Hackman (2008) er en ledergruppe effektiv i den grad den oppnår sine mål ovenfor virksomheten, skaper motivasjon, bidrar til utvikling for medlemmene, samt øker sin evne til å samarbeide fremover. Hvorvidt en ledergruppe er effektiv kan derfor vurderes ut ifra tre nivåer; individnivå, gruppenivå og organisasjonsnivå (Hackman, 2002).

I 2012 publiserte organisasjonspsykologene Henning Bang og Thomas Nettet Midelfart en modell for effektivitet i ledergrupper. Målet var å presentere et helhetlig bilde av de ulike faktorene som påvirker effektivitet i ledergrupper (Bang & Midelfart, 2012; Bang & Midelfart, 2017). Modellen er en klassisk input-process-output (IPO) modell, og flere forskere tar i bruk denne modelltypen i teamforskning (for eksempel McGrath, 1964; Hackman, 2002). Modellen er forskningsbasert og bygger på internasjonal teamforskning, samt Bang og Midelfarts egen forskning på ledergrupper, andre typer arbeidsgrupper og team som ligner ledergrupper i kompleksitet, slik som beslutnings- og problemløsningsgrupper (Bang, 2008; Bang, 2010, Bang & Midelfart, 2012; Bang & Midelfart, 2017).

Ledergruppeinventoriet *effect*

effect er et selvrapporingsskjema designet for å måle faktorer som er viktige for hvilke resultater ledergruppen oppnår. *effect* måler samtlige av de 27 faktorene i Bang og Midelfarts (2012) modell for effektivitet i ledergrupper. Av de 27 er 19 faktorer kategorisert som input- og prosessfaktorer, altså faktorer som påvirker de resterende åtte outputfaktorene.

De 27 faktorene er representert ved 27 indikatorer bestående av mellom to og syv enkeltledd. Disse besvares på en syvpunkts Likert-skala, fra 1 = helt uenig, 4 = verken uenig eller enig, til 7 = helt enig. Spørreskjemaet består av til sammen 124 ledd. En høy gjennomsnittsskåre på en indikator tilsier at medlemmene er enige i at denne indikatoren i stor grad karakteriserer gruppen, mens en lav gjennomsnittsskåre tilsier at indikatoren i liten grad karakteriserer gruppen.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

I tillegg til de 124 leddene består *effect* av fem åpne spørsmål: 1) Hvilke resultater skaper ledergruppen som organisasjonen har nytte av? 2) Hva fungerer spesielt bra i ledergruppen? 3) Hva kan forbedres i måten ledergruppen fungerer på? 4) Hva mener du er det viktigste man kan gjøre for å få denne ledergruppen til å fungere bedre? og 5) Hvordan tror du andre i organisasjonen oppfatter ledergruppen?

Analyser av *effect* viser at de 27 indikatorene har tilfredsstillende reliabilitet, hvor samtlige av indikatorene har en Cronbachs alfa på over .70 (Bang, 2017). Det er gjennomført en studie av faktorstrukturen til 20 av indikatorene i ledergruppeinventoriet *effect*, som indikerte tilfredsstillende diskriminant og konvergent validitet for disse 20 indikatorene (Hansen, 2017). Det er imidlertid ikke gjort noen valideringsstudier av sammenhengen mellom skårene på indikatorene i *effect* og eksterne indikatorer som er tiltenkt å måle de samme fenomenene som *effect* hevder at det gjør.

Validitet

Mennesker prøver å forstå verden rundt seg ved å ta i bruk tolkninger, konklusjoner og slutninger. Når man tolker, konkluderer eller trekker en slutning, så gjør man dette basert på tilgjengelig og observerbar informasjon (Taylor, 2013). For å kunne trekke en gyldig slutning er man avhengig av at observerbar informasjonen er en reell representasjon av virkeligheten. Validitet sier noe om gyldigheten til slutninger eller tolkninger (Shadish et al., 2002). En slutning eller tolkning er valid i den grad den sier noe om det faktiske forholdet mellom begreper man ønsker å undersøke (DeVellis, 2012; Lund & Haugen, 2006). Dette betyr at et resultat eller en metode ikke innehar validitet i seg selv. Derimot er det slik at validitet sier noe om gyldigheten til slutningene eller konklusjonene man foretar, basert på resultatene man besitter og metodene som er tatt i bruk for å oppnå resultatene (Cook & Campbell, 1979). Eksempelvis har ikke spørreskjemaet *effect* validitet i seg selv, men *slutningene* man trekker på grunnlaget av resultatene fra spørreskjemaet, kan ha validitet i varierende grad.

Det eksisterer flere former for validitet (Cook & Campbell, 1979; Cronbach & Meehl, 1955; Messick, 1995). I 1979 postulerte Cook og Campbell et validitetssystem, som i senere tid har blitt videreutviklet av Shadish et al. (2002). Validitetssystemet til Shadish et al. (2002) er i dag et anerkjent validitetssystem (Haugen & Lund, 2006), og blir derfor brukt som teoretisk utgangspunkt i denne studien. Validitetssystemet består av fire forskjellige typer slutninger, og deres respektive validiteter (se tabell 1).

Tabell 1

Validitetssystemet til Shadish et al. (2002)

Fire former for validitet

Statistisk slutningsvaliditet

Gyldigheten av statistiske slutninger, altså slutningen om en sammenheng mellom to variabler.

Indre validitet

Gyldigheten av en kausal slutning, altså om observert samvariasjon mellom to variabler reflekterer et årsaksforhold der variabel A fører til en endring i variabel B.

Begrepsvaliditet

Gyldigheten av en slutning om at indikatorer representerer bestemte begreper.

Ytre validitet

Gyldigheten av slutninger om at det antatte forholdet mellom to variabler kan generaliseres til andre settinger, tidsrom og populasjoner.

Begrepsvaliditet

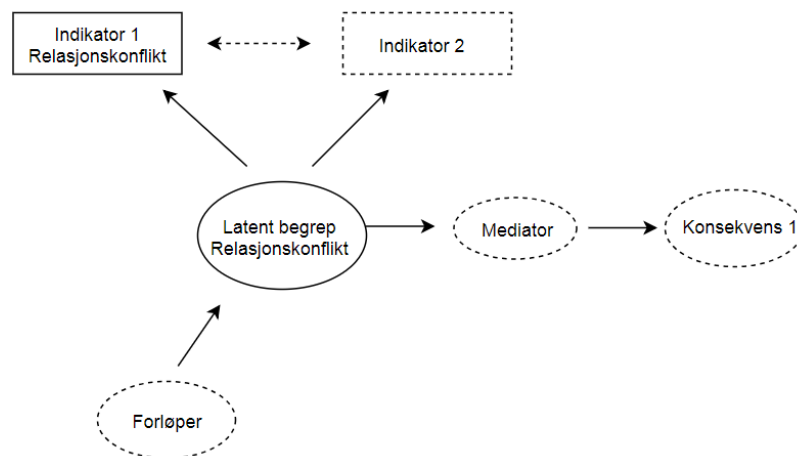
I tråd med oppgavens formål har jeg valgt å fokusere på begrepsvaliditet. Jeg vil i det følgende redegjøre nærmere for denne formen for validitet.

I psykologisk forskning er man ofte interessert i å studere og måle komplekse begreper som ikke er direkte observerbare, eksempelvis intelligens og personlighet (Cronbach & Meehl, 1955). For å måle slike begreper er det nødvendig å operasjonalisere dem (Kleven, 2002; Shadish et al., 2002). Når man utvikler spørreskjemaer for å måle begreper, vil operasjonaliseringen av et begrep gjerne innebære at man utformer et sett med enkeltledd som til sammen skal representere det teoretiske begrepet man ønsker å måle. Disse leddene vil samlet utgjøre en indikator som dermed skal være en representasjon av det latente begrepet (MacKenzie, Podsakoff & Podsakoff, 2011). Det er imidlertid slik at operasjonalisering av et begrep i seg selv, ikke nødvendigvis gir god begrepsvaliditet. Derfor er det viktig å undersøke hvorvidt slutninger om forholdet mellom begrepsindikatoren og det underliggende begrepet er gyldige (Lund, 2002). I testpsykologien betyr dette at en respondents skåre på en indikator bør representere en reell skåre på det underliggende begrepet man ønsker å måle (Messick, 1995). Å undersøke begrepsvaliditeten til et spørreskjema vil være å undersøke i hvilken grad operasjonaliseringen av et teoretisk begrep korresponderer til det faktiske begrepet av interesse (Machery, 2007).

En prosedyre for begrepsvalidering

Campbell og Fiske (1959) hevder at begrepsvaliditet kan undersøkes ved å bruke flere mål og ulike metoder for å undersøke det samme begrepet. Begrepsvaliditet kan undersøkes ved å sammenligne resultater fra et sett med indikatorer opp mot et eksternt sammenligningsgrunnlag. Man kan for eksempel se på sammenhengen mellom begrepet av interesse, lignende begreper og ulike indikatorer som er ment å representere det samme begrepet (MacKenzie et al., 2011). Slike undersøkelser baserer seg på diskriminant og konvergent validitet. Konvergent validitet undersøkes ved at indikatorer som er tiltenkt å måle det samme latente begrepet skal korrelere sterkt med hverandre. Diskriminant validitet undersøkes ved at indikatorer som ikke er tiltenkt å måle det samme latente begrepet skal korrelere svakt med hverandre (Shadish et al., 2002).

Ifølge Cronbach og Meehl (1955) har ethvert begrep et nomologisk nettverk. Et begreps nomologiske nettverk består av andre begreper eller indikatorer som på en eller annen måte er relatert til begrepet. Relasjonen kan variere, fra å være en forløper til begrepet, en konsekvens av begrepet, eller en annen indikator som korrelerer med begrepet (MacKenzie et al., 2011).



Figur 1. En illustrasjon av et nomologisk nettverk med *effect*-indikatoren Relasjonskonflikt som eksempel, inspirert av MacKenzie et al. (2011).

Et begreps nomologiske nettverk utvider seg ettersom begrepet forskes videre på. Et relativt nytt begrep vil ha et lite nomologisk nettverk, men etter hvert som mengden forskning på begrepet økes, vokser nettverket. Prosedyren for å undersøke det nomologiske nettverket til et begrep er å undersøke forholdet mellom begrepet og antatte korrelater, eksempelvis en

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

ny indikator (MacKenzie et al., 2011). Å undersøke forholdet mellom to indikatorer som er ment å representere det samme underliggende begrepet, vil derfor være en måte å undersøke begrepsvaliditet på. Adcock og Collier (2001) påpeker at en måte å gjennomføre en begrepsvalidering på, er å undersøke om en gitt indikator er empirisk assosiert med andre indikatorer av samme begrep, ettersom de er forventet å korrelere på bakgrunn av teoretiske fellestrekk.

I denne oppgaven er det respondentenes kvalitative beskrivelser av ledergruppens styrker og utviklingsområder som blir benyttet som eksterne indikatorer, og utgjør dermed sammenligningsgrunnlaget for indikatorene i *effect*.

Studiens oppbygning

Jeg har valgt å dele metode- og resultatdelen inn i tre deler. I første del gjør jeg rede for fremgangsmåten jeg har anvendt for å komme frem til de fenomenene ledergruppene kvalitativt beskriver som sine styrker og utviklingsområder. Resultatene fra denne delen danner grunnlaget for analysen i del to. I andre del beskrives prosessen som gikk ut på å sammenligne ledergruppens beskrivelser av styrker og utviklingsområder opp mot et utvalg av indikatorene i *effect*. Funnene fra studiens andre del danner grunnlaget for studiens siste analyse, der jeg undersøker hvorvidt det er en sammenheng mellom hva ledergruppene kvalitativt beskriver som styrker og utviklingsområder og ledergruppens skårer på indikatorene i *effect*. Til slutt diskuterer jeg hvorvidt de utvalgte indikatorene i *effect* faktisk måler de latente begrepene de er ment å måle.

Del 1

I første del av oppgaven skal jeg gjøre rede for studiens utvalg og fremgangsmåten som er tatt i bruk for å identifisere ledergruppers styrker og utviklingsområder. Deretter presenteres resultatet av den tematiske analysen, samt en beskrivelse av hvert tema og kategori.

Metode

Utvalg og datainnsamling

Datautvalget består av 16 toppledergrupper i statsforvaltningen i Norge (departementer, direktorater og tilsyn). Toppledergruppene deltok i et 9 måneder langt ledergruppeutviklingsprogram i regi av Direktoratet for forvaltning og IKT (Difi). Datainnsamlingen foregikk i 2017, og dataen ble innhentet av Henning Bang og Difi.

Innsamlingen skjedde i forkant av at ledergruppene skulle starte på programmet. Samtlige 16 ledergrupper har gitt samtykke til bruk av dataene fra sine *effect*-rapporter til forskningsrelaterte formål. Forskningsprosjektet «Effektivitet i ledergrupper» som denne studien er en del av, er registrert og godkjent av NSD.

Hvert enkelt medlem i toppledergruppene har svart på ledergruppeinventoriet *effect*. Toppledergruppens skåre på de 27 faktorene er presentert i en rapport som gjennomsnittet av medlemmenes vurdering på de leddene som utgjør faktoren. Hvert enkelt medlem har også svart på fem åpne spørsmål, hvorav to av spørsmålene er utgangspunktet for denne studien: «Hva fungerer spesielt bra i ledergruppen?» og «Hva kan forbedres i måten ledergruppen fungerer på?» Svarene på disse spørsmålene gir et uttrykk for hva medlemmene opplever som ledergruppens styrker og utviklingsområder.

Induktiv tematisk analyse

En induktiv tematisk analyse er en kvalitativ metode som benyttes for å identifisere kategorier og temaer uavhengig av teori (Guest, MacQueen & Namey, 2012). Jeg har brukt tematisk analyse for å identifisere temaer og kategorier som representerer medlemmenes egne beskrivelser av hva ledergruppen deres er særlig sterk på, og hva ledergruppen kan utvikle seg på. Innholdet i disse kategoriene er derfor deskriptivt fremfor tolkende. Jeg har valgt å benytte meg av en induktiv tematisk analyse, ettersom Sandelowski og Barroso (2003) anbefaler å ta i bruk enten innholdsanalyse eller tematisk analyse hvis behovet for tolkning er relativt lite.

I denne oppgaven vil jeg bygge på Braun og Clarkes (2006) metodikk for tematisk analyse. Jeg har også latt meg inspirere av måten Berger og Harr (2013) benyttet tematisk analyse for å identifisere ulike formål med ledergrupper. Følgende fremgangsmåte ble anvendt:

Steg 1. Bli kjent med datasettet. Sidene med de kvalitative utsagnene fra to av spørsmålene ble tatt ut fra de 16 opprinnelige rapportene, og lest gjennom flere ganger. Spørsmålene som datamaterialet er basert på er: 1) Hva fungerer spesielt bra i ledergruppen? og 2) Hva kan forbedres i måten ledergruppen fungerer på?

Steg 2. Identifisere meningsenheter. Identifiseringen av meningsenhetene er basert på Boyatzis (1998) definisjon av begrepet: «Det mest grunnleggende segmentet, eller elementet av rådata eller informasjon som kan vurderes på en meningsfull måte» (Boyatzis, 1998, s. 64 *oversettelse av Berger og Harr, 2013*). Inndelingen i meningsenheter ble gjort for å forenkle kodingsprosessen. For å opprettholde ledergruppenes anonymitet, er setninger der

ledergruppens tilhørighet ble nevnt, byttet ut med *virksomheten* eller *organisasjonen*. Se vedlegg A for prosedyren jeg har benyttet for identifisering av meningsenheter.

Steg 3. Utvikle koder. Hver enkelt meningsenhet ble klassifisert i en kode. Kodene ble utviklet ved at meningsenheter som ikke passet inn i allerede eksisterende koder, ble tildelt en ny kode i tråd med en induktiv tilnærming (Boyatzis, 1998; Saldaña, 2009). Det ble laget ett sett med koder for styrker, og ett sett med koder for utviklingsområder. Se vedlegg B for mer informasjon om kodingsprosessen.

Steg 4. Identifisere kategorier og temaer. Kodene ble vurdert ut ifra innhold, sjargong og repetisjoner (Ryan & Bernard, 2003) og kategorisert i ulike kategorier. Eksempelvis ble kodene «god stemning» og «god tone» plassert i samme kategori. Kategorier med beslektet meningsinnhold ble slått sammen og overordnede temaer ble identifisert. Det ble tatt i bruk papirlapper bestående av navnet på én kode per lapp, for sortering og utprøving av koder i mulige kategorier og temaer (Braun & Clarke, 2006).

Steg 5. Gjennomgang og revurdering av koder, kategorier og temaer. Etter gjennomgang av samtlige koder, kategorier og temaer, ble prosessen gjennomgått på nytt, ettersom kodingsprosessen i en tematisk analyse er en iterativ prosess (Braun & Clarke, 2006). Flere koder ble slettet underveis, noen ble slått sammen og antall kategorier og temaer ble redusert. Målet med dette steget var å sørge for at ingen koder kunne plasseres i mer enn ett tema, og at innholdet i kodene innenfor hvert tema hadde et meningsfullt forhold til hverandre.

Steg 6. Navngi og definere temaer og kategorier. Siste steg i den tematiske analysen var å navngi og definere kategoriene og temaene. Tema- og kategorinavn ble inspirert av ord og utsagn som respondentene selv benyttet seg av.

Reliabilitet i kodingsprosessen

Ved bruk av tematisk analyse kan reliabilitet defineres som «konsistens innenfor observasjoner, koding eller tolkning» (Boyatzis, 1998, s.144, *min oversettelse*). I kvalitativ forskning er det vanlig å benytte seg av inter-rater reliabilitet som et estimat på konsistens i analysen (Guest et al., 2012; Joffe & Yardley, 2004). Inter-rater reliabilitet ble undersøkt ved at en ekstern koder parete et tilfeldig utvalg meningsenheter opp mot kategoriene. Deretter ble disse sammenlignet med mine egne kodinger av de samme meningsenhetene. Grad av samsvar viste en Cohen's Kappa på .76 som indikerer høy grad av enighet (Landis & Koch, 1977).

I tillegg ble det gjennomført en test-retest etter anbefaling fra Joffe og Yardley (2004). Jeg kodet et tilfeldig utvalg meningsenheter inn i allerede eksisterende kategorier. For å redusere sannsynligheten for overføringseffekter i henhold til hukommelse, gjennomførte jeg testen 3 måneder etter den tematiske analysen var ferdigstilt. Grad av samsvar mellom test-retest viste en Cohen's Kappa på .97 som indikerer en nesten perfekt overenstemmelse (Landis & Koch, 1977).

Resultater

Induktiv tematisk analyse

I det følgende presenteres funnene fra den tematiske analysen. Analysen baserte seg på totalt 796 meningsenheter, hvorav 749 ble kodet inn i meningsbærende koder, mens 47 ble kodet i en residualkategori, derav en total residualprosent på 5,9 %.

Den tematiske analysen resulterte i 19 kategorier, dekket av 9 temaer. 7 temaer ble identifisert på tvers av styrker og utviklingsområder, ett tema var særegent for utviklingsområder og ett tema var kun å finne blant styrker. De 7 temaene på tvers av styrker og utviklingsområder er: 1.Relasjoner og fellesskap, 2.Trygghet og åpenhet, 3.Beslutninger, 4.Saker, 5.Diskusjonskvalitet, 6.Ledermøter og sakspapirer, og 7.Ledergruppens sammensetning. Temaet som kun ble nevnt i forbindelse med ledergruppens utviklingsområder er Ledergruppens formål, mens temaet Kompetanse kun ble nevnt i sammenheng med hva som fungerer bra i ledergrupper.

Det temaet som opptrer hyppigst når medlemmene blir bedt om å beskrive hva som fungerer bra i ledergrupper, er Relasjoner og fellesskap (14 % av meningsenhetene). Når det gjelder utviklingsområder, fremgår det av analysen at temaet Saker ser ut til å oppta majoriteten av toppledergruppene i studien. Seksten prosent av meningsenhetene går ut på å beskrive hvordan arbeidet med ledergruppens saker kan utvikles og forbedres, og 15 av 16 ledergrupper trekker fram dette temaet som et utviklingsområde. Dette gjenspeiler seg i resultatene fra spørsmålet om ledergruppens styrker, hvor kun 3 % av meningsenhetene ble kodet inn under temaet Saker. I tabell 2 fremgår det en oversikt over temaene og hvor stor prosentandel av totalt antall meningsenheter som tilhører de ulike temaene.

Tabell 2

Presentasjon av temaer og frekvensfordeling av meningsenheter

Temaer	Styrke % av menings- enheter styrker	Utviklingsområde % av meningsenheter utviklingsområder	Totalt % av totalt antall meningsenheter
1. Relasjoner og fellesskap	14 %	9 %	23 %
2. Saker	3 %	16 %	19 %
3. Diskusjonskvalitet	9 %	6 %	15 %
4. Ledermøter og sakspapirer	8 %	7 %	15 %
5. Trygghet og åpenhet	9 %	2 %	11 %
6. Kompetanse	5 %	-	5 %
7. Ledergruppens sammensetning	1 %	4 %	5 %
8. Beslutninger	2 %	2 %	4 %
9. Ledergruppens formål	-	3 %	3 %
			= 100 %

Bemerk. Prosentandelen er regnet ut basert på antall meningsenheter og ikke respondenter, da respondentene kan uttale seg om flere temaer i samme utsagn.

Beskrivelse av temaer og kategorier

I det følgende presenteres temaene med tilhørende kategorier, samt hva som er essensen i hvert tema og hver kategori. Temaene er presentert i uthevet skrift, mens kategoriene er presentert i kursiv. Formålet med studien er å begrepsvalidere et utvalg indikatorer i *effect* ved å sammenligne indikatorene opp mot de fenomener ledergruppene kvalitativt beskriver som gruppens styrker og utviklingsområder. Dette medfører at det primære fokuset i denne delen av oppgaven er på det teoretiske innholdet i temaene og kategoriene, og ikke hvor ofte et tema ble nevnt. Kategoriene er definert ut ifra kodene som utgjør kategorien. Jeg anså det som unødvendig å definere hver enkelt kode, ettersom kodene er så spesifikke og ligger tett opp mot hver enkelt meningsenhet. Se vedlegg C for full oversikt over temaer, kategorier, koder og antall meningsenheter per kode.

Tema 1: Relasjoner og fellesskap. Relasjoner og fellesskap omhandler i hvilken grad ledergruppen oppleves som en samkjørt ledergruppe fremfor en gruppe enkeltledere, og i

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

hvilken grad medlemmene vil hverandre vel, samarbeider og bidrar til det felles beste. Temaet består av fire kategorier:

Fellesskap og samhold. Kategorien Fellesskap og samhold defineres som i hvilken grad gruppen fokuserer på fellesskapet fremfor individet, ser hele organisasjonsbildet og ikke bare egen avdeling/enhet, og hvor lojalitet ovenfor gruppen og dens medlemmer står sterkt. Kategorien kan illustreres av følgende utsagn: «Et fellesskap der vi kan diskutere, få forståelse [...]» (styrke) og «Hver mann er en øy» (utviklingsområde).

Samarbeid. Kategorien defineres som i hvilken grad man hjelper hverandre i og utenfor ledermøtene, bidrar til gruppen og foster et inkluderende miljø gjennom godt samarbeid. Det handler ikke bare om samarbeid under ledermøtene, men også samarbeid på tvers av avdelinger. Kategorien kan illustreres av følgende utsagn: «Ledergruppen fungerer også godt i en-til-en diskusjoner utenom ledermøtene, der vi støtter hverandre og gir hverandre innspill» (styrke) og «Vi bør jobbe for å forbedre samhandlingen i konfliktsaker» (utviklingsområde).

Relasjoner. Kategorien defineres som kvaliteten på de mellommenneskelige forholdene i ledergruppen, og omhandler i hvilken grad gruppen preges av konflikter, utfordringer knyttet til enkeltpersoner, respekt og/eller trivsel. Kategorien kan illustreres av følgende utsagn: «Stort sett har jeg en opplevelse av at man vil hverandre vel – at jeg skal lykkes med mine oppgaver» (styrke) og «Vi er nokså upersonlige» (utviklingsområde).

Stemning. Kategorien defineres som hva slags stemning som preger ledermøtene. En god stemning kjennetegnes av en hyggelig og uformell tone hvor latteren sitter løst, og det er en konstruktiv tilnærming til arbeidsoppgavene. Kategorien kan illustreres med følgende utsagn: «Lett tone og mye godt humør» (styrke) og «Ledergruppen må holde et positivt fokus, og medlemmene må ikke snakke ledergruppen ned» (utviklingsområde).

Tema 2: Saker. Temaet omhandler hvorvidt sakene som diskuteres i ledergruppen er relevante og viktige, om de hører hjemme i ledergruppen, og om sakene blir tatt opp på et hensiktsmessig tidspunkt i prosessen. Temaet består av to kategorier:

Diskutere riktige saker. Diskutere riktige saker defineres som i hvilken grad medlemmene opplever at saker diskutert i ledermøtene er relevante for ledergruppen og at saksfremleggene belyser sakene på en god måte. Det tydeliggjøres hvorfor sakene blir tatt opp i ledergruppen, og eventuelt hvilke saker om bør diskuteres et annet sted i organisasjonen. Kategorien kan illustreres ved følgende utsagn: «I hovedsak relevante saker på sakslisten» (styrke) og «Diskutere de store steinene før vi diskuterer grus og sand» (utviklingsområde).

Tidspunkt for saker. Kategorien defineres som hvorvidt tidspunktet en sak diskuteres på er til fordel for videre saksgang. Saker som diskuteres på et for tidlig stadium kan fort bli generelle, mens saker som kommer inn for sent, er ofte avgjort på forhånd, eller prosessen har gått så langt at det er vanskelig å få til endringer. Kategorien kan illustreres av følgende utsagn: «Få inn større saker på en tidligere tidspunkt» (utviklingsområde). Det var ingen utsagn der Tidspunkt for saker ble beskrevet som en styrke.

Tema 3: Diskusjonskvalitet. Temaet omhandler måten ledergruppen diskuterer og utveksler informasjon på. Temaet består av to kategorier:

Tydelige diskusjoner. Kategorien defineres som i hvilken grad ledermøtene preges av tydelig kommunikasjon i diskusjoner med konkrete og ryddige innspill med lite avsporinger. Kategorien kan illustreres av følgende utsagn: «[...] lite utenomsnakk» (styrke) og «For mange høytflyvende innlegg» (utviklingsområde).

Åpen meningsutveksling. Kategorien defineres som i hvilken grad ulike synspunkter kommer frem i diskusjonen, hvorvidt medlemmene lytter til hverandre, og er interessert i andres meninger og perspektiver. Åpen meningsutveksling handler også om at medlemmer tør å utfordre hverandre i diskusjoner. Kategorien kan illustreres av følgende utsagn: «Ulike meninger og synspunkter kommer frem og drøftingene blir gode» (styrke) og «Enda mer rom for å innta "djevelens advokat-rollen", og utfordre hverandre enda mer uten at det går på bekostning av de positive verdiene med gode relasjoner» (utviklingsområde).

Tema 4: Ledermøter og sakspapirer. Dette temaet omhandler gjennomføring av ledergruppens møter, tilhørende forberedelser, papirarbeid samt møteledelse. Temaet består av tre kategorier:

Gjennomføring av ledermøtene. Kategorien defineres som i hvilken grad ledermøtene har møteagendaer og oppsummeringer, og at tidsfordelingen per sak opprettholdes samt at det skrives referater for hvert møte. Kategorien kan illustreres av følgende utsagn: «Struktur og rutiner rundt møteavvikling» (styrke) og «Møter kan til tider bli mer effektive» (utviklingsområde).

Møteforberedelser og sakspapirer. Kategorien defineres som hvor godt hver enkelt medlem er forberedt før ledermøter, og hvor godt sakene og de tilhørende sakspapirene er forberedt. I hvilken grad blir sakspapirer sendt ut i god tid før møtet, slik at medlemmer får mulighet til å lese gjennom og være forberedt til ledermøtet? Inneholder sakspapirene nok informasjon som tydeliggjør hva saken dreier seg om? Kategorien kan illustreres av følgende utsagn: «Godt forberedte saker og innledere» (styrke) og «Det at få har rukket å lese papirene gjør at diskusjonene ikke blir så forberedt som de kunne ha vært» (utviklingsområde).

Møteledelse. Denne kategorien defineres som i hvilken grad lederen av ledergruppen utøver tydelig møteledelse, oppsummerer på slutten av møtet, oppfordrer til deltakelse og er tydelig i sine beslutninger. Kategorien kan uttrykkes av følgende utsagn: «God og tydelig møteledelse fra ledergruppens leder skaper en god ramme» (styrke). Ingen respondenter trakk frem Møteledelse som et utviklingsområde.

Tema 5: Trygghet og åpenhet. Trygghet og åpenhet defineres som i hvilken grad ledergruppen preges av politisk spill, hvorvidt takhøyden for såkalte «dumme» innspill er høy, og om ledergruppen er et sted der åpenhet stimuleres. Temaet består av to kategorier:

Politisk spill. Kategorien defineres som i hvilken grad gruppen preges av skjulte agendaer, maktkamper, koalisjoner og taktikkeri. Politisk spill innebærer også at saker og beslutninger er avgjort på forhånd eller avgjort utenfor ledermøtene. Kategorien kan illustreres av følgende utsagn: «Det er ingen skjulte agendaer» (styrke) og «Det er ulike koalisjoner» (utviklingsområde).

Trygg og åpen atmosfære. Kategorien defineres som i hvilken grad det eksisterer en atmosfære der medlemmer kan legge frem utfordringer av faglig og personlig art, og kan ytre meninger uten å være redd for negative konsekvenser. I en åpen og trygg atmosfære er det lav terskel for ulike synspunkter, og det er liten grad av forutinntatte holdninger blant medlemmene. Kategorien kan illustreres av følgende utsagn: «Vi tør å fremstå som menneskelige» (styrke) og «Litt mer interesse for hverandre som mennesker for å skape tillitt» (utviklingsområde).

Tema 6: Kompetanse. Temaet omhandler i hvilken grad medlemmene besitter faglig kompetanse og hvorvidt medlemmene lærer av hverandre ved å dele kunnskap. Temaet består av to kategorier:

Kompetanse. Kategorien defineres som i hvilken grad medlemmene besitter nødvendig kompetanse og kunnskap for å bidra til faglige diskusjoner og problemløsning. Kategorien kan illustreres av følgende utsagn: «Høy fagkompetanse på de fleste områder» (styrke). Ingen respondenter trakk frem Kompetanse som et utviklingsområde.

Kompetansedeling. Kategorien defineres som i hvilken grad medlemmene deler erfaringer, kompetanse og kunnskap som bidrar til læring i gruppen. Kategorien kan illustreres av følgende utsagn: «Vi lærer av hverandre» (styrke). Ingen respondenter trakk frem Kompetansedeling som et utviklingsområde.

Tema 7: Ledergruppens sammensetning. Temaet defineres som kvaliteten på sammensetningen av gruppens medlemmer, både når det gjelder gruppestørrelse og i hvilken grad medlemmene utfyller hverandre. Temaet består av to kategorier:

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Størrelse. Kategorien defineres som hvorvidt gruppen oppleves som passe stor. Kategorien kan illustreres av følgende utsagn: «Gruppestørrelsen er slik at diskusjonene flyter godt» (styrke) og «Stor ledergruppe gjør at diskusjonene tar lang tid» (utviklingsområde).

Mangfold. Kategorien defineres som i hvilken grad medlemmene i ledergruppen representerer et fagfelt eller ansvarsområde som komplementerer hverandre, og hvorvidt medlemmene har innsikt i hverandres fagområder. Kategorien kan illustreres av følgende utsagn: «Medlemmene utfyller hverandre med tanke på bakgrunnskunnskap» (styrke) og «Vi jobber med ulike fagområder og det preger gruppen» (utviklingsområde).

Tema 8: Beslutninger. Temaet defineres som kvaliteten på beslutninger med tanke på hvorvidt det er klart for medlemmene av gruppen når beslutninger blir tatt, på hvilket grunnlag de blir tatt og hvilken beslutning som er fattet. Temaet har ingen kategorier. Temaet kan illustreres av følgende utsagn: «[...] og konklusjonene er tydelige» (styrke) og «Tydelighet i beslutninger, spesielt de som er krevende å gjennomføre» (utviklingsområde).

Tema 9: Ledergruppens formål. Dette temaet omhandler hvorvidt medlemmer av ledergruppen er klar over ledergruppens funksjon i organisasjonen, hvilke saker som hører hjemme i ledergruppen og hvilke forventninger som stilles til ledergruppen og dens medlemmer. Temaet består av to kategorier:

Formålsavklaring. Kategorien defineres som i hvilken grad medlemmene opplever at det er avklart hvilke forventninger som stilles til ledergruppen og at det er en felles enighet angående ledergruppens formål og funksjon. Kategorien kan illustreres av følgende utsagn: «Det er uklart om man ønsker en strategisk ledergruppe eller en administrativ ledergruppe» (utviklingsområde). Ingen respondenter trakk frem Formålsavklaring som en styrke.

Saksavklaring. Kategorien defineres som i hvilken grad det er tydelig for ledergruppens medlemmer hva slags saker ledergruppen skal behandle og hvorvidt ledergruppen diskuterer hvilke saker som er de riktige sakene å diskutere. Kategorien kan illustreres av følgende utsagn: «Vi må avklare hvilke saker som skal diskuteres i ledergruppen» (utviklingsområde). Ingen respondenter trakk frem Saksavklaring som en styrke.

Del 2

I følgende del skal jeg gjøre rede for hvordan jeg gikk frem for å avgjøre hvilke av indikatorene i *effect* som utgjør utvalget for begrepsvalideringen.

Metode

Deduktiv tematisk analyse

I en deduktiv tematisk analyse anvendes en ovenfra-og-ned (top-down) tilnærming. En slik analyse benyttes hvis formålet er å identifisere mønstre eller likheter ut ifra et allerede eksisterende rammeverk (Boyatzis, 1998). Jeg benyttet en deduktiv tematisk analyse for å se etter likheter mellom temaene og kategoriene fra den induktive analysen og de 27 indikatorene som utgjør ledergruppeinventoriet *effect*.

Temaene og kategoriene fra den induktive tematiske analysen ble parett opp mot de 27 indikatorene i *effect* med Bang og Midelfarts (2012; 2017) modell for effektivitet i ledergrupper som teoretisk rammeverk. Først ble de teoretiske definisjonene til samtlige av *effect*-indikatorene gjennomlest. Videre ble innholdet i de kvalitative temaene og kategoriene gjennomlest i lys av definisjonene til indikatorene i *effect*. Fellestrekk mellom kategorier og temaer og *effect*-indikatorer ble identifisert, og avslutningsvis ble temaer og kategorier parett opp mot hver sin *effect*-indikator. Inter-rater reliabilitet ble undersøkt ved at en ekstern koder leste definisjonene til temaene, kategoriene og indikatorene i *effect*, for så å ta i bruk informasjonen til å pare kategoriene og temaene fra den tematiske analysen opp mot indikatorene i *effect*. Cohen's Kappa var på .82 som tilsier høy grad av enighet (Banerjee, Capozzoli, McSweeney & Sinha, 1999).

Resultater

Deduktiv tematisk analyse

Den deduktive tematiske analysen viste at 2 temaer og 10 kategorier fra den induktive tematiske analysen har teoretiske fellestrekk med 12 av indikatorene i *effect*. Dette muliggjør en begrepsvalidering av 12 *effect*-indikatorer i denne studien. Se vedlegg D for en full oversikt over tema- og kategoridefinisjonene sammenlignet med definisjonene til de 12 *effect*-indikatorene, hentet fra *effect* (Bang & Midelfart, 2012; 2017).

For å lette lesingen, har jeg videre i oppgaven valgt å omtale de 2 temaene og de 10 kategoriene som deler fellestrekk med indikatorene i *effect* som kategorier. Se tabell 3 for resultatet av den deduktive analysen.

Tabell 3

Resultatet av den deduktive tematiske analysen.

Kvalitative kategorier	<i>effect</i> -indikatorer
1. Ledergruppens formål	1. Klart formål
2. Saker	2. Riktige saker
3. Størrelse	3. Riktig størrelse
4. Kompetanse	4. Riktig kompetanse
5. Møteforberedelser og sakspapirer	5. Møteforberedelser
6. Tydelige diskusjoner	6. Fokuseret kommunikasjon
7. Relasjoner	7. Fravær av relasjonskonflikt
8. Politisk spill	8. Fravær av politisk spill
9. Åpen meningsutveksling	9. Dialog
10. Møteledelse	10. Effektiv teamledelse
11. Trygg og åpen atmosfære	11. Gruppepsykologisk trygghet
12. Fellesskap og samhold	12. Lagånd

Del 3

Studiens siste del omhandler hvorvidt det er en sammenheng mellom kategoriene og indikatorene i *effect*. Først skal jeg gjøre rede for de ulike fremgangsmåtene jeg har benyttet og hvilke hypoteser jeg skal undersøke. Deretter presenteres metodedelen hvor det fremgår hvilke analyser jeg har tatt i bruk for å undersøke en eventuell sammenheng mellom indikatorer og kategorier. Avslutningsvis presenteres resultatene, etterfulgt av en diskusjon av begrepsvalideringen.

Begrunnelse og hypoteser

For å undersøke hvorvidt det er en overenstemmelse mellom kvalitative og kvantitative resultater har jeg anvendt følgende fremgangsmåter:

Den første fremgangsmåten undersøker sammenhengen mellom en kategori og en indikator basert på følgende antagelse: Ledergrupper der minst ett medlem kvalitativt beskriver en kategori som en styrke ved ledergruppen, skårer høyere på tilsvarende *effect*-indikator sammenlignet med ledergrupper der minst ett medlem kvalitativt beskriver den samme kategorien som et utviklingsområde. I de tilfellene hvor en kvalitativ kategori ble

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

betegnet både som en styrke og et utviklingsområde innad i en ledergruppe, er det antall medlemmer som avgjør hvorvidt en kvalitativ kategori blir betraktet som noe som fungerer bra eller noe som kan forbedres i ledergruppen. For eksempel beskrev fire medlemmer i samme ledergruppe kategorien Trygg og åpen atmosfære som en styrke ved ledergruppen. Derimot var det ett medlem i samme ledergruppe som kvalitativt påpekte en Trygg og åpen atmosfære som et utviklingsområde. Det kan tenkes at en Trygg og åpen atmosfære er en styrke ved ledergruppen, ettersom majoriteten av medlemmene i ledergruppen kvalitativt beskrev en Trygg og Åpen atmosfære som en styrke. Se vedlegg E for en oversikt over hvor mange av medlemmene per ledergruppe som hadde utsagn kodet inn i de ulike kategoriene.

Antagelsen om at ledergrupper der minimum ett medlem kvalitativt beskriver en kategori som en styrke ved ledergruppen, skårer høyere på tilsvarende *effect*-indikator enn ledergrupper der minimum ett medlem kvalitativt beskriver den samme kategorien som et utviklingsområde, gir følgende hypoteser:

H1: Ledergrupper som kvalitativt beskriver Møteforberedelser og sakspapirer som en styrke, skårer høyere på effect-indikatoren Møteforberedelser enn ledergrupper som kvalitativt beskriver Møteforberedelser og sakspapirer som et utviklingsområde.

H2: Ledergrupper som kvalitativt beskriver Tydelige diskusjoner som en styrke, skårer høyere på effect-indikatoren Fokuset kommunikasjon enn ledergrupper som kvalitativt beskriver Tydelige diskusjoner som et utviklingsområde.

H3: Ledergrupper som kvalitativt beskriver Relasjoner som en styrke, skårer høyere på effect-indikatoren Fravær av relasjonskonflikt enn ledergrupper som kvalitativt beskriver Relasjoner som et utviklingsområde.

H4: Ledergrupper som kvalitativt beskriver fravær av Politisk spill som en styrke, skårer høyere på effect-indikatoren Fravær av politisk spill enn ledergrupper som kvalitativt beskriver Politisk spill som et utviklingsområde.

H5: Ledergrupper som kvalitativt beskriver Åpen meningsutveksling som en styrke, skårer høyere på effect-indikatoren Dialog enn ledergrupper som kvalitativt beskriver Åpen meningsutveksling som et utviklingsområde.

H6: Ledergrupper som kvalitativt beskriver Trygg og åpen atmosfære som en styrke, skårer høyere på effect-indikatoren Gruppepsykologisk trygghet enn ledergrupper som kvalitativt beskriver Trygg og åpen atmosfære som et utviklingsområde.

H7: Ledergrupper som kvalitativt beskriver Fellesskap og samhold som en styrke, skårer høyere på effect-indikatoren Lagånd enn ledergrupper som kvalitativt beskriver Fellesskap og samhold som et utviklingsområde.

Sammenligningen mellom kategorien Størrelse og *effect*-indikatoren Riktig størrelse skiller seg ut fra hypotesene som er nevnt over, ettersom en skåre på 4 i *effect* tilsier at ledergruppen oppleves som passe stor. En skåre på over 4 for indikatoren Riktig størrelse, indikerer at medlemmene opplever at ledergruppen er for stor, mens skårer under 4 indikerer at medlemmene opplever gruppen som for liten. For å undersøke forskjeller i skårer mellom ledergrupper som kvalitativt beskrev størrelsen på ledergruppen som en styrke og ledergrupper som kvalitativt beskrev ledergruppens størrelse som et utviklingsområde, setter jeg opp følgende hypotese:

H8: Ledergrupper som kvalitativt beskriver Størrelse som en styrke, skårer nærmere 4 (gruppen oppleves passe stor) på effect-indikatoren Riktig størrelse enn ledergrupper som kvalitativt beskriver Størrelse som et utviklingsområde.

Ingen av medlemmene i de 16 ledergruppene trakk frem kategoriene Møteledelse og Kompetanse som utviklingsområder ved ledergruppen. I disse to tilfellene blir ledergruppens skårer på tilsvarende indikatorer (Effektiv teamledelse og Riktig Kompetanse) sammenlignet ut i fra antagelsen om at ledergrupper som ikke kvalitativt beskriver en kategori som en styrke, skårer lavere på *effect*-indikatoren enn ledergrupper som kvalitativt beskriver en kategori som en styrke ved ledergruppen. Dette gir følgende hypoteser:

H9: Ledergrupper som kvalitativt beskriver Møteledelse som en styrke, skårer høyere på effect-indikatoren Effektiv teamledelse enn ledergrupper som ikke påpeker Møteledelse som en styrke ved ledergruppen.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

H10: Ledergrupper som kvalitativt beskriver Kompetanse som en styrke, skårer høyere på effect-indikatoren Riktig kompetanse enn ledergrupper som ikke påpeker Kompetanse som en styrke ved ledergruppen.

Kategorien Ledergruppens formål ble kun beskrevet som noe som kan forbedres i ledergruppen. I dette tilfellet blir ledergruppenes skårer på indikatoren Klart formål sammenlignet ut i fra antagelsen om at ledergrupper som kvalitativt beskriver en kategori som et utviklingsområde, skårer lavere på tilsvarende *effect*-indikator enn ledergrupper som ikke kvalitativt beskriver en kategori som et utviklingsområde. Dette gir følgende hypotese:

H11: Ledergrupper som kvalitativt beskriver Ledergruppens formål som et utviklingsområde, skårer lavere på effect-indikatoren Klart formål enn ledergrupper som ikke påpeker Ledergruppens formål som et utviklingsområde.

Det fremgår av resultatene til den tematiske analysen at majoriteten av ledergruppene i studien er opptatt av hvorvidt de diskuterer de riktige sakene til riktig tid. Funnene viser at 15 av 16 ledergrupper kvalitativt påpekte Saker som utviklingsområde. Videre var det bare én ledergruppe som var fornøyd med ledergruppens Saker. I tillegg viser de kvalitative dataen at noen ledergrupper skiller seg ut fra andre ledergrupper ved at en stor andel av gruppens medlemmer beskrev Saker som et utviklingsområde.

I dette tilfellet har jeg valgt å undersøke sammenhengen mellom kategorien Saker og indikatoren Riktige saker ved å sammenligne *effect*-skårene til ledergrupper der mange av medlemmene (tre eller flere) beskrev saker som et utviklingsområde, med *effect*-skårene til ledergrupper der kun ett eller to av medlemmene i gruppen beskrev Saker som et utviklingsområde. Dette gir hypotesen:

H12: Ledergrupper der tre eller flere av medlemmene kvalitativt beskriver Saker som et utviklingsområde, skårer lavere på effect-indikatoren Riktige saker enn ledergruppene der kun ett til to medlemmer kvalitativt beskriver Saker som et utviklingsområde.

Metode

Aggregering av data

effect-skårene er aggregerte skårer. Det vil si at skåren på en indikator er gjennomsnittet av hvert enkelt ledergruppemedlems skåre på indikatoren. I Bang (2017) fremgår det at de 27 indikatorene i *effect* viser tilfredsstillende grad av enighet mellom medlemmene innad i en ledergruppe, målt med r_{WG} (within-group rater agreement). For de 12 indikatorene som benyttes i denne studien, viser Bang (2017) at r_{WG} varierer mellom .59 og .84. Ifølge Biemann, Cole og Voelpel (2012) og LeBreton og Senter (2008) indikerer en r_{WG} koeffisient fra .00 til .30 «mangel på enighet», .31 til .50 en «svak enighet», .51 til .70 en «moderat enighet», .71 til .90 en «sterk enighet» og .91 til 1 indikerer en «veldig sterk enighet» blant medlemmene innad i ledergruppen. Det vil si at indikatorene som benyttes i denne studien har en «moderat til sterk enighet» mellom medlemmene, som rettferdiggjør bruken av aggregerte skårer.

Statistiske analyser

En t-test for uavhengige utvalg kan benyttes for å undersøke hvorvidt det eksisterer en signifikant forskjell i gruppegjennomsnitt mellom to utvalg (Howell, 2013; Welch, 1958). En forutsetning for å benytte seg av denne type t-test, er at ledergruppene som sammenlignes må være uavhengige av hverandre (Field, 2013). Eksempelvis kan ikke ett medlem tilhøre to ledergrupper, og respondenter skal ha svart på *effect* uten påvirkning fra andre ledergrupper. Kravet om uavhengighet er derfor tilfredsstillt i denne studien.

Utvalget i studien var på $n = 16$. I tillegg er det slik at ledergruppene beskrev ulike styrker og utviklingsområder i de åpne kommentarene i *effect* som medførte en svært liten utvalgsstørrelse i majoriteten av t-testene. For eksempel var det kun fire ledergrupper som kvalitativt beskrev fravær av Politisk spill som en styrke, mens tre ledergrupper beskrev Politisk spill som et utviklingsområde.

Jeg har likevel valgt å benytte meg av t-test for uavhengige utvalg, ettersom Winter (2013) påpeker at en t-test for uavhengige utvalg er utviklet med forbehold om små utvalgsstørrelser ($n \leq 5$). Ved liten utvalgsstørrelse ($n \leq 20$) anbefaler Stevens (1996) å justere signifikansnivået fra $p \leq .05$ til et eksempelvis $p \leq .10$ og $p \leq .15$. Årsaken er at en liten utvalgsstørrelse kan medføre ikke-signifikante resultater på grunn av manglende statistisk styrke, selv om det faktisk foreligger en forskjell i gruppegjennomsnitt for de to utvalgene. Dermed kan en liten utvalgsstørrelse øke sannsynligheten for å begå en type II feil, som er å feilaktig bekrefte nullhypotesen (Howell, 2013). Samtlige av analysene ble utført i SPSS v25.

Resultater

Statistiske analyser

Hypotese 1 - Møteforberedelser. En t-test for uavhengige utvalg ble gjennomført for å sammenligne skårene på *effect*-indikatoren Møteforberedelser, mellom ledergrupper som kvalitativt trakk frem Møteforberedelser og sakspapirer som en styrke og ledergrupper som kvalitativt trakk frem Møteforberedelser og sakspapirer som et utviklingsområde.

t-testen viste en signifikant forskjell ($p \leq .01$) i skårene på indikatoren Møteforberedelser mellom ledergrupper som kvalitativt beskrev Møteforberedelser og sakspapirer som en styrke ($M = 5.63$, $SD = .31$) og ledergrupper som kvalitativt beskrev Møteforberedelser og sakspapirer som et utviklingsområde ($M = 4.58$, $SD = .28$; $t(9) = 5.75$, $p = .00$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på 1.05 (95% *CI*: .63 til 1.46). Hypotese 1 støttes.

Hypotese 2 - Fokusert kommunikasjon. Det var ingen signifikant ($p > .15$) forskjell i skårer på indikatoren Fokusert kommunikasjon mellom ledergrupper som kvalitativt beskrev Tydelige diskusjoner som en styrke ($M = 5.56$, $SD = .52$) og ledergrupper som kvalitativt beskrev Tydelige diskusjoner som et utviklingsområde ($M = 5.1$, $SD = .0$; $t(5) = .92$, $p = .19$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på .46 (95 % *CI*: -.6 til 1.3). Hypotesen støttes ikke.

Hypotese 3 - Fravær av relasjonskonflikt. Det var en signifikant forskjell ($p \leq .01$) i skårer på indikatoren Fravær av relasjonskonflikt mellom ledergrupper som kvalitativt beskrev Relasjoner som en styrke ($M = 6.03$, $SD = .51$) og ledergrupper som kvalitativt beskrev Relasjoner som et utviklingsområde ($M = 3.96$, $SD = .3$; $t(10) = 6.42$, $p = .00$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på 2.06 (95 % *CI*: 1.35 til 2,78). Hypotese 3 støttes.

Hypotese 4 - Fravær av politisk spill. Det var en signifikant forskjell ($p \leq .15$) i skårer på indikatoren Fravær av politisk spill mellom ledergrupper som kvalitativt påpekte lite Politisk spill som en styrke ($M = 5.72$, $SD = .26$) og ledergrupper som kvalitativt påpekte Politisk spill som et utviklingsområde ($M = 5.03$, $SD = .73$; $t(5) = 1.78$, $p = .135$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på .69 (85 % *CI*: .03 til 1.35). Hypotese 4 støttes.

Hypotese 5 - Dialog. Det var en signifikant forskjell ($p \leq .15$) i skårer på indikatoren Dialog mellom ledergrupper som kvalitativt beskrev Åpen meningsutveksling som en styrke ($M = 5.26$, $SD = .39$) og ledergrupper som kvalitativt beskrev Åpen meningsutveksling som et

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

utviklingsområde ($M = 4.88$, $SD = .51$; $t(14) = 1.64$, $p = .12$, to-halet).

Gjennomsnittsforskjellen mellom de to utvalgene var på $.37$ (85 % CI : $.02$ til $.72$). Hypotese 5 støttes.

Hypotese 6 - Gruppepsykologisk trygghet. Det var en signifikant forskjell ($p \leq .10$) i skårer på indikatoren Gruppepsykologisk trygghet mellom ledergrupper som kvalitativt beskrev Åpen og trygg atmosfære som en styrke ($M = 5.7$, $SD = .55$) og ledergrupper som kvalitativt beskrev Åpen og trygg atmosfære som et utviklingsområde ($M = 4.9$, $SD = .8$; $t(12) = 1.85$, $p = .09$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på $.80$ (90 % CI : $.03$ til 1.61). Hypotese 6 støttes.

Hypotese 7 - Lagånd. Det var ingen signifikant forskjell ($p > .15$) i skårer på indikatoren Lagånd for ledergrupper som kvalitativt beskrev Fellesskap og samhold som en styrke ($M = 5.44$, $SD = .21$) og ledergrupper som kvalitativt beskrev Fellesskap og samhold som et utviklingsområde ($M = 5.27$, $SD = .69$; $t(12) = .64$, $p = .53$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på $.16$ (95% CI : $-.39$ til $.71$). Hypotese 7 støttes ikke.

Hypotese 8 - Riktig størrelse. Det var en signifikant forskjell ($p \leq .10$) i skårer på indikatoren Riktig størrelse mellom ledergrupper som kvalitativt beskrev Størrelse som en styrke ($M = 4$, $SD = 0$) og ledergrupper som kvalitativt beskrev Størrelse som et utviklingsområde ($M = 4.75$, $SD = 0.07$; $t(1) = -8.66$, $p = 0.07$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på $.75$ (90 % CI : -1.2 til $-.20$). Hypotese 8 støttes.

Hypotese 9 - Effektiv teamledelse. Det var ingen signifikant forskjell ($p > .15$) i skårer på indikatoren Effektiv teamledelse mellom ledergrupper som kvalitativt påpekte Møteledelse som en styrke ($M = 5.8$, $SD = .33$) og ledergrupper som ikke påpekte Møteledelse som en styrke eller et utviklingsområde ($M = 5.75$, $SD = .47$; $t(14) = .2$, $p = .83$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på $.04$ (95% CI : $.2$ til $-.41$). Hypotese 9 støttes ikke.

Hypotese 10 - Riktig kompetanse. Det var ingen signifikant forskjell ($p > .15$) i skårer på indikatoren Riktig kompetanse mellom ledergrupper som kvalitativt beskrev Kompetanse som en styrke ($M = 5.62$, $SD = .4$) og ledergrupper som ikke påpekte Kompetanse som en styrke ($M = 5.44$, $SD = .79$; $t(14) = .2$, $p = .83$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på $.04$ (95 % CI : $-.41$ til $.49$). Hypotese 10 støttes ikke.

Hypotese 11 - Klart formål. Det var en signifikant forskjell ($p \leq .05$) i skårer på indikatoren Klart formål mellom ledergrupper kvalitativt beskrev Ledergruppens formål som et utviklingsområde ($M = 5.13, SD = .47$) og ledergrupper som ikke påpekte Ledergruppens formål som et utviklingsområde ($M = 5.9, SD = .54; t(13) = 2.64, p = .02$, to-halet). Gjennomsnittsforskjellen mellom de to utvalgene var på .77 (95% *CI*: .13 til 1.38). Hypotese 11 støttes.

Hypotese 12 - Riktige saker. Det var en signifikant forskjell ($p \leq .10$) i skårer på indikatoren Riktige saker mellom ledergrupper hvor tre eller flere av medlemmene kvalitativt beskrev Saker som et utviklingsområde ($M = 4.77, SD = .38$) og ledergrupper der ett til to av medlemmene beskrev Saker som et utviklingsområde ($M = 5.22, SD = .43; t(13) = 2.05, p = .06$, to – halet). Gjennomsnittsforskjellen mellom de to utvalgene var på .45 (90% *CI*: .06 til .83). Hypotese 12 støttes. I tabell 4 fremgår det en oversikt over resultatene for samtlige av t-testene.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Tabell 4

Resultater fra-tester med deskriptiv statistikk

H	effect-indikator	Utvalg									
		LG- Styrker			LG-Utviklingsområder			95% CI	t	df	p-verdi
M	SD	n	M	SD	n						
1	Møteforberedelser	5.63	.31	6	4.58	.28	5	.63, 1.46	5.75	9	.00
2	Fokusert kommunikasjon	5.56	.52	5	5.1	.0	2	-.6, 1.3	.92	5	.19
3	Fravær av relasjonskonflikt	6.03	.51	9	3.96	.3	3	1.35, 2.78	6.42	10	.00
4	Fravær av politisk spill	5.72	.26	4	5.03	.73	3	*.03, 1.35	1.78	5	.13
5	Dialog	5.26	.39	10	4.88	.51	6	*.02, .72	1.64	14	.12
6	Gruppepsykologisk trygghet	5.7	.55	12	4.9	.8	2	** .03, 1.61	1.85	12	.09
7	Lagånd	5.44	.21	5	5.27	.69	9	-.39, .71	.64	12	.53
8	Riktig størrelse	4	.0	1	4.75	.07	2	** -1.2, -.20	-8.6	1	.07
		LG-Styrker			LG-ikke påpekt som styrke			95% CI	t	df	p-verdi
		M	SD	n	M	SD	n				
9	Effektiv teamledelse	5.8	.33	7	5.75	.47	9	.2, -.41	.2	14	.83
10	Riktig kompetanse	5.62	.4	11	5.44	.79	5	-.41, .49	.2	14	.83
		LG-Utviklingsområde			LG-ikke påpekt som utviklingsområde			95% CI	t	df	p-verdi
		M	SD	n	M	SD	n				
11	Klart formål	5.13	.47	11	5.9	.54	4	.13, 1.38	2.64	13	.02
		LG-Utviklingsområde (1-2 medlemmer)			LG-Utviklingsområde (3 eller flere medlemmer)			95% CI	t	df	p-verdi
		M	SD	n	M	SD	n				
12	Riktige saker	5.22	.43	5	4.77	.38	11	** .06, .83	2.05	13	.06

* 85 % CI

** 90 % CI

Diskusjon

Formålet med studien er å begrepsvalidere et utvalg indikatorer i *effect*. Funnene fra t-testen indikerer støtte for en sammenheng mellom 8 av indikatorene fra *effect* og deres respektive kvalitative kategorier. De åtte indikatorene er: Riktige saker, Klart formål, Riktig størrelse, Gruppesykologisk trygghet, Dialog, Fravær av politisk spill, Fravær av relasjonskonflikt og Møteforberedelser. Dette tyder på at forskjeller i ledergruppens skårer på indikatorene er i overensstemmelse med forskjeller i ledergruppens kvalitative utsagn. Dette illustreres med følgende eksempel: Gjennomsnittsforskjellen i skårer på indikatoren Fravær av relasjonskonflikt er på 2.06 mellom ledergrupper som kvalitativt beskrev Relasjoner som en styrke ($M = 6.03$) og ledergrupper som kvalitativt beskrev Relasjoner som et utviklingsområde ($M = 3.96$). En gjennomsnittsskåre på 3.96 indikerer at ledergruppene opplever konflikter og mellommenneskelige gnisninger i ledergruppen, som er i tråd med ledergruppens kvalitative utsagn. Videre indikerer en gjennomsnittsskåre på 6.03 at ledergruppene kjennetegnes av fravær av relasjonskonflikt, som er i tråd med de kvalitative utsagnene for disse ledergruppene. Overensstemmelse mellom kvalitative og kvantitative resultater tyder på at en skåre på *effect*-indikatoren Fravær av relasjonskonflikt kan være en reell representasjon av det nivået av mellommenneskelige spenninger og gnisninger som faktisk eksisterer i ledergruppen.

Messick (1975) hevder at målet med å undersøke begrepsvaliditet er å gjøre rede for konsistens i atferd eller svarrespons hos respondenter. I denne studien fant jeg en sammenheng mellom ledergruppens kvantitative skårer på åtte av indikatorene i *effect* og de fenomenene medlemmene kvalitativt beskrev som ledergruppens styrker og utviklingsområder. Det betyr at det var konsistens i respondentenes svarrespons på de kvantitative og kvalitative spørsmålene i *effect*. Et slikt funn tyder på at ledergruppens skårer på de åtte indikatorene kan representere reelle nivåer av de latente begrepene indikatorene er operasjonalisert for å måle (Machery, 2007; Messick, 1995).

Resultatene fra t-testen viser imidlertid ikke støtte for en sammenheng mellom fire av indikatorene fra *effect* og deres respektive kvalitative kategorier, noe som kan indikere manglende begrepsvaliditet for disse indikatorene. De fire indikatorene er: Riktig kompetanse, Effektiv teamledelse, Fokuseret kommunikasjon og Lagånd. Det kan imidlertid også være forhold ved selve fremgangsmåten for begrepsvalideringen som kan ha påvirket resultatet.

Ifølge resultatene fra den tematiske analysen var det ingen av ledergruppene som kvalitativt beskrev Kompetanse som et utviklingsområde. Ettersom ingen av ledergruppene

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

beskrev Kompetanse som et utviklingsområde, er hypotesen som ligger til grunn for t-testen formulert slik at ledergrupper som kvalitativt beskrev Kompetanse som en styrke, forventes å skåre høyere på indikatoren Riktig kompetanse sammenlignet med ledergrupper som verken beskrev kompetanse som en styrke eller et utviklingsområde (fem ledergrupper).

Resultatet fra t-testen viser at disse fem ledergruppene ikke skårer signifikant lavere på indikatoren Riktig kompetanse, enn ledergruppene som kvalitativt beskrev Kompetanse som en styrke. En mulig forklaring kan være at fravær av kvalitative kommentarer som omhandler Kompetanse i ledergruppen, ikke nødvendigvis er ensbetydende med en opplevelse av lite Kompetanse. For eksempel kan ledergruppene oppleve at andre aspekter ved gruppen fungerer spesielt bra, og velger derfor å fremheve de aspektene som styrker i det åpne spørsmålet i *effect*, fremfor Kompetanse. Dette kan være en årsak til at ledergruppene som ikke valgte å beskrive Kompetanse som en styrke, likevel skårer tett opp mot ledergruppene som valgte å beskrive Kompetanse som en styrke. En annen årsak kan være det at kategorien Kompetanse kun innebærer faglig kompetanse, mens indikatoren Riktig kompetanse også favner mellommenneskelig kompetanse, i tillegg til faglig og intellektuell kompetanse. Dette kan tyde på at indikatoren og kategorien ikke representerer det samme latente begrepet, som kan være en forklaring på manglende samsvar mellom indikatoren og kategorien.

Videre var det ingen av ledergruppene som kvalitativt påpekte ledelsen av ledergruppen som et utviklingsområde. Dermed ble ledergruppenes skåre på indikatoren Effektiv teamledelse sammenlignet mellom ledergrupper som kvalitativt påpekte ledelsen som en styrke ved ledergruppen og ledergrupper som valgte å ikke beskrive ledelsen i de åpne spørsmålene. Resultatet fra t-testen viser en ikke signifikant forskjell i gjennomsnitt mellom de to utvalgene ledergrupper.

De teoretiske definisjonene til indikatoren Effektiv teamledelse og kategorien Møteledelse noe ulike hverandre. Majoriteten av utsagnene som utgjør kategorien Møteledelse omhandler hvorvidt lederen er «dyktig» eller «flink». Det er lite informasjon om selve atferden til lederen. Hvilke handlinger utfører lederen som fører til at medlemmene opplever lederen som dyktig? Innholdet i enkeltleddene som utgjør indikatoren Effektiv teamledelse går ut på å kartlegge i hvilken grad ledergruppens leder utfører handlinger som er viktige for resultatene ledergruppen oppnår (Bang, 2017). Det kan derfor tenkes at kategorien Møteledelse og indikatoren Effektiv teamledelse ikke representerer det samme latente begrepet. Dette kan således være årsaken til manglende stadfesting av en sammenheng mellom indikatoren og kategorien.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Når ingen av ledergruppene kvalitativt beskrev Kompetanse og Møteledelse som et utviklingsområde er det vanskelig å si med sikkerhet hvorvidt det er en sammenheng mellom de kvalitative og kvantitative resultatene. Således indikerer funnene at man skal være forsiktig med å trekke for vide slutninger om hva indikatorene Riktig kompetanse og Effektiv teamledelse måler, inntil man har flere valideringsstudier som kan si noe om hvorvidt slutningene man trekker på bakgrunn av skårene på disse to indikatorene er valide (Lund & Haugen, 2002; Shadish et al., 2002).

Jeg fant ikke støtte for en sammenheng mellom kategorien Tydelige diskusjoner og indikatoren Fokuset kommunikasjon. Ledergruppene som kvalitativt beskrev Tydelige diskusjoner som en styrke ved ledergruppen skåret ikke signifikant høyere på indikatoren Fokuset kommunikasjon, sammenlignet med ledergrupper som kvalitativt beskrev Tydelige diskusjoner som et utviklingsområde. Dette kan tyde på at indikatoren Fokuset kommunikasjon ikke representerer det latente begrepet det er ment å måle. På den annen siden var det bare var to ledergrupper som kvalitativt påpekte Tydelige diskusjoner som et utviklingsområde, derav en svært lav utvalgsstørrelse i t-testen. I tillegg skårer de to ledergruppene helt likt på indikatoren Fokuset kommunikasjon. Det betyr at det ikke er noe varians i skårene, som impliserer et standardavvik på 0 og fravær av et normalfordelt utvalg. Dette er faktorer som øker sjansen for å begå type II feil, som er å feilaktig beholde nullhypotesen mens det egentlig er en signifikant forskjell i gjennomsnittskårer (Stevens, 2002). Dette betyr at det er en mulighet for at en t-test med et større utvalg ledergrupper hadde gitt et annet resultat, men dette er vanskelig å si med sikkerhet. Det bør derfor vises forsiktighet med å trekke for vide slutninger om hva indikatoren Fokuset kommunikasjon faktisk måler, inntil man har flere studier som sier noe om begrepsvaliditeten til denne indikatoren.

Videre fant jeg ikke støtte for en sammenheng mellom indikatoren Lagånd og kategorien Fellesskap og Samhold. Dette er noe overraskende ettersom kategorien og indikatoren deler felles teoretiske definisjoner, som kan tyde på at indikatoren Lagånd har dårlig begrepsvaliditet. Det kan også tenkes at manglende samvariasjon mellom kategorien og indikatoren kan forklares ut ifra følgende forhold ved studien:

For det første er det en forutsetning at de kvalitative utsagnene er reelle representasjoner av ledergruppens styrker og utviklingsområde. En svakhet ved å anvende en induktiv tematisk analyse er økt risiko for subjektivitet i kodingsprosessen (Boyatzis, 1998). For det andre bør den kvalitative kategorien representere det samme latente begrepet som indikatoren i *effect* hevder at den måler. For det tredje er formuleringen av det åpne

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

spørsmålet som utgjør ledergruppens utviklingsområder, utformet slik at ledergruppene nødvendigvis ikke svarer ut ifra hva som er deres absolutte svakheter, men heller ut i fra hva som kan forbedres i ledergruppen.

Underveis i begrepsvalideringen ble det tatt forhåndsregler for å unngå at slike faktorer hadde muligheten til å påvirke begrepsvalideringen. For å sørge for konsistens i den tematiske analysen, ble det blant annet gjennomført en test-retest av et utvalg meningsenheter med 3 måneders mellomrom. I tillegg ble inter-rater reliabilitet undersøkt ved at en ekstern koder kodet et utvalg meningsenheter inn i allerede identifiserte kategorier. Cohen's Kappa for overenstemmelse var tilfredsstillende for begge tilfeller (.97 og .76), som tyder på konsistens i kodingsprosessen (Landis & Koch, 1977). Dessuten undersøkte jeg hvorvidt en ekstern koder var enig i resultatene fra den deduktive tematiske analysen, herunder paring av kategoriene opp mot indikatorene i *effect*. Inter-rater reliabilitet var tilfredsstillende med en Cohens's Kappa på .82 (Landis & Koch, 1977). Det ble ikke tatt hensyn til spørsmålsformuleringen av det åpne spørsmålet som utgjør ledergruppens utviklingsområder.

De teoretiske definisjonene til kategorien Fellesskap og Samhold og indikatoren Lagånd har et meningsfullt innhold til hverandre. Derfor kan det tenkes at indikatoren Lagånd ikke er operasjonalisert på en hensiktsmessig måte, som fører til at en respondents skåre på indikatoren ikke representerer et reelt nivå av det underliggende begrepet man ønsker å måle (Messick, 1995). Av den grunn bør det vises ekstra forsiktighet ved å trekke slutninger basert på ledergruppens skåre på indikatoren Lagånd, ettersom det er usikkert hvorvidt en skåre representerer det faktiske nivået av i hvilken grad en ledergruppe er sammensveiset, hvor godt medlemmene liker hverandre, inntar et helhetsperspektiv og er stolte av ledergruppen.

Funnene i denne studien indikerer at 8 av 12 indikatorer i *effect* evner å måle det reelle nivået av det latente begrepet indikatoren var designet for å måle, men dette er imidlertid vanskelig å si med sikkerhet. Peter (1981) påpeker at en enkelt studie ikke er nok til å etablere hvorvidt slutninger har god begrepsvaliditet. Videre hevder Cronbach (1971) at begrepsvalidering er en langvarig prosess som krever en kombinasjon av flere og ulike reliabilitet- og validitetsstudier.

Implikasjoner

Resultatene fra begrepsvalideringen kan være til nytte for forskere som tar i bruk ledergruppeinventoriet *effect* i sin forskning på ledergrupper, ettersom funnene i denne studien tyder på at 8 av indikatorene måler det de er ment for å måle. I tillegg indikerer

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

funnene at man skal være forsiktig med å trekke for vide slutninger om hva indikatorene Effektiv teamledelse, Riktig kompetanse, Fokuseret kommunikasjon og Lagånd faktisk måler, inntil man har flere studier som sier noe om begrepsvaliditeten til disse fire indikatorene.

Den eksplorative metoden for begrepsvalidering kan være til inspirasjon for forskere og studenter. Spesielt for forskere som vil benytte seg av en kombinasjon av metoder (mixed method design) for å undersøke begrepsvaliditet. Til min kjennskap er det ingen andre studier der det fremgår en steg for steg oppskrift i å begrepsvalidere et spørreskjema ved å ta i bruk respondenters egne utsagn som et eksternt sammenligningsgrunnlag. Forskere eller studenter som ønsker å gjennomføre en lignende valideringsprosedyre, kan la seg inspirere av fremgangsmåten postulert i denne studien.

Funnene kan også ha betydning for konsulenter som tar i bruk *effect* i sitt arbeid med utvikling av ledergrupper. Konsulenter kan i enda større grad være sikker på at en skåre på en indikator i *effect* faktisk gjenspeiler det reelle nivået av det underliggende begrepet indikatoren er ment å måle. Videre kan ledergrupper som tar i bruk *effect* kjenne seg igjen i resultatene fra rapporten *effect* genererer, og dermed oppleve utviklingsarbeidet som meningsfylt.

Det fremgår av den tematiske analysen at toppledergrupper i statsforvaltningen i Norge opplever hvorvidt ledergruppen arbeider med de riktige sakene som sitt største utviklingsområde. Videre viser resultatene av den tematiske analysen at toppledergrupper har behov for å diskutere strategiske saker i større grad enn det de allerede gjør. Slik informasjon kan være til nytte for konsulenter som arbeider med å utvikle toppledergrupper. Likeledes kan informasjonen være til nytte for toppledergrupper og andre ledergrupper for å øke bevisstheten rundt hvilke saker ledergruppen bør diskutere og hvorvidt ledergruppen opplever å ha et tilfredsstillende fokus på strategiske saker.

Studiens begrensninger

Studiens begrensninger diskuteres i lys av faktorer som kan ha påvirket resultatet av begrepsvalideringen. Dette er: kvantifisering av data, spørsmålene som ligger til grunn for de kvalitative kategoriene, spørreskjemaets utforming, bruk av tematisk analyse og utvalget av ledergrupper.

Kvantifisering av kvalitative data. Jeg valgte å sammenligne ledergruppenes skårer på indikatorene i *effect* ut i fra tilstedeværelsen eller fravær av en kvalitativ kategori. Det vil si at i *hvor stor grad* en ledergruppe opplever et fenomen som en styrke eller et utviklingsområde ikke ble tatt hensyn til i sammenligningen mellom kvantitative og

kvalitative data. En forenklet kvantifisering av kvalitative data kan medføre en reduksjon av meningsinnholdet (Driscoll, Appiah-Yeboah, Salib & Rupert, 2007). For å unngå å eliminere graden av styrken i et meningsutsagn, anbefaler Boyatzis (1998) å kvantifisere kvalitative data ut i fra hvor ofte en kode blir nevnt, eller i hvilken grad respondenter har tatt i bruk adjektiv for å beskrive fenomener av interesse. Eksempelvis kan utsagn kodes med en numerisk verdi som tilsier i hvilken grad en ledergruppe opplever en kategori som enten en styrke eller et utviklingsområde ved ledergruppen (1 = liten, 2 = middels, 3 = stor). Det er nærliggende å tenke at sammenligningene mellom de kvalitative kategoriene og *effect*-indikatorerne hadde økt i sensitivitet og nøyaktighet ved å ta i bruk en mer gradert skala.

På den annen side kan høy frekvens av et ord eller en setning ikke nødvendigvis indikerer økt betydning, men kan derimot være et resultat av en person eller en gruppes ønske om eller villighet til å nevne akkurat det ordet (Joffe & Yardley, 2004; Vaismoradi, Turunen & Bondas, 2013). Dessuten er de originale utsagnene som ligger til grunn for de kvalitative kategoriene begrenset i omfang ved at majoriteten av utsagnene gir lite utdypende beskrivelser av ledergruppens styrker og utviklingsområder. Eksempelvis svarte flere av medlemmene kun «Møteledelse» eller «Trygghet» på hva som fungerer bra og hva som kan forbedres i gruppen. Denne formen for svar gir lite informasjon om i hvilken grad ledelse og trygghet kjennetegner en ledergruppe.

Spørsmålene som ligger til grunn for de kvalitative kategoriene. Spørsmålet som utgjør ledergruppens utviklingsområder omhandler hva som kan forbedres i ledergruppen. Hva som kan forbedres er ikke nødvendigvis ensbetydende med hva som ikke fungerer i ledergruppen. For eksempel kan det tenkes at ledergrupper til en viss grad opplever å ha et klart bilde av hva ledergruppen er til for. Likevel kan ledergrupper ha et ønske om og behov for å forbedre opplevelsen av et Klart formål. Dette betyr at selv om ledergrupper skårer høyt på *effect*-indikatoren Klart formål, kan medlemmer benytte seg av det åpne spørsmålet i *effect* til å påpeke at ledergruppens formål kan forbedres. Muligheten for å beskrive forbedringsområder i motsetning til klare svakheter, kan være en forklaring på uoverensstemmelser i de kvalitative og kvantitative resultatene.

En annen svakhet ved bruk av åpne spørsmål er fravær av mulighet til å oppklare tvetydige svar hos respondenter (Lund & Haugen, 2006). For å unngå feiltolkninger i kategorier og temaer, var det viktig å eliminere tvetydige og ufullstendige svar fra analysen. I de tilfellene der svarene til respondentene var utydelige eller uforståelige ble de kodet inn i en residualkategori. Det er likevel en mulighet for at respondentene kan ha misforstått betydningen av de åpne spørsmålene, eller tolket dem på forskjellige måter innad i samme

ledergruppe. Dette kan ha påvirket resultatet av den tematiske analysen, og således selve begrepsvalideringen.

Utforming av spørreskjemaet. Mange og ulike typer spørsmål kan påvirke respondentens svar og dermed være en trussel for studiens indre validitet (Lasorsa, 2003; Schwarz, Hippler, Deutsch & Strack, 1985). Forhåndspåvirkning (priming effect) er når respondentenes svar er påvirket av tidligere spørsmål i spørreskjemaet (Moss & Lawrence, 1997; Tulving & Schacter, 1990). Dersom respondentenes svar på de åpne spørsmålene i *effect* er påvirket av spørreskjemaets 124 ledd om faktorer som er viktige for hvilke resultater ledergruppen oppnår, kan dette føre til at sammenhengen mellom de kvantitative og kvalitative resultatene fremstår sterkere enn den egentlig er.

Majoriteten av svarene på *effects* åpne spørsmål er korte fremfor lange. Det kan tenkes at respondentene ble slitne av å svare på samtlige av leddene (124 stk.) før de kom til delen med de åpne spørsmålene. Dette er kjent som «survey fatigue» (Lasorsa, 2003) og er et vanlig fenomen i omfattende måleinstrumenter og spørreundersøkelser. Dette kan ha påvirket hvor mye arbeid respondentene la i svarene på de kvalitative spørsmålene. Ulike respons bias slik som «survey fatigue» og «priming effect» kan problematisere tolkningen av sammenhengen mellom de kvalitative kategoriene og indikatorene i *effect*, og utgjør således en trussel mot studiens indre validitet.

Tematisk analyse. En fjerde begrensning ved studien er bruk av tematisk analyse. Etersom det eksterne sammenligningsgrunnlaget som anvendes i begrepsvalideringen er et resultat av funnene fra den tematiske analysen, er det viktig at disse funnene er troverdige. Ifølge Graneheim og Lundman (2004) handler troverdighet om i hvilken grad resultatene (temaer, kategorier og koder) faktisk representerer respondentenes originale utsagn. Resultatene av en tematisk analyse kan inneholde elementer av forskerens egne tolkninger av rådataen (Boyatzis, 1998; Braun & Clarke, 2006), og utgjør derfor en svakhet med tanke på kategoriseringsprosessens objektivitet.

Utvalget. Begrepsvalideringen er et resultat av data fra 16 toppledergrupper i statsforvaltningen i Norge (departementer, direktorater og tilsyn). For det første medførte en liten utvalgsstørrelse en begrenset mengde kvalitative data, som også førte til en svært liten utvalgsstørrelse i majoriteten av t-testene. Liten utvalgsstørrelse ($n \leq 20$) gir manglende statistisk styrke (Pallant, 2016). Dette kan ha hatt en betydning for resultatet av t-testene, ettersom det er en større sannsynlighet for å begå både type I og type II feil.

For det andre er utvalget kun basert på toppledergrupper fra statsforvaltningen og det er derfor problematisk å stadfeste hvorvidt funnene er gjeldene for ledergrupper og andre

typer team. Til slutt er det vanskelig å si om funnene kan generaliseres til andre land enn Norge.

Anbefalinger til fremtidig forskning

I lys av studiens begrensninger bør fremtidig forskning benytte seg av et større utvalg og andre metoder for innsamling av data. Et større utvalg vil øke mengden kvalitative data, som eventuelt kan føre til identifiseringen av flere kategorier og temaer som kan sammenlignes med indikatorene i *effect*. Videre vil et større utvalg medføre økt statistisk styrke for t-testene som undersøker sammenhengen mellom indikatorene og de kvalitative kategoriene. Bedre statistisk styrke gir reduksjon i både type I og type II feil (Howell, 2013; Stevens, 2002). Dette betyr at man i større grad kan være sikker på at de observerte forskjellene som fremgår av t-testene ikke skyldes tilfeldigheter, samtidig som man øker sannsynligheten for å identifisere faktiske forskjeller mellom utvalg.

Den andre anbefalingen er å ta i bruk intervjuer eller fokusgrupper fremfor åpne spørsmål for innhenting av kvalitative data (Sandelowski, 2000). Bruk av intervju og/eller fokusgruppe gir medlemmer av ledergrupper mulighet til å reflektere rundt og beskrive med egne ord i hvilken grad en faktor kjennetegner ledergruppen. Informasjonen kan således benyttes til å utvikle eksterne indikatorer som kan sammenlignes med skårene på indikatorene i *effect*.

For det tredje kan fremtidig forskning undersøke hvorvidt indikatorene i *effect* representerer refleksive eller formative begreper. Forskning (se MacKenzie, 2003; Mackenzie et al., 2011) viser at indikatorer ofte utvikles med forbehold om at indikatorene representerer refleksive begreper. Et refleksivt begrep er et begrep der man antar at et nivå av et latent begrep er årsaken til en skåre på en indikator (Bagozzi, 2007). Ifølge forskning (se Bagozzi, 2007; MacKenzie et al., 2011) er mange begreper i organisasjonspsykologi formative av natur. Det vil si at kombinasjonen av ulike indikatorer eller enkeltledd skaper et begrep (Bagozzi, 2007). Flere av indikatorene i *effect* består av enkeltledd som er innholdsmessige ulike hverandre. For eksempel består indikatoren Riktig kompetanse av et enkeltledd som illustrerer faglig kompetanse, mens et annet enkeltledd skal representere mellommenneskelig kompetanse. Likeledes gjelder dette indikatoren Dialog, som består av enkeltledd som omhandler det å lytte på hverandre, men også det å informere hverandre. Det kan tenkes at det er kombinasjonen av skårene på de ulike enkeltleddene i Dialog som utgjør det helhetlige begrepet Dialog, fremfor en underliggende g-faktor som påvirker skåren på samtlige av leddene i Dialog.

Manglende stadfesting av hvorvidt et begrep er formativt eller refleksivt kan medføre feilaktig konseptualisering og operasjonalisering av begreper, unødvendig eliminering av enkeltledd med lav Cronbachs alfa samt misvisende faktoranalyser (MacKenzie, 2003; MacKenzie et al., 2011).

Konklusjon

God begrepsvaliditet bør etterstrebnes i utvikling av spørreskjemaer, ettersom det er viktig å forsikre seg om at slutningene man foretar på bakgrunn av innsamlet data er gyldige (Shadish et al., 2002). I ledergruppeinventoriet *effect* er det av betydning for forskere og konsulenter hvorvidt en skåre på *effects* 27 indikatorer faktisk er en reell representasjon av de underliggende begrepene som indikatorene er designet for å måle, ettersom begrepene er viktige for kvaliteten på de resultatene ledergruppen oppnår (Bang & Midelfart, 2012).

Denne studien er den første til å begrepsvalidere et utvalg indikatorer i *effect* ved å undersøke sammenhengen mellom *effect*-indikatorene og eksterne indikatorer som er tiltenkt å representere de samme latente begrepene som indikatorene i *effect*. I denne oppgaven har jeg benyttet meg av 16 toppledergruppers egne beskrivelser av ledergruppens styrker og utviklingsområder som et eksternt sammenligningsgrunnlag. Ut i fra innholdet i de kvalitative kommentarene som omhandler styrker og utviklingsområder, fikk jeg muligheten til å begrepsvalidere 12 av indikatorene i *effect*.

Resultatet av begrepsvalideringen gir støtte til at slutninger man foretar på bakgrunn av skårene til 8 *effect*-indikatorer kan være gyldige. Det vil si at det kan tenkes at en skåre på indikatoren representerer det reelle nivået av det underliggende begrepet man ønsker å måle. De 8 indikatorene er: Riktige saker, Klart formål, Riktig størrelse, Gruppepsykologisk trygghet, Dialog, Fravær av politisk spill, Fravær av relasjonskonflikt og Møteforberedelser. Videre tilsier funnene at man bør være forsiktig med å trekke for vide slutninger om hva indikatorene Effektiv teamledelse, Riktig kompetanse, Fokuseret kommunikasjon og Lagånd faktisk måler, inntil flere studier kan si noe om begrepsvaliditeten til disse fire indikatorene.

Resultatene fra én enkelt valideringsstudie gir ikke tilstrekkelig nok informasjon til å konkludere med sikkerhet hvorvidt de 8 indikatorene faktisk har god begrepsvaliditet (Peter, 1981). Denne studien er derfor én av forhåpentligvis flere valideringsstudier av ledergruppeinventoriet *effect*.

Litteraturliste

- Adcock, R., & Collier, D. (2001) Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529-546. doi: 10.1017/S0003055401003100
- Bagozzi, R. P. (2007). On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods*, 12(2), 229-237. doi:10.1037/1082-989X.12.2.229
- Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 27(1), 3-23. doi: 10.2307/3315487
- Bang, H. (2008). Effektivitet i ledergrupper – hva er det, og hvilke faktorer påvirker det? *Tidsskrift for Norsk Psykologforening*, 45(3), 272-286.
- Bang, H. (2010). *Effektivitet i ledergrupper: En studie av sammenhengen mellom gruppeprosesser og teameffektivitet i ledermøter* (Doktoravhandling). Psykologisk institutt, Universitetet i Oslo.
- Bang, H. (2017). Technical manual on *effect* – An inventory of measuring management team effectiveness. Upublisert manuskript, Psykologisk institutt, Universitetet i Oslo.
- Bang, H., & Midelfart, T.N. (2012). *Effektive ledergrupper*. Oslo: Gyldendal Akademiske Forlag.
- Bang, H., & Midelfart, T. N. (2017). What characterizes effective management teams? A research-based approach. *Consulting Psychology Journal: Practice and Research*, 69(4), 334-359. doi: 10.1037/cpb0000098
- Barrick, M., Bradley, B., Kristof-Brown, A., & Colbert, A. (2007). The moderating role of top management team interdependence: Implications for real teams and working

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

groups. *The Academy of Management Journal*, 50(3), 544-557. doi:
10.2307/20159872

Berger, A., & Harr, M. (2013). *Norske toppledergruppers formål: Utvikling av en formålstypologi* (Masteroppgave). Oslo. Hentet fra:
<https://www.duo.uio.no/handle/10852/36758>

Biemann, T., Cole, M. S., & Voelpel, S. (2012). Within-group agreement: On the use (and misuse) of rWG and rWG (J) in leadership research and some best practice guidelines. *The Leadership Quarterly*, 23(1), 66-80. doi:10.1016/j.leaqua.2011.11.006

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analyses and code development*. Thousand Oaks, CA: Sage Publications.

Braun, V., & Clarke, V. (2006). Using thematic analyses in psychology. *Qualitative research in psychology*, 3(2), 77-101. doi: 10.1191/1478088706qp063oa

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105. doi:
10.1037/h004601

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand-McNally.

Cronbach, L. J. (1971). Test validation. I R. L. Thorndike (red.), *Educational measurement* (2.utg, s. 443-507). Washington, D.C: American Council on Education.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4). doi: 10.1037/h0040957

DeVellis, R. F. (2016). *Scale development: Theory and applications* (3.utg.). London: Sage publications.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Driscoll, D. L., Appiah-Yeboah, A., Salib, P., & Rupert, D. J. (2007). Merging qualitative and quantitative data in mixed method research: How to and why not. *Ecological and Environmental Anthropology*, 3(1), 19-28.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4.utg.). London: Sage Publications.

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse education today*, 24(2), 105-112. doi:10.1016/j.nedt.2003.10.001

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. doi: 10.4135/9781483384436.n4

Hackman, J. R (2002). *Leading teams. Setting the stage for great performances*. Boston: Harvard Business School Press.

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *The Academy of Management Review*, 9(2), 193- 206. doi: 10.2307/258434

Hansen, L.B (2017). *The Dimensionality of Management Team Effectiveness: A psychometric analysis of the team inventory "effect"* (Masteroppgave). Oslo. Hentet fra: <https://www.duo.uio.no/handle/10852/57724>

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967-988. doi: 10.4135/9781483384436.n4

Howell, D. C. (2013). *Statistical methods for psychology* (8. utg.). Wadsworth: Cengage Learning.

Joffe, H., & Yardley, L. (2004). Content and thematic analysis. In D. F. Marks, & L. Yardley (Red.), *Research methods for clinical and health psychology* (s.56-68). London: Sage

Publications.

Kleven, T. A. (2002). Begrepsoperasjonalisering. I T. Lund (red.), *Innføring i forskningsmetodologi* (s.141-183). Oslo: Unipub forlag.

Kraut, A. I. (1996). *Organizational surveys: Tools for assessment and change*. San Francisco: Jossey-Bass.

Landis, J.R., & Koch, G.G (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. doi:10.2307/2529310

Lasorsa, D. L. (2003). Question-order effects in surveys: The case of political interest, news attention, and knowledge. *Journalism and Mass Communications Quarterly*, 80(3), 499-512. doi: 10.1177/10.7769900308000302

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11(4), 815-852. doi: 10.1177/1094428106296642

Lund, T. (2002) Metodologiske prinsipper og referanserammer. I T. Lund (red.), *Innføring i forskningsmetodologi* (s.79-121). Oslo: Unipub forlag.

Lund, T., & Haugen, R. (2006). *Forskningsprosessen*. Oslo: Unipub Forlag.

Machery, E. (2007). 100 years of psychology concepts: The theoretical notion of concepts and its operationalization. *Science Direct*, 38(1), 63-84. doi: 10.1177/10.7769900308000302

MacKenzie, S. B. (2003). The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science*, 31(3), 323-326. doi: 10.1177/0092070303254130

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293-334. doi: 10.1177/10.7769900308000302

- McGrath, J.E. (1964). *Social psychology*. New York: Holt, Rinehart and Winston.
- McIntyre, M.G. (1998). *The management team handbook: Five key strategies for maximizing group performance*. San Francisco: Jossey-Bass.
- Messick, S. (1975). The standard problem: Meaning and values in measurement and evaluation. *American Psychologist*, 30(10), 955-966. doi: 10.1037/0003066x.30.10.955
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. doi: 10.1037/0003-066X.50.9.741
- Moss, S.E., & Lawrence, K.G. (1997). The effects of priming on the self-reporting of perceived stressors and strains. *Journal of Organizational Behavior*, 18(4), 393-403.
- Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (6.utg.). Berkshire, UK: McGraw-Hill Education.
- Peter, J. P. (1981). Construct validity: A review of basic issues and marketing practices. *Journal of Marketing Research*, 18(2), 133-145. doi: 10.2307/3150948
- Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, 15(1), 85-109. doi: 10.1177/1525822X02239569
- Saldaña, J. (2009). *The Coding Manual for Qualitative Researchers*. London: Sage Publications.
- Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, 23(3), 246-255.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

- Sandelowski, M., & Barroso, J. (2003). Classifying the findings in qualitative studies. *Qualitative Health Research, 3*(7), 905-923. doi: 10.1177/1049732303253488
- Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response scales effects of category of range on reported behavior and comparative judgments. *Public Opinion Quarterly, 49*(3), 368-395.
- Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin Company.
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4.utg). Mahwah: Lawrence Erlbaum.
- Stone, E. F. (1978). *Research methods in organizational behavior*. Santa Monica: Goodyear Publishing Company, Inc.
- Taylor, C. S. (2013). *Validity and validation: Understanding Statistics*. New York: Oxford University Press.
- Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306.
- Vaismoradi, B., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences, 15*(3), 398-405. doi: 10.1111/nhs.12048
- Wageman, R., Nunes, D. A., Burruss, J. A., & Hackman, J. R. (2008). *Senior leadership teams: What it takes to make them great*. Boston: Harvard Business School.
- Wageman, R., & Hackman, J. R. (2010). What makes teams of leaders leadable? I N. Nohria., & R. E. Khurana (Red.), *Advancing leadership* (s. 475-507). Boston: Harvard Business School Press.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Welch, B. L. (1958). "Student" and small sample theory. *Journal of the American Statistical Association*, 53(284), 777-788. doi: 10.2307/2281951

Winter, J.C.F. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation*, 18(10), 1-12.

Vedlegg
VEDLEGG A

Identifisering av meningsenheter

Identifisering av meningsenheter er basert på Boyatzis (1998) definisjon av en meningsenhet: «Det mest grunnleggende segmentet, eller elementet av rådata eller informasjon som kan vurderes på en meningsfull måte» (Boyatzis, 1998, s.64, *oversettelse av Berger & Harr*, 2013). Det kvalitative datamaterialet er et resultat av to åpne spørsmål hentet fra ledergruppeinventoriet *effect*. Hvor mye tekst hver respondent har valgt å svare, varierer fra person til person, alt fra et par korte stikkord, til 4-5 setninger. Følgende illustreres to eksempler på utsagn basert på spørsmålet: Hva fungerer spesielt bra i ledergruppen? og to utsagn basert på spørsmålet: Hva kan forbedres i måten ledergruppen fungerer på?

Hva fungerer spesielt bra i ledergruppen?

«Åpen dialog, god tone og at vi liker å arbeide sammen» og «Folk med mye kunnskap. Alle bringer inn ulike perspektiver for å belyse problemstillingene. Evne til å tenke helhetlig til det beste for organisasjonen».

Hva kan forbedres i måten i ledergruppen fungerer på?

«Bedre forberedelse i forhold til hva vi ønsker å diskutere og hvordan hver enkelt kan bidra» og «I større grad behandle prinsipielle saker. Klargjøring av ledergruppens rolle. Mindre saksdokumenter i noen saker. Litt mer interesse for hverandre som mennesker for å skape tillit (vi er nokså upersonlige)».

Utsagnene er representative for datamateriale ettersom materialet i hovedsak består av en kombinasjon av korte setninger og oppramsinger. For å lette lesningen samt opprettholde ledergruppens anonymitet, er setninger der ledergruppens tilhørighet er nevnt, byttet ut med *virksomheten* eller *organisasjonen*. På lik linje er eksempelvis *departementsråden* byttet ut med ordet *leder/overordnede* og *han/hun* byttet ut med *leder/overordnede* dersom det er klart at det snakkes om ledergruppens leder. Dette er gjort for å ikke miste meningsenhetens betydning ved identifisering og oppdeling av setninger.

Prosedyren for identifisering av meningsenheter ble utviklet i samarbeid med veileder. I de tilfellene hvor utsagn som beskrev hva som fungerer bra i ledergruppen inneholdt flere adjektiv, var det naturlig å identifisere meningsenheten ved å dele opp setningen og gjenta subjektet/substantivet. Eksempelvis så ble utsagnet «Åpen og god tone. Ikke spisse albuer»

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

delt inn i meningsenhetene «Åpen tone», «God tone» og «Ikke spisse albuer». Dette er fordi det ble tatt i bruk en semantisk prosedyre (Saldaña, 2009) som betyr at man skiller meningsenheter med ulike betydninger for å få et mest mulig rikt datasett. Det kan tenke seg at de tre meningsenhetene eksemplifisert ovenfor vil falle under samme kategori senere i analysen, men siden meningsenhetene er identifisert ved en induktiv tilnærming bør man ikke la seg begrense av mulige like meningsenheter. Følgende prosess ble gjennomført:

1. Veileder og undertegnede diskuterte og ble enige om felles retningslinjer for identifisering av meningsenheter.
2. For å sørge for konsistens i prosessen identifiserte vi meningsenheter i én ledergrupper hver for oss. Vi gikk så over hverandres meningsenheter for å undersøke hvorvidt vi var enige i hverandres beslutninger. Der det var inkonsistens diskuterte vi oss fram til retningslinjer for videre identifisering. Videre identifiserte vi to ulike ledergrupper hver for oss mens vi diskuterte eventuelle problemstillinger underveis. Da vi var sikre på at samtlige utfordringer var diskutert og retningslinjer for videre koding var etablert, delte vi inn de resterende ledergruppene mellom oss for videre identifisering.
3. Der vi var usikre på identifisering av meningsenheter i et enkelt utsagn, markerte vi dette med en stjerne. Etter at samtlige ledergrupper var gjennomgått, gikk jeg gjennom hver enkelt meningsenhet markert med stjerne, for å sørge for korrekt behandling av hvert utsagn.

VEDLEGG B

Kodeskjema

En kode består som oftest av et ord eller en kort setning som representerer datamaterialets innhold og essens (Saldaña, 2009). Kodingsprosessen er iterativ, det vil si at man koder i flere runder for å sørge for at samtlige meningsenheter blir merket med passende koder, uavhengig hvor man er i kodingsprosessen. Følgende illustreres kodereglene som jeg har benyttet meg av i denne oppgaven:

1. En meningsenhet kan kun merkes med én kode. Hvis det er et spørsmål hvorvidt en meningsenhet tilfaller flere koder, bør meningsenheten revurderes og eventuelt deles opp i flere meningsenheter.
2. Meningsenheter som tilhører samme respondent kan kodes under samme kode, så lenge respondentens ID er inntak. Dette er viktig for å kunne si noe om frekvensen av hvert tema og kategori som oppstår i en ledergruppe.
3. Hver meningsenhet skal så langt det lar seg gjøres bli merket med en kode. I de tilfellene der meningsenhetene består av for lite informasjon eller er uforståelige, skal de kodes som «residual».

Kodereglene nevnt ovenfor ble tatt i bruk og resulterte i følgende kodingsprosess:

1. Hver meningsenhet ble gjennomgått og merket med en kode i Excel.
2. Kodelisten ble til etter hvert i prosessen og bestod av navnet på koden, koden representert ved et tall og en kort forklaring av hver kode.
3. Hver gang én meningsenhet ikke passet inn i en eksisterende kode, ble det laget en ny kode.
4. Noen meningsenheter hadde for lite informasjon, eller var uforståelige. To eksempler på slike utsagn er: «*Ingen spesielle innspill*» og «*Det er for tidlig å kunne gi en vurdering av dette*». Slike utsagn ble tildelt koden «residual».
5. Gjennomgang av koder med veileder.
6. Gjentakelse av kodingsprosessen (iterativ prosess).

VEDLEGG C

Oversikt over temaer, kategorier, koder og antall meningsenheter per kode

Temaer og kategorier	Koder	Antall ME
	<u>Styrker</u>	
Relasjoner og fellesskap		
<i>Fellesskap og samhold</i>	Godt gruppefellesskap	4
	Felles mål	3
	Lojalitet	3
	Helhetsperspektiv	7
<i>Samarbeid</i>	God gruppedynamikk	1
	Godt samarbeid	10
	Alle medlemmene bidrar	7
	Gir hverandre innspill utenom møtene	2
	Hjelper hverandre	4
	Inkluderende atmosfære	1
<i>Relasjoner</i>	Gode relasjoner	4
	Trives sammen	6
	Ingen eller få konflikter	2
	Respekt for hverandre	6
	Vil hverandre vel	9
<i>Stemning</i>	God stemning	9
	God tone	10
	Godt klima	2
	Godt humør	14
	Diktlesning	2
	God energi	1
Saker		
<i>Diskutere riktige saker</i>	Diskuterer strategi	9
	Viktige saker	10
	Diskuterer utvikling	5
Diskusjonskvalitet		
<i>Tydelige diskusjoner</i>	Direkte tone	1
	Lite avsporinger	2
	Konkrete diskusjoner	1
	Ryddige diskusjoner	1
	Tydelig informasjonsutveksling	1
<i>Åpen meningsutveksling</i>	Meningsbrytning	7
	Djevelens advokat	1
	Nysgjerrig på hverandres utfordringer	1

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

	Åpen meningsutveksling	3
	God dialog	8
	Åpne diskusjoner	9
	Gode diskusjoner	27
	God informasjonsutveksling	5
Ledermøter og sakspapirer		
<i>Gjennomføring av ledermøtene</i>	Strukturerte møter	15
	Oppsummering i ledermøtene	5
	Effektive ledermøter	7
<i>Møteforberedelser og sakspapirer</i>	Møteforberedelser	12
	Sakspapirer	4
<i>Møteledelse</i>	God møteledelse	13
	Dyktig leder	4
Trygghet og åpenhet		
<i>Politisk spill</i>	Ingen eller få skjulte agendaer	5
	Ingen maktkamper	2
	Ingen eller lite spill	2
	Sakene er ikke avgjort på forhånd	2
<i>Trygg og åpen atmosfære</i>	Åpenhet	28
	Høyt under taket	6
	Trygghet	18
	Tillitt	14
	Lite prestisje	1
	Rom for innspill	4
	Konstruktive	1
Kompetanse		
<i>Kompetanse</i>	Kompetanse	25
	Faglige diskusjoner	4
	Løsningsorientert	2
<i>Kompetansedeling</i>	Kompetansedeling	5
	Erfaringsdeling	2
	Læring	5
Ledergruppens sammensetning		
<i>Størrelse</i>	Riktig størrelse	3
<i>Mangfold og variasjon i ansvar</i>	Viktige medlemmer	1
	Mangfold	2
Beslutninger	Gode beslutninger	14

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Temaer og kategorier	Koder	Antall ME
	<u>Utviklingsområder</u>	
Relasjoner og fellesskap		
<i>Fellesskap og samhold</i>	Teamfølelse	13
	Helhetsperspektiv	12
<i>Samarbeid</i>	Samarbeid på tvers av avd.	14
	Samhandling i vanskelige saker	11
	Konflikthåndtering	1
<i>Relasjoner</i>	Bli kjent med hverandre	3
	Anerkjenne hverandre	1
	Støtte hverandre	1
	Aksept for uenigheter	1
	Utfordringer med enkeltpersoner	7
<i>Stemning</i>	Energi i gruppen	2
	Positivt fokus	4
Saker		
<i>Diskutere riktige saker</i>	Diskutere andre saker	6
	Diskutere strategiske saker	51
	Diskutere prioriteringer	5
	Diskutere viktige saker	9
	Fokus på organisasjonskultur	1
	Flere fellesprosjekter	1
	Flere faglige saker	4
	Diskutere riktige saker	28
	Refleksivitet rundt ledergruppens arbeid	2
<i>Riktig tidspunkt</i>	Riktig tidspunkt for å diskutere riktige saker	17
Diskusjonskvalitet		
<i>Tydeligere diskusjoner</i>	Mindre høytflyvende innlegg	1
	Avsporinger	1
	Tydelige diskusjoner	1
	Bedre innspill	1
<i>Åpen meningsutveksling</i>	Nysgjerrige på hverandres perspektiver	5
	Utforske meninger	3
	Åpenhet i diskusjoner	3
	Friskere diskusjoner	7
	Kritiske innspill	2
	Andre diskusjonsformer	1
	Si i fra på en konstruktiv måte	3
	Lytte på hverandre i diskusjoner	3

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

	Utfordre hverandre	11
Ledermøter og sakspapirer		
<i>Gjennomføring av ledermøtene</i>	Bedre møterutiner	3
	Lengden på ledermøter	4
	Riktig tidsbruk på saker	10
	Effektive møter	5
	Mindre administrasjon	2
<i>Møteforberedelser og sakspapirer</i>	Møteforberedelser	20
	Færre saksdokumenter	1
	Bedre sakspapirer	4
Trygghet og åpenhet		
<i>Politisk spill</i>	Koalisjoner	1
	Sakene er avgjort på forhånd	8
	Gehør basert på uformell posisjon	1
<i>Trygg og åpen atmosfære</i>	Trygghet	3
	Tillitt	1
	Åpenhet	2
Ledergruppens sammensetning		
<i>Størrelse</i>	For stor gruppe	7
<i>Mangfold og variasjon i ansvar</i>	Kompetanse utenfra	2
	Perspektiver utenfra	1
	Ulike ansvarsområder	3
	Innsikt i hverandres fagfelt	2
	Forståelse for hverandres saker	1
Beslutninger	Bedre konklusjoner	2
	Tydligere beslutninger	11
	Oppfølging av beslutninger	1
Ledergruppens formål		
<i>Formålsavklaring</i>	Formålsavklaring	12
<i>Saksavklaring</i>	Saksavklaring	13

VEDLEGG D

Oversikt over kategoriene og *effect*-indikatorene

Kategorier (tematisk analyse)	<i>effect</i> -faktor (definisjoner hentet fra <i>effect</i> 2.0)
<p>Ledergruppens formål</p> <p>I hvilken grad opplever medlemmene at det er avklart forventninger til ledergruppen, og at det er en felles enighet angående ledergruppens formål og hvilke saker ledergruppen skal behandle.</p>	<p>Klart formål</p> <p>Medlemmene har en klar og felles oppfatning av hvorfor ledergruppen er til, hva den skal levere og hvilke saker og temaer som hører hjemme der.</p>
<p>Saker</p> <p>I hvilken grad ledergruppen opplever å diskutere viktige saker som er relevante ledergruppen, og hvorvidt sakene kommer inn til diskusjon i ledergruppen til riktig tidspunkt.</p>	<p>Riktige saker</p> <p>Ledergruppen bruker tiden sin på viktige saker som er i tråd med ledergruppens formål, og som trenger hele ledergruppens oppmerksomhet. Møtene har en god balanse mellom orienteringssaker, diskusjonssaker og beslutningssaker.</p>
<p>Størrelse</p> <p>Hvorvidt gruppen oppleves som passe stor.</p>	<p>Riktig størrelse</p> <p>Med tanke på ledergruppens formål og oppgaver er ledergruppen passe stor.</p>
<p>Kompetanse</p> <p>I hvilken grad medlemmene besitter nødvendig faglig kompetanse og kunnskap for å bidra til faglige diskusjoner og problemløsning.</p>	<p>Riktig kompetanse</p> <p>Ledergruppens medlemmer har den nødvendig faglige, intellektuelle og mellommenneskelige kompetansen for å skape gode resultater.</p>
<p>Møteforberedelser og sakspapirer</p> <p>I hvilken grad medlemmene er forberedt til ledermøtene, og hvorvidt sakene og de tilhørende sakspapirene er av god kvalitet og sendt ut til tide.</p>	<p>Møteforberedelser</p> <p>Medlemmene stiller godt nok forberedt til ledermøtene, sakspapirene er av høy kvalitet, og de sendes ut i god nok tid før ledermøtet.</p>
<p>Tydelige diskusjoner</p> <p>I hvilken grad ledermøtene preges av tydelig kommunikasjon med konkrete og ryddige innspill med lite avsporinger.</p>	<p>Fokusert kommunikasjon</p> <p>Medlemmene holder seg til saken, unngår omstendelige innlegg og utenomstakk, og det blir oppsummert og konkludert på slutten av hver diskusjon i ledermøtene.</p>

Relasjoner

Hvorvidt gruppen preges av konflikter og utfordringer knyttet til enkeltpersoner. I hvilken grad medlemmene trives med hverandre og respekterer hverandre.

Fravær av relasjonskonflikt

Medlemmene kommer godt overens med hverandre, og det er få eller ingen gnisninger, negative spenninger eller personlige uoverensstemmelser i gruppen.

Politisk spill

I hvilken grad ledergruppen preges av skjulte agendaer, maktkamper, koalisjoner og taktikkeri. Sakene er ikke avgjort på forhånd.

Fravær av politisk spill

Ledergruppen preges ikke av maktkamper og skjulte agendaer, og det er ingen som bevisst holder tilbake informasjon som kan være viktig for ledergruppen.

Åpen meningsutveksling

I hvilken grad ulike synspunkter kommer frem og hvorvidt medlemmene lytter på hverandre og er interessert i andres meninger og perspektiver.

Dialog

Kommunikasjonen i ledergruppen preges av at medlemmene åpent uttrykker sine syn og meninger, samtidig som de nysgjerrig utforsker hverandres synspunkter. De diskuterer uten å nedvurdere hverandre, og bygger ofte på hverandres ideer.

Møteledelse

Hvorvidt lederen av ledergruppen utøver tydelig møteledelse; oppsummerer, oppfordrer til deltakelse og er tydelig i sine beslutninger.

Effektiv teamledelse

Gruppens leder sørger for at alle kritiske funksjoner for effektiv fungering blir ivarettatt. Lederen skaper et trygt klima og styrer ledermøtene på en effektiv måte.

Trygg og åpen atmosfære

Hvorvidt medlemmer kan dele utfordringer av faglig og personlig art og ytre meninger uten å være redd for negative konsekvenser.

Gruppepsykologisk trygghet

Det oppleves som trygt å ta opp problemer, gi uttrykk for usikkerhet og å be om hjelp. Det er ingen frykt for at det vil straffe seg om man åpent sier hva man mener.

Fellesskap og samhold

I hvilken grad ledergruppen fokuserer på fellesskapet fremfor individet, ser hele organisasjonsbildet fremfor egen vinning og lojalitet til gruppen og dens medlemmer står sterkt.

Lagånd

Gruppen er sammensveiset og uttrykker seg utad med én stemme. Medlemmene liker hverandre, er stolte av gruppen, og inntar et helhetsperspektiv i saker.

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

VEDLEGG E

Oversikt over hvor mange medlemmer per ledergruppe som kvalitativt beskrev en kategori enten som en styrke (S) eller som et utviklingsområde (U).

Kategorier	Ledergrupper (1-8)															
	1		2		3		4		5		6		7		8	
	S	U	S	U	S	U	S	U	S	U	S	U	S	U	S	U
Ledergruppens formål	-	2	-	1	-	1	-	1	-	1	-	-	-	2	-	1
Saker	-	3	-	4	1	2	-	4	-	5	-	2	1	-	-	1
Størrelse	-	-	-	-	-	-	-	4	-	-	-	-	-	-	-	-
Kompetanse	-	-	2	-	2	-	-	-	4	-	2	-	2	-	1	-
Møteforberedelser og sakspapirer	-	4	-	-	-	4	1	-	2	1	-	-	2	-	-	-
Tydelige diskusjoner	1	3	-	-	2	-	1	-	-	-	1	-	1	-	-	-
Relasjoner	2	-	3	1	1	4	2	-	2	1	-	-	-	-	2	-
Politisk spill	-	-	1	-	-	1	1	-	-	1	-	-	-	-	-	-
Åpen meningsutveksling	1	2	-	2	4	1	7	1	3	2	1	-	5	-	2	-
Møteledelse	2	-	-	-	3	-	4	-	3	-	-	-	-	-	-	-
Trygg og åpen atmosfære	-	1	4	-	2	-	5	1	4	1	2	-	-	-	3	-
Samhold og fellesskap	1	-	1	-	1	-	-	1	1	-	-	1	-	-	1	-

BEGREPSVALIDERING AV ET UTVALG INDIKATORER I EFFECT

Ledergrupper (9-16)																
Kategorier	9		10		11		12		13		14		15		16	
	S	U	S	U	S	U	S	U	S	U	S	U	S	U	S	U
Ledergruppens formål	-	2	-	2	-	-	-	1	-	2	-	-	-	2	-	-
Saker	-	7	1	6	1	1	1	6	-	3	-	5	-	3	-	3
Størrelse	-	-	-	-	2	-	-	-	-	-	-	1	-	-	-	-
Kompetanse	2	-	-	-	3	-	-	-	-	-	1	-	1	-	1	-
Møteforberedelser og sakspapirer	-	3	-	-	1	2	-	-	2	-	1	-	-	1	1	-
Tydelige diskusjoner	-	-	-	1	-	-	-	1	-	-	-	-	-	-	-	-
Relasjoner	1	-	-	2	-	-	1	-	3	1	3	2	-	-	1	-
Politisk spill	-	-	-	-	1	2	-	-	3	-	-	-	-	-	1	-
Åpen meningsutveksling	3	-	1	3	5	1	1	-	2	3	5	3	-	1	1	2
Møteledelse	-	-	-	-	-	-	1	-	1	-	1	-	-	-	-	-
Trygg og åpen atmosfære	4	-	-	1	2	-	1	-	6	-	4	1	3	-	-	-
Samhold og fellesskap	2	-	-	1	-	-	2	-	-	4	-	2	-	-	-	1