

Pernille Boye Koch og Carina Saxlund Bischoff

Laster og dyder i embedsværk og forskning – en metodekritisk analyse af Bo Smith-udvalgets undersøgelser af embedsværkets overholdelse af regler og normer

Bo Smith-udvalgets konklusioner om tilstanden i det danske embedsværk har sat markante aftryk i den offentlige debat. Konklusionerne indeholder en klar afvisning af, at der er systemiske fejl i samspillet mellem politikere og embedsmænd, og det fremhæves, at de klassiske embedsmandsnormer efterleves i betydelig udstrækning. Vi gennemgår her de undersøgelser, som konklusionerne hviler på, og afdækker væsentlige brud på forskningskriterierne og metodiske fejlgreb, som svækker validiteten af udvalgets konklusioner. Faktisk kunne dele af datagrundlaget lige så godt tolkes omvendt, når det handler om lovlighedsnormens overholdelse. Det fremgår nemlig, at et bemærkelsesværdigt højt antal embedsmænd er villige til at medvirke til ulovlige handlinger. Vi understreger derfor, at der er behov for flere og bedre undersøgelser for at få relevant viden om normoverholdelse i centraladministrationen.

Med Bo Smith-udvalgets rapport, *Embedsmanden i det moderne folkestyre*, fik vi det seneste bidrag til den eksisterende viden om forholdet mellem politikere og embedsmænd i Danmark. Bo Smith-udvalget (BS)¹ blev nedsat i kølvandet på en række kritisable sager (fx Jægerbogsagen, Christiania-sagen, GGGI-sagen, Zornig-sagen, skattesagen og statsløsesagen), som skabte debat i offentligheden om forholdene i centraladministrationen og etikken i embedsværket. Et af hovedformålene med udvalgets arbejde var at foretage en aktuel analyse af samarbejdet mellem den politiske ledelse og embedsværket, herunder en vurdering af ”om embedsmanden er fulgt med tiden og løser sine opgaver med respekt for gældende regler og normer” (BS, 2015: 14).

Udvalget gennemførte på den baggrund en række empiriske undersøgelser, sammenfattede dem i en rapport og præsenterede under stor offentlig bevågenhed den overordnede konklusion: Der er ikke nogen systemiske fejl i samspillet mellem politikere og embedsmænd, og embedsmandsdyderne har det grundlæggende set godt. Konklusionen er umiddelbart overraskende i lyset af de

nævnte sager og den udbredte offentlige debat om embedsmandssystemet (se fx Koch og Knudsen, 2014; Tynell, 2015; Loft og Rosted, 2016). På den baggrund besluttede vi at se nærmere på rapporten og særligt de data og metoder, som udvalget bygger deres konklusion på.

Det er vigtigt at nævne, at Bo Smith-udvalgets rapport skriver sig ind i en længere række af betænkninger og redegørelser, som de seneste 25-30 år har beskæftiget sig med forholdet mellem ministre og embedsmænd.² Rapporten refererer og sammenfatter disse undersøgelser og inddrager ligeledes relevante resultater fra forvaltningsforskningen på de spørgsmål, som den behandler. Nogle af de spørgsmål, som udvalget tager op, har imidlertid ikke tidligere været genstand for systematiske undersøgelser i Danmark.³ Udvalget præsenterer således resultater, som ikke kan holdes op imod eksisterende viden. På disse områder er Bo Smith-undersøgelserne således særligt værdifulde, men det betyder også, at det er ekstra vigtigt at undersøge grundlaget for de konklusioner, som rapporten præsenterer.

Vi fokuserer i artiklen på to af udvalgets tre selvstændige undersøgelser, nemlig undersøgelsen om udviklingen i antal og fælles karakteristika for de "kritiske sager" og undersøgelsen om embedsmandsnormernes forankring.⁴ På baggrund af den første undersøgelse drager Bo Smith-udvalget den overordnede konklusion, at der ikke er kommet flere problematiske sager siden 1980 (BS, 2015: 20, 43-45), at der ikke er nogen fællesnævner ved sagerne, som peger på tilstedeværelsen af et underliggende problem (p. 49), og at efterfølgende undersøgelser ikke altid har identificeret alvorlige fejl og forsømmelser (p. 48). På baggrund af undersøgelsen om embedsmandsnormerne skriver udvalget, at der ikke er tegn på, at embedsmændene i højere grad end tidligere misligholder deres forpligtelser (p. 49). De klassiske embedsmandsnormer ser ud til at blive efterlevet i betydelig udstrækning, og dette gælder ifølge udvalget i særlig grad lovlighedsforpligtelsen (pp. 197, 239). Udvalget konkluderer dog også, at der er behov for at styrke embedsmændenes kendskab til normerne.

I artiklen vurderer vi de empiriske undersøgelser, som udvalget baserer sine konklusioner på, ud fra fire centrale og alment accepterede forskningskriterier. Vi er opmærksomme på, at udvalgets sammensætning er speciel, idet udvalget både består af medlemmer med en relevant faglig og forskningsmæssig baggrund samt medlemmer med en mere praktisk og erfaringsbaseret viden om politikere og embedsmænd. Også selve rapporten er som genre lidt speciel, idet den fremstår som en blanding af vidensformidling, afrapportering af forskningsresultater og anbefalinger af mere retspolitisk karakter. Når vi alligevel mener, at undersøgelserne bør vurderes ud fra forskningskriterierne, er det, fordi de er designet og gennemført som forskningsstudier, ligesom de indhentede

data indgår i forskningsartikler (se fx Smith og Christensen, 2016 samt Christensen, Wang og Opstrup, 2016), men også fordi Bo Smith-udvalgets konklusioner er med til at etablere vidensgrundlaget for helt principielle diskussioner om embedsværkets rolle i demokratiet. Det er også vigtigt at understrege, at vi ikke foretager en tilsvarende granskning af de dele af rapporten, der har karakter af fx formidling af eksisterende viden eller politiske anbefalinger.

Vi finder en række grundlæggende problemer i undersøgelserne i forhold til følgende forskningskriterier: transparens og replicerbarhed, undersøgelsesvaliditet (overensstemmelse mellem problemstilling og undersøgelsesdesign) og målevaliditet. Endelig sætter vi spørgsmålstegn ved den eksterne validitet. Vi uddyber, hvad der ligger i disse kriterier, og hvilken betydning de har i forbindelse med gennemgangen af undersøgelserne nedenfor. Gennemgangen af materialet viser desuden, at data fra spørgeskemaundersøgelsen på udvalgte punkter kan foranledige helt andre konklusioner end dem, Bo Smith-udvalget drager.

Strukturen i artiklen er som følger. Efter en kort præsentation af det tilgængelige datamateriale vil vi diskutere grundlaget for den talmæssige opgørelse af “antal kritiske sager”. Der rejses kritik af den anvendte optællingsmetode, og ved om det i praksis er muligt at lave valide opgørelser af denne type. I forhold til fællestræk ved sagerne savnes en systematisk transparent analyse af sagerne, idet det ikke angives, hvilke karakteristika ved sagerne der er undersøgt, og hvordan de er klassificeret. Centrale karakteristika synes desuden at være helt udeladt af analysen. Endeligt sættes der spørgsmålstegn ved opgørelsen af sagernes konsekvenser.

I den anden del af artiklen beskæftiger vi os med den spørgeskemaundersøgelse, som danner grundlag for konklusionerne om embedsmænds kendskab til og anvendelse af de klassiske embedsmandsnormer. Først ses på det overordnede undersøgelsesdesign og stikprøven, hvorefter der fremlægges en kritik af spørgeskemaet, hvor der peges på problematiske fejlgreb i formuleringen af både spørgsmål og “vignetter”, som svækker anvendeligheden af store dele af datamaterialet.

Til sidst sammenfattes resultaterne af vores undersøgelse, det diskuteres hvilke implikationer den har for vores viden om feltet, og hvilke behov der er for yderligere undersøgelser.

Tilgængeligt datamateriale

Selve rapporten og dens bilag udgør en vigtig del af datagrundlaget. I forhold til undersøgelsen vedrørende de “kritiske sager” har vi kun haft det datamateriale, som er offentliggjort i selve rapporten, hvilket er en liste over de i alt

49 sager, som i perioden 1980-2015 har været gjort til genstand for særlige undersøgelser (BS, 2015: 289-290, bilag 6). I forhold til undersøgelsen af embedsmandsnormernes forankring, har vi været så heldige, at udvalget har stillet datasættet fra spørgeskemaundersøgelsen til vores rådighed. Det er vi meget taknemlige for, da det har muliggjort videre efterprøvning af resultaterne.

Bo Smith-udvalget har også afholdt interview og høringer med mere end 130 nuværende og forhenværende politikere, embedsmænd, forskere og medierepræsentanter. Det fremgår af rapporten (BS, 2015: 20), at interview og høringer er afrapporteret skriftligt til udvalget, og at ”vigtige pointer fra interviewene er indsat som citater i teksten” (BS, 2015: 168). Det er i bilag til rapporten angivet navnene på deltagere i høringer og interview, og der er også optrykt et skriftligt diskussionsoplæg. Desværre sammenfattes og præsenteres *resultaterne* af disse høringer og interview ikke i nogen systematisk form.⁵ Vi har endvidere ikke haft mulighed for at se materialet, som offentligheden (og andre forskere) er forment adgang til med henvisning til, at de deltagende er sikret anonymitet. Vi kan derfor ikke vurdere denne del af udvalgets undersøgelse og dens betydning for dets konklusioner.

Undersøgelsen af kritiske sager siden 1980

Bo Smith-udvalget begrundede denne undersøgelse med, at der i den offentlige debat om centraladministrationen synes at være taget udgangspunkt i antagelser om, at der er kommet mange flere sager, at der er en fælles rød tråd i sagerne, og at de er blevet mere alvorlige. Derfor vil udvalget se på de kritiske sager i et lidt længere tidsperspektiv med henblik på at teste disse antagelser.

Udviklingen i antallet af sager

Den centrale metodiske problematik i forhold til optællingen af antal kritiske sager er spørgsmålet om målevaliditet, som refererer til sammenhængen mellem det begreb, man gerne vil udtale sig om, og den empiriske måling, der foretages. Høj målevaliditet fordrer en præcis og relevant definition af det begreb, man vil undersøge, og en klar overensstemmelse mellem dette begreb og den empiriske måling som foretages (se Adcock og Collier, 2001; Gerring, 2012: 155ff). Målevaliditeten er fundamental for vores evne til at drage konklusioner på baggrund af vores data og analyser (Johnson, Reynolds og Mycoff, 2008: 118).

I denne sammenhæng vil udvalget gerne vide noget om omfanget af fejl og forsømmelser i centraladministrationen, altså omfanget af sager, hvor centrale normer er blevet tilsidesat. Det er imidlertid kun muligt at tælle det antal sager, som bliver offentligt kendt, mens andre sager ikke bliver opdaget, og omfanget

af sådanne uopdagede sager – også kaldet ”mørketallet” – er selvsagt vanskeligt at gisne om og kan variere over tid (se også Christensen, 1996: 256 ff).

Det er efter alt at dømme en kompleks cocktail af faktorer, der afgør, i hvor høj grad problematiske sager kommer til offentlighedens kendskab, herunder omfanget af lækager/whistleblowing, udviklingen i adgangen til aktindsigt, mediernes ressourcer til undersøgende journalistik, hvorvidt de relevante sager handler om kontroversielle ministre eller politikområder mv. Det er helt uklart, hvilken rolle det reelle antal fejl faktisk spiller for det kendte tal. Man kan derfor med god grund hævde, at det er meningsløst at udtale sig om udviklingen i det samlede omfang af kritiske sager på baggrund af en opgørelse af de offentligt kendte sager. Bo Smith-udvalget er opmærksom på problemet, når de anfører, at udvalgets afdækning ikke kan ”tage højde for eventuelle problemsager, som af forskellige grunde går ’under radaren’” (BS, 2015: 42).

Bo Smith-udvalget vælger alligevel at foretage en optælling af ”skandalesager” i perioden 1980-2015. Men man indfører et yderligere kriterium, idet man kun medregner de sager, som har været genstand for en særlig officiel *undersøgelse*. Udvalgets begrundelse for denne fremgangsmåde er en henvisning til, at ”det er ifølge forholdets natur umuligt at vurdere, men det forekommer ikke sandsynligt, at der skulle være en sådan underskov af sager med en anden udvikling end i de sager, der gennem perioden er blevet genstand for undersøgelser. Udvalget har endvidere noteret sig, at hverken ombudsmandens udtalelser eller Rigsrevisionens beretninger giver grundlag for en slutning i denne retning” (BS, 2015: 43), hvilket dog vanskeligt kan ses som et argument, idet Ombudsmanden og Rigsrevisionen jo også i vid udstrækning baserer deres arbejde på sager, der allerede er kommet frem i offentligheden.

Den vigtigste afgrænsning i optællingen – udover den tidsmæssige – er, at udvalgets undersøgelse kun omfatter sager, som har været genstand for *særlige undersøgelser*, dvs. dommerundersøgelser, undersøgelseskommissioner, advokatundersøgelser, undersøgelser i parlamentarisk regi samt undersøgelser i form af udvalg, arbejdsgrupper, revisorundersøgelser, tjenstlige undersøgelser mv. Dette betyder imidlertid, at selv kendte ”problemsager” ikke indgår i optællingen. Af velkendte sager, som ikke indgår i Bo Smith-udvalgets undersøgelse, kan fx nævnes Eritrea-sagen, sagen om overbetaling af privathospitaler, lækagesagen vedrørende jægerkorpset og sagen om hemmelige CIA-overflyvninger af Grønland (Koch og Knudsen, 2014). Heller ikke de sager, som journalist Jesper Tynell afdækkede i Beskæftigelsesministeriet, og som indbragte ham Cavling-prisen, er med i opgørelsen (Tynell, 2015).⁶

Problemet med kun at inddrage sager, som har været genstand for en offentlig undersøgelse, er, at dette antal i udpræget grad er afhængigt af de givne

politisk-parlamentariske forhold i den relevante periode (jf. Christensen, 2006: 258). Hvorvidt der skal nedsættes en undersøgelseskommission eller iværksættes andre særlige undersøgelser er først og fremmest en politisk beslutning, der forudsætter en velvilligt indstillet regering eller et politisk pres fra et flertal i Folketinget. Da man i 1999 vedtog lov om undersøgelseskommissioner, besluttede man ikke – som fx Europarådet anbefaler i resolution 1601, guideline 2.2.8 – at indbygge en mindretalsbeskyttelse, så et kvalificeret mindretal i Folketinget kan kræve en undersøgelseskommission nedsat. Heller ikke andre undersøgelsesformer kan kræves af et parlamentarisk mindretal. Dette betyder, at i perioder med mere stabile flertalskonstellationer – som fx årene 2002-2011, hvor VK-regeringen havde DF som fast støtteparti – vil problematiske sager ikke nødvendigvis blive undersøgt.

Bo Smith-udvalget er for så vidt opmærksom på denne svaghed ved sin egen opgørelse, idet man anfører, at ”der kan være sager, som ud fra en substantiel vurdering kan være potentielt lige så alvorlige, men som ikke bliver gjort til genstand for en sådan undersøgelse, fordi der ikke er politisk tilslutning til det” (BS, 2015: 42) og senere anerkender, at ”det lave antal sager i perioden 2001-2010 skal med stor sandsynlighed ses i lyset af den daværende regerings meget stabile forhold til sit parlamentariske grundlag.”

Det bemærkelsesværdige er imidlertid, at disse erkendelser ikke ændrer ved Bo Smith-udvalgets konklusioner. For eksempel angives det i hovedoverskriften: ”Der er ikke kommet flere sager” og med fed skrift i afslutningen af afsnittet anføres følgende:

Samlet set er konklusionen, at der siden 1980 nok har været udsving i antallet af sager, der er undersøgt, men at der ikke er kommet flere sager end tidligere. Det er naturligvis muligt, men efter udvalgets vurdering ikke videre sandsynligt, at der reelt er en vækst i antallet af sager, men at relativt færre bliver undersøgt. Det er i den forbindelse væsentligt at være opmærksom på, at nutidens ministre i meget høj grad er eksponerede over for såvel Folketingets stående udvalg som medierne.⁷

Problemerne med målevaliditeten i denne del af undersøgelsen er så alvorlige, at det sætter spørgsmålstegn ved udvalgets konklusion. Antallet af sager, som har været genstand for en offentlig undersøgelse, kan ikke tages som et validt mål for omfanget og udviklingen af sager, hvor der er sket brud på embedsmandsnormerne.⁸

Er der en rød tråd i sagerne?

I den næste del af undersøgelsen stiller udvalget et mere kvalitativt orienteret spørgsmål (BS, 2015: 38): "Er der et mønster i sagerne med hensyn til, hvilken slags problemer sagerne handlede om, hvilke konklusioner undersøgelserne kom frem til, og hvilke konsekvenser de fik?" Formålet er at afdække, om sagerne udtrykker et fælles underliggende problem, altså om der går en rød tråd gennem sagerne.

Udvalgets svar er her ganske entydigt: Der er ikke nogle særlige fællestræk i sagerne, der "har beskæftiget sig med vidt forskellige konkrete problemstillinger og berørt mange grene af den statslige forvaltning" (BS, 2015: 45). Nogle sager har handlet om korrekt og fyldestgørende information af Folketinget, men langt fra alle. Nogle sagsforløb har udspillet sig i betydelig afstand fra det ministerielle hierarki, andre ikke. Udvalget pointerer også, at i det omfang undersøgelserne har givet anledning til kritik, har kritikken – med statsløsesagen som en nylig undtagelse – ikke rettet sig mod samspillet mellem ministre og embedsmænd. Endelig konkluderer man, at de gennemførte særlige undersøgelser "i et vist omfang, men langt fra altid, har fundet, at der var begået fejl og forsømmelser i de undersøgte myndigheder. Der er ikke over tid nogen tendens i udviklingen" (BS, 2015: 48).

Et afgørende problem i denne del af undersøgelsen er den manglende replicerbarhed og transparens. Replicerbarhed er et af de mest fundamentale forskningskriterier og indebærer, at analysemetoder og procedurer for dataindsamling og -behandling ekspliciteres, således at andre forskere har mulighed for at gentage undersøgelsen. Transparens sikrer, at man kan følge processen, der leder fra data til konklusion (se blandt andre Andersen, 2012: 97; Gerring, 2012: 92-95). Hvordan de undersøgte sager er klassificeret og ud fra hvilke karakteristika, beskrives dog hverken i hovedrapporten eller i bilagsmateriale. Der er derfor ikke transparens i forhold til, hvordan undersøgelsens data leder til konklusionen. Når analysen ikke fremlægges, er det umuligt at vurdere og efterprøve konklusionen.

Et andet centralt problem berører princippet om undersøgelsesvaliditet, dvs. spørgsmålet om, hvorvidt der er tilstrækkelig overensstemmelse mellem en undersøgelse og dens problemstilling (se blandt andre Gerring, 2012: 95 ff). Det handler grundlæggende om, hvorvidt den empiriske undersøgelse er designet til at kaste lys over den teori eller problemstilling, man ønsker at belyse. Man kan således lave en valid empirisk undersøgelse, hvor resultaterne alligevel ikke giver mulighed for hverken at be- eller afkræfte de hypoteser eller den teori, der er baggrunden for undersøgelsen. Hvis man fx er interesseret at få afdækket folkeskolereformens betydning for læring og laver en state-of-the-art undersø-

gelse af børns trivsel i skolen før/efter reformen, så vil resultaterne kun være perifert relevante. Man argumenterer måske for, at trivsel er væsentlig for læring, men hvis ikke man undersøger læring direkte, så står man med et resultat, som ikke besvarer det spørgsmål, som er stillet.

Et vigtigt baggrundstæppe for udvalgets arbejde er debatten om, hvorvidt de kritiske sager er et udtryk for, at politiske hensyn i stigende grad trumfer de klassiske embedsmandsdyder (se fx BS, 2015: 225). Men i stedet for at se på, om dette aspekt er en fællesnævner ved sagerne, er det tilsyneladende kun andre træk ved sagerne, der undersøges. For eksempel har udvalget set på, *om* sagerne berører problematisk sagsbehandling eller bevillingsspørgsmål (altså på sagens konkrete "emne" eller type af forvaltningssag), *om* de undersøgte forløb har været tæt på ministeren eller længere fra det ministerielle hierarki, eller *om* spørgsmålet om embedsmændenes konkrete rådgivning af ministeren og spørgsmålet om ministrenes information til Folketinget har været dominerende temaer.⁹ Ud fra disse kriterier konkluderer udvalget så, at der er stor spredning i sagerne, og at det "ikke i særlig grad er bestemte typer problemstillinger, som bringer embedsmændene i uføre" (BS, 2015: 49).

Udvalget medgiver dog, at "ser man på de seneste i rækken af sager, har langt de fleste haft et politisk element", men man går ikke videre med at undersøge, om dette også betyder, at politiske hensyn har trumfet klassiske embedsmandsdyder. Det er mærkeligt, at udvalget ikke undersøger dette potentielle fællestræk, idet man i et af rapportens indledningskapitler om kritikken i kølvandet på sagerne anfører, at "Kritikken er formuleret i forskellige varianter, men fælles for dem er den opfattelse, at funktionen som politisk-taktisk hjælper er kommet til at dominere, så fagligheden ligger under for politiske hensyn" (BS, 2015: 33).

Det er ikke muligt på det foreliggende grundlag at acceptere udvalgets konklusion om, at der ingen fælles rød tråd er i de undersøgte sager. Betingelserne for replicerbarhed og transparens er ikke opfyldt, så vi kan ikke følge, hvorledes data og analyse leder frem til konklusionen. Derudover er der et problem med undersøgelsesvaliditeten (overensstemmelsen mellem problemstilling og undersøgelsesdesign), idet udvalget tilsyneladende har undladt at undersøge det mulige fællestræk, at politiske hensyn trumfer embedsmandsnormerne. Denne undladelse betyder øget risiko for type II fejl (eller med andre ord: *missing mountains*), hvor man drager en fejlagtig slutning om, at der ikke er et problem.

Hvor alvorlige konsekvenser fik undersøgelse?

Udvalget gennemgår endelig de gennemførte undersøgelses resultater med henblik på at få et overblik over, hvor alvorlige konsekvenser de fik. De kon-

kluderer, at undersøgelserne langtfra altid har fundet fejl og forsømmelser, og at de kun i begrænset omfang har ført til retlige sanktioner mod de involverede embedsmænd og andre involverede (BS, 2015: 48-49). Dette bruger udvalget som en indikator på, om der er alvorligere problemer i embedsværket, hvilket umiddelbart kan synes rimeligt nok.

Det kan dog indvendes, at man med lov om undersøgelseskommissioner i 1999 fik cementeret et system for undersøgelser, der opererer med meget omfattende retssikkerhedsgarantier for de involverede ministre og embedsmænd, og som også er forpligtet af en ganske markant sikkerhedsdoktrin, når det kommer til at drage mere håndfaste konklusioner om embedsmænds og – særligt – ministres ansvar. En sikkerhedsdoktrin, der forpligter undersøgelseskommissionen til at være forsigtig i forhold til konklusioner om enkeltpersoner (Henrichsen, 2015: 30 ff; Justitsministeriet, 2014: 58-63).

Udover sikkerhedsdoktrinen betød lov om undersøgelseskommissioner også, at de involverede ministre og embedsmænd blev udstyret med omfattende retssikkerhedsgarantier i forbindelse med afhøringerne i en undersøgelseskommission (Betænkning 1315: 15, 103-104), således at de reelt har en beskyttelse, der svarer til den, man som sigtet har i en egentlig straffesag. Det betyder blandt andet, at de involverede ikke har pligt til at udtale sig under vidnepligt, hvis de hermed risikerer at skade deres egen sag og risikerer straf (forbuddet mod selvinkriminering) (Talevski et al., 2002: 104-105). Alle får i øvrigt også tildelt bisiddere (advokater), som rådgiver dem i forbindelse med afhøringer og processen i undersøgelseskommissionen i øvrigt. Disse udførlige retssikkerhedsgarantier indebærer, at der ofte forekommer ganske mange modstridende forklaringer om de relevante begivenhedsforløb, og at der ofte henvises til, at de implicerede ikke kan huske, hvad der mere konkret foregik.

Sikkerhedsdoktrinen kombineret med manglen på vidnepligt (begge elementer indført ved lov nr. 357 af 2. juni 1999) betyder, at undersøgelserne (i det omfang der er tale om egentlige undersøgelseskommissioner) efter 1999 har haft svært ved at nå frem til mere alvorlige og håndfaste konklusioner, nok også sværere end førhen. Dette har direkte betydning for troværdigheden af Bo Smith-udvalgets konklusioner om, hvor alvorlige forhold undersøgelserne af de seneste års skandalesager har afdækket.

Ud fra udvalgets egne optællinger kunne man lige såvel have fokuseret på, at der – på trods af de begrænsede muligheder for at komme med håndfaste konklusioner om enkeltpersoners ansvar – blev påvist fejl og forsømmelser i knap halvdelen af sagerne i perioden 1999-2015 (48 pct.), mens det kun var tilfældet for 40 % af sagerne i perioden 1980-1999. Eller at undersøgelserne blev fulgt op med yderligere procedure med henblik på retlig ansvarsplacering i 23 pct. af

sagerne i den tidlige periode (1980-1999), mens dette var tilfældet for 48 pct. af sagerne i den seneste periode (1999-2015) (BS, 2015: 48 (tabel 4.2.)).

Igen er det dog uigennemskueligt, hvad udvalget baserer sin analyse på, og hvilke kategoriseringer, skøn og vurderinger der er foretaget for de enkelte sagers vedkommende. Der er med andre ord også her problemer i forhold til replicerbarhed og transparenskriterierne. Der fremlægges ingen steder i rapporten en oversigt over, hvordan de i alt 49 sager, som udvalget inddrager i sin undersøgelse, er klassificeret i forhold til det afgørende kriterium om, hvorvidt undersøgelsen påviste fejl og forsømmelser, og om der efterfølgende blev iværksat yderligere undersøgelser eller draget konsekvenser.

Pointen i denne sammenhæng er ikke at fremføre argumenter for, at der nødvendigvis er blevet flere alvorligere sager, men derimod at problematisere grundlaget for Bo Smith-udvalgets opgørelser og den dertil knyttede konklusion. Når udvalget således konkluderer, at de seneste års sager ikke giver grundlag for at hævde, at embedsmændene i højere grad end tidligere misligholder deres forpligtelser, eller at det i særlig grad er bestemte typer problemstillinger, der bringer embedsmændene i uføre, er denne konklusion ikke tilstrækkeligt videnskabeligt underbygget.

Undersøgelsen af udbredelsen af de klassiske embedsmandsnormer

Udvalget har gennemført selvstændige empiriske undersøgelser af ”embedsmænds kendskab til og tolkning af normerne for korrekt embedsmandsadfærd” (BS, 2015:168). Som tidligere nævnt fokuserer vi her på spørgeskemaundersøgelsen blandt embedsmænd på sagsbehandler- og mellemliderniveau i seks ministerier. I alt 3795 embedsmænd fik spørgeskemaet, og der indkom 2611 besvarelser (68,8 pct.). Baggrunden for valget af de seks ministerier begrundes ikke i rapporten, og det kan ikke afvises, at dette valg af population for undersøgelsen kan have indflydelse på generaliserbarheden af resultaterne. Med andre ord antages det, at embedsmænd i de seks ministerier ikke afviger i deres svarmønstre fra ansatte i andre ministerier, når undersøgelsens resultater anvendes til at sige noget om embedsmænd generelt. Det er usikkert, om dette er en korrekt antagelse, men vi har ikke mulighed for at efterprøve den.

Der kan også peges på det potentielt problematiske i, at alle ansatte i ministerierne var inkluderet i populationen (gruppen af mulige *respondenter*) til spørgeskemaundersøgelsen. Problemet med dette er, at de stillede spørgsmål slet ikke er relevante for alle typer ansatte. De fleste spørgsmål fordrer, at man sidder med ministerbetjening, analyse- og policy-opgaver eller sagsbehandling. En gennemgang af materialet viste imidlertid, at stikprøven indeholdt 28 pct.

ansatte med andre stillingsfunktioner end disse, herunder beskæftigede 9 pct. sig med intern økonomistyring og administration, og 19 pct. med forskellige typer IT, konsulentopgaver, laboratoriefunktioner mv. 70 respondenter bemærker selv uopfordret, at de ikke oplever at kunne give relevante svar på de stillede spørgsmål. Vi vender tilbage til dette spørgsmål i sidste afsnit.

Spørgeskemaets udformning – kendskab til normerne

Formålet med spørgeskemaundersøgelsen er at kortlægge ”embedsmændenes kendskab til og deres individuelle tolkning af normerne for korrekt embedsmandsadfærd, når de blev stillet i en situation, der involverede dem i rådgivning af ministeren, og hvor de skulle afveje lydhørhedskravet over for lovligheds-, sandheds- og faglighedskravet” (BS, 2015: 172).

I den ene del af spørgeskemaet søger man at afdække embedsmændenes normkendskab ved at spørge på følgende måde:

Hvilket af nedenstående udsagn er du mest enig i?

“Jeg kender til normerne for embedsværkets rådgivning og bistand, og de kan oftest guide mig, når der opstår dilemmaer eller svære sager, som jeg skal tage stilling til”

eller

“Jeg konsulterer min nærmeste chef, når jeg bliver mødt af svære dilemmaer eller problemstillinger, og vi sparrer om, hvilke løsninger, der kan bringes i anvendelse”

Ud fra et surveymetodisk synspunkt kan man pege på flere problemer ved spørgsmålet. For det første er de to relevante udsagn, som respondenterne skal vælge imellem, flerdimensionelle/tvetydige. Det er ved udformning af spørgeskemaer yderst problematisk, hvis respondenter bliver bedt om at svare på spørgsmål eller tilkendegive en holdning til udtalelser, som indeholder mere end ét enkelt udsagn (Converse og Presser, 1986: 13; se også Olsen, 2006: 26). Der opstår nemlig den situation, at respondenterne kan blive forvirret over, hvad der præcis spørges til: det ene, det andet eller begge to? Det umuliggør samtidig en meningsfuld fortolkning af svarene.

I den første af de to udtalelser, som embedsmændene skal tage stilling til, er der to forskellige udsagn. Det første går på normkendskabet (”jeg kender til normerne”), og det andet om man oplever disse normer som handlingsanvisende i vanskelige situationer (”og de kan oftest guide mig, når der opstår dilemmaer eller svære sager, som jeg skal tage stilling til”). Problemet er her, at det er fuldt ud muligt at kende til normerne og samtidig *ikke* mene, at de kan

“guide” i en vanskelig sag. Det er ikke blot to udsagn, men to udsagn som ikke er logisk forbundne.

Den anden udtalelse inkluderer også to forskellige udsagn. Det første går på, hvorvidt man rådfører sig med sin nærmeste chef (“Jeg konsulterer min nærmeste chef”), og det andet hvorvidt man sparrer med sin chef om valget af løsninger (“og vi sparrer om, hvilke løsninger, der kan bringes i anvendelse”). Problemet er, at man godt kan gå til sin nærmeste chef, når man står i et dilemma, men ikke opleve at der finder en ”sparring” sted og derfor fravælge enighed i udsagnet på det grundlag.

Derudover er det et problem, at respondenterne skal vælge mellem to udtalelser, som ikke er enten modsætninger eller gensidigt udelukkende. Hvis man havde valgt udtalelser, som stod i modsætningsforhold – fx én som udtrykker, at man kender normerne, og én hvor man udtrykker tvivl om normerne – så kunne man bruge svarene til at vurdere, i hvilket omfang embedsmændenes *oplever* at have kendskab til normerne. Men man kan sagtens kende normerne og alligevel ønske at konsultere sin chef. Formuleringen ”svære dilemmaer eller problemstillinger” tyder på situationer, hvor det netop er vanskeligt at træffe en selvstændig afgørelse, og på den baggrund forekommer det vel meget naturligt, at over 80 pct. af respondenterne i hierarkiske organisationer vælger det andet alternativ.

Udformningen af spørgeskemaet, hvor embedsmændene er blevet bedt om at vælge mellem disse to udsagn, giver derfor efter vores opfattelse slet ikke mulighed for at drage meningsfulde konklusioner om embedsmændenes normkendskab.

I Bo Smith-rapporten bliver resultatet imidlertid anvendt til at konkludere, at ”kun 17 pct. af de embedsmænd, der har besvaret udvalgets spørgsmål oplyser, at de kender normerne for rådgivning og bistand ...” (BS, 2015: 173). Senere sammenfattes det, at ”kun et mindretal af embedsmændene giver udtryk for at kende normerne så godt, at de selv ser sig i stand til at tage stilling i de beskrevne dilemmasituationer. De rådfører sig i stedet med deres umiddelbart overordnede”. Og i resuméet konkluderes, at ”der på dette felt er både et behov for og en betydelig efterspørgsel efter at styrke embedsmændenes kendskab til de grundlæggende spilleregler for deres politiske bistand og rådgivning”.

Problemet med disse konklusioner er imidlertid for det første, at vi slet ikke har fået besvaret det opstillede spørgsmål om embedsmændenes kendskab til normerne. For det andet kan man ikke sige noget generelt om, hvorvidt embedsmændene forlader sig mere eller mindre på egen tolkning af normer og implikationer eller støtter sig til deres chefer, fordi det er rent faktisk ikke det, som er blevet undersøgt.

Udover disse meget grundlæggende problematiske forhold vedrørende spørgsmålsformulering og svaralternativer er det bekymrende, at der ikke er valgt en mere relevant metode til at undersøge normkendskabet på. Hvis man gerne vil have en valid måling af ”viden”, er det en tvivlsom metode at spørge de relevante personer direkte til, om de besidder denne viden. Svaret bliver reelt, om embedsmændene oplever/vurderer, at de har kendskab til normerne for rådgivning og bistand – ikke om de rent faktisk har det. Særligt i situationer hvor respondenter oplever et vist pres – fx grundet sociale normer – for at give et bestemt svar, undermineres validiteten af spørgeskemaundersøgelsen. Der opstår en *social-desirability bias*, altså respondenter har en tendens til at give ønskværdige snarere end sande svar (Krumpal, 2012; Fisher, 1993) om deres kendskab til normer for arbejdet.

Vignetter – embedsmandsnormer i praksis

Udvalgets forfattere skriver selv, at spørgeskemaer ikke er det bedste redskab til at kortlægge viden om og efterlevelse af embedsmandsnormer, idet fx grundige casestudier via dybtgående interview eller omfattende aktstudier ville være mere egnede (BS, 2015: 176), men dog meget ressourcekrævende og vanskelige at generalisere ud fra. Derfor har man i stedet valgt den såkaldte vignette metode. Vignetter er korte beskrivelser af virkelighedstro situationer, hvor respondenterne bedes om at angive, hvordan de ville handle i den skitserede situation. Metoden er således velegnet til at ”afdække normative forestillinger i professionelle grupper”, som udvalget skriver (BS, 2015: 177). Et helt centralt problem for undersøgelser af normer i spørgeskemaundersøgelser er nemlig, at respondenter vil have en tendens til svare, som de mener de ”bør” svare ud fra gængse normative forestillinger, snarere end at ærligt give udtryk for de holdninger og den adfærd, de faktisk har. Denne social desirability bias vanskeliggør undersøgelser af denne type. Vignetter er en måde, at få respondenterne til i højere grad at svare ærligt i forhold til direkte spørgsmål, der går på holdning til og overholdelse af normer. Man kan dog naturligvis ikke helt undgå denne bias.

De syv scenarier, som opridses i vignetterne, beskriver udvalget som ægte dilemmaer, ”der har en sådan karakter, at man ikke uden videre kan give et entydigt svar.” De ”beskriver hypotetiske situationer, hvor de forskellige normer trækker i hver sin retning.” Vignetterne skitserer situationer, som bringer normerne om lovlighed, faglighed og sandhed i forvaltningen i spil. Alle vignetterne afsluttes med et spørgsmål om, hvorvidt man ville medvirke til eller udføre en given handling, som den foregående tekst har angivet karakteren af. Der er fire svarkategorier (se nedenfor).

Vignetmetoden åbner nogle meget interessante muligheder for at afdække holdninger, normer og viden, men desværre betyder den aktuelle udformning af vignetterne, at man ikke i tilstrækkelig grad får mulighed for at besvare det, som er undersøgelsens spørgsmål, nemlig ”i hvilket omfang embedsmændene i det daglige efterlever normerne, når de på den ene eller anden måde medvirker til ministerrådgivning og -bistand” (BS, 2015: 176).

Formuleringen af de enkelte vignetter

De to første vignetter, vi vil diskutere (V3 og V5), berører normen om *lovlighed*. For at give et generelt indtryk af vignetternes form citeres disse to dilemmaer i deres helhed:

Figur 1:

Lovlighed

V3: Din minister ønsker meget hurtigt at få ændret nogle regler på det område, du arbejder med. Han mener, at det kan klares ved en bekendtgørelse. Din afdeling vurderer, at det kræver en ændring af loven, da der ikke er tilstrækkelig klar hjemmel til at udstede bekendtgørelsen. Ministeren mener imidlertid, at sagen er for lille til at ulejlige Folketinget med og fastholder, at ændringen må kunne foretages i bekendtgørelsesform. Vil du medvirke til det?

V5: Du behandler en konkret sag og får besked om, at din minister gerne ser, at ansøgeren får medhold. Det er åbenbart fra mediedækningen af sagen, at der er stor offentlig sympati for en sådan løsning, og at ministeren vil komme i strid modvind ved et afslag. Loven og praksis giver efter din vurdering ikke mulighed for at nå det ønskede resultat. Vil du medvirke til at give tilladelsen?

Figur 2:

Svarmuligheder for alle vignetter

1. Ja, uden forbehold
2. Ja, men jeg vil rejse mine betænkeligheder/forbehold over for min leder
3. Kun efter direkte ordre, og jeg vil advare min leder utvetydigt
4. Jeg vil sige klart fra og ikke medvirke

Den måde, lovlighedsvignetterne er udformet på, er dog ikke helt så dilemmafyldt, som udvalget giver indtryk af i sin beskrivelse af vignetternes indhold. Begge scenarier repræsenterer situationer, hvor embedsmænd bør sige nej til at medvirke. Vi har ladet to professorer i forvaltningsret fra Københavns Universitet (Michael Gøtze og Carsten Henrichsen) vurdere de to vignetter, og begge

udtaler, at der ikke er nogen tvivl om, at handlingen er ulovlig, og at embedsmændene derfor skal sige fra. Det er naturligvis korrekt, at loyalitetspligten her kommer i konflikt med lovlighedspligten, men vejledningen i den skitserede situation er ifølge begge professorer klar; nemlig at lovlighedspligten ikke bør vige for politisk loyalitet, og at det derfor vil være helt forkert at vælge svarmulighed 1 eller 2.¹⁰ Det er i lyset heraf problematisk, at Bo Smith-udvalget tilsyneladende ikke anerkender, at de to skitserede situationer har et juridisk korrekt svar, idet udvalget konsekvent karakteriserer alle vignetter som dilemmaer hvor ”forskellige normer trækker i forskellig retning”.

Anderledes forholder det sig med de vignetter, som vedrører *sandheds-* og *faglighedsnormerne* (de resterende fem vignetter). Her er de beskrevne situationer langt mere komplekse og tvetydige. For eksempel handler en af sandhedsvignetterne (V4) om, at en redegørelse til Rigsrevisionen indeholder ”de relevante oplysninger” men i embedsmændens vurdering fokuserer ”noget ensidigt” på de positive sider af ministerens rolle på bekostning af de negative sider. Vignetten skitserer således en situation, hvor der ikke er en klar konflikt med sandheden (ingen løgne), men hvor præsentationen vinkles og dermed ikke fremstår neutral. I de resterende vignetter er der heller ikke en (lige så) klar konflikt mellem det politiske ønske og den relevante grundnorm.

Det kan måske forekomme besynderligt, at vi her begiver os ud i egne fortolkninger af, hvordan dilemmaerne bør fortolkes og besvares. Det skyldes imidlertid, at udvalget ikke selv redegør for, hvad der er korrekt embedsmandsadfærd i de skitserede situationer. Man beder altså embedsmænd om at forholde sig til en række fiktive situationer for at få information om deres anvendelse af embedsmandsnormerne i praksis, men man giver ikke efterfølgende nogle retningslinjer for, hvordan de burde agere. Alligevel anvendes vignetterundersøgelsen til at drage konklusioner om, hvorvidt embedsmændene forstår at anvende normerne i praksis.

Efter vores opfattelse er det kun lovlighedsvignetterne, der egner sig til at vurdere embedsmændenes evne til at anvende relevante normer. Disse vignetter kan teste embedsmændenes faktiske kendskab til og villighed til at efterleve lovlighedsnormen, mens der for de andre vigneters vedkommende ikke kan drages klare konklusioner vedrørende kendskabet til og anvendelsen af fagligheds- og sandhedsnormen. Da vignetterne ikke beskriver samme niveau af konflikt med grundnormen, kan svarene på dem ikke sammenlignes på en meningsfuld måde.

Der er således ikke belæg for at konkludere, at ”der er en iøjnefaldende forskel i embedsmændens oplevelse af hhv. lovligheds-, fagligheds- og sandhedspligten” (BS, 2015: 181), da forskellene mellem vignetterne umuliggør en

sådan sammenligning. Udvalget anerkender selv, at ”dette skyldes delvis, at vignetterne er formuleret, så der kan argumenteres for at lægge vægt på lydhørheden, uden at de går ud over pligten til at tale sandt og pligten til at sikre et godt fagligt grundlag” (BS, 2015: 182). Det vil sige, udvalget erkender selv, at vignetterne ikke kan sammenlignes men lader ikke denne erkendelse slå igennem i deres konklusioner.

Også den anden centrale konklusion, nemlig at ”embedsmanden er således meget tilbageholdende med at tilsidesætte lovlighedskravet”, kan der sættes spørgsmålstejn ved. Ser man nærmere på data på dette spørgsmål, vil det være mere oplagt at drage den stik modsatte konklusion. Selv hvis en forsigtig tilgang vælges (dvs. både svar 3 og 4 accepteres som juridisk korrekte), så er der 26,9 pct. af embedsmændene, som svarer forkert på vignette 3, og 9,8 pct. som svarer forkert på vignette 5. Det betyder med andre ord, at de direkte ville medvirke til en ulovlig handling. Dette forekommer at være en foruroligende høj andel også set i lyset af en forventning om social desirability bias i svarene, som betyder, at en højere andel vil angive det normativt ønskelige svar, selvom det ikke nødvendigvis afspejler, hvordan de ville handle i virkeligheden. Vi ser nærmere på svarene på lovlighedsvignetterne i det næste afsnit.

Nærmere undersøgelse af lovlighedsnormen

Det kan bemærkes, at stikprøven af ansatte i centraladministrationen som udgangspunkt ikke skelner mellem embedsmænd med relevante funktioner (fx ministerbetjening, analyse- og policyopgaver samt sagsbehandling) og embedsmænd med irrelevante funktioner i forhold til de skitserede dilemmaer (fx IT-support, laboratoriearbejde og visse konsulentfunktioner). Det kunne derfor tænkes, at de høje andele forkerte svar om lovlighedsnormen skyldes svar fra respondenter, som ikke sidder i relevante stillingsfunktioner og dermed ikke kan forventes at kende pligterne. Vi gennemgik derfor datasættet og sorterede respondenter fra, som ikke har relevante arbejdsopgaver, og gentog analyserne med kun de relevante respondenter. Resultaterne af de nye analyser, hvor antal besvarelser på begge vignetter er reduceret med cirka 20 pct., er gengivet i tabel 1. Tabel 1 præsenterer andelen af respondenter, der giver svar, som ikke er juridisk korrekte (dvs. angiver svarmulighed 1 eller 2). Tabellen indeholder fire bivariate analyser, hvor sammenhængen mellem henholdsvis primære arbejdsopgaver, uddannelse, ansættelsessted (departement/styrelse), personaleansvar og andelen af juridisk forkerte svar præsenteres.

Af tabellen kan vi se, at det ikke ændrer det samlede billede særligt meget, hvis vi frasorterer irrelevante respondenter. Der er stort set den samme høje andel embedsmænd (25,4 og 8,9 pct. i forhold til 26,9 og 9,4 når alle respon-

Table 1: Lovlighedsvignetterne

	Hjemmel (V3)		Sagsafgørelse (V5)	
	Medvirke ulovligt (pct.)	n	Medvirke ulovligt (pct.)	n
Relevante respondenter	24,5	1758	8,9	1731
Primære opgaver				
Ministerbetjening	19,4	361	7,3	356
Analyse og policy	22,2	568	8,0	556
Sagsbehandling	34,3	501	10,9	495
Andet (evt. kombination)	24,3	328	9,3	324
Uddannelse				
Jur.	15,5	473	6,6	469
Polit./econ.	18,7	139	3,6	137
Scient.pol.	23,3	296	5,8	291
Anden samfundsvid.	31,9	282	13,4	277
Anden universitetsudd.	33,8	507	16,3	55
Andet	31,0	58	16,3	55
Departement/styrelse				
Department	15,0	433	5,4	424
Styrelse	28,8	1325	10	1307
Personaleansvar				
Chef	16,5	237	6,0	234
Medarbejder	26,9	1503	9,4	1479
Arbejdssted (anonym)				
Sted med højeste andel	39,1	87	14,5	152

Note: V3: Variablen primære opgaver er signifikant ($p < 0,05$), og der er signifikante forskelle mellem sagsbehandling og ministerbetjening; sagsbehandling og analyse og policy; sagsbehandling og andet. Uddannelsesvariablen er signifikant ($p < 0,05$), og der er signifikante forskelle mellem jur. og anden samfundsvid.; jur. og anden universitetsudd.; jur. og andet; polit./econ. og anden samfundsvid.; polit./econ. og andet universitetsudd.; scient.pol. og anden samfundsvid.; scient.pol. og anden universitetsudd. Departement/styrelse og personaleansvar er også signifikante.

V5: Variablen primære opgaver er ikke signifikant ($p < 0,05$). Uddannelsesvariablen er signifikant, og der er signifikante forskelle mellem jur. og anden samfundsvid.; jur. og anden universitetsudd.; jur. og andet; polit./econ. og anden samfundsvid.; polit./econ. og anden universitetsudd.; polit./econ. og andet; scient.pol. og anden samfundsvid.; scient.pol. og anden universitetsudd.; scient.pol. og andet. Departement/styrelse-variablen er også signifikant. Personaleansvar er ikke signifikant.

denter regnes med), som vælger at medvirke til en ulovlig handling i de to skitserede scenarier. Enkelte “arbejdssteder” er andelen så høj som 39,1 og 14,5 pct.

Hvis man udelukkende fokuserer på de personer, som rent faktisk angiver at arbejde med opgaver meget tæt på de skitserede problemstillinger (nemlig ministerbetjening for V3 og sagsbehandling for V5), ser det ikke meget bedre ud. Godt nok er personerne, som primært arbejder med ministerbetjening, signifikant bedre end sagsbehandlerne til at anvende lovlighedsnormen i forhold til spørgsmålet om hjemmel i lov. Men det er stadig omkring en ud af fem, som foretager en forkert vurdering, hvilket må siges at være en meget høj andel. Ser vi på de personer, som primært arbejder med sagsbehandling, kan vi ikke se en tilsvarende signifikant forskel mellem deres evne til at anvende lovlighedsnormen i forbindelse med en sagsafgørelse i forhold til andre embedsmænd.

Andelen af medarbejdere i departementet, som giver det juridisk korrekte svar, er signifikant højere end i styrelserne. Der er for begge vignetter cirka dobbelt så mange fejlvurderinger i styrelserne sammenlignet med departementerne, hvilket kan skyldes den generelt stærkere ekspertise på det retlige felt i departementerne (se BS, 2015: 184). Ser vi på respondenternes placering i hierarkiet, er det på sin vis betryggende, at signifikant flere chefer svarer rigtigt på det fiktive lovlighedsdilemma om hjemmel end deres ansatte, mens der ikke er signifikant forskel i forhold til vignetten om sagsafgørelser. Det er dog stadig relativt høje andele – 16,5 pct. henholdsvis 6 pct. – af cheferne, der direkte angiver at ville medvirke til noget ulovligt. Uddannelse viser sig også at have signifikant betydning for anvendelsen af lovlighedsnormen, hvor juristerne ikke overraskende klarer sig bedst, men dog ikke signifikant forskelligt fra polit./econ.uddannede. Det vil dog kræve multivariate analyser at undersøge, hvilke faktorer der viser sig at have betydning, når der udføres kontrol.

Konklusion og perspektiver

Samlet set peger vores kritiske gennemgang af Bo Smith-udvalgets undersøgelser på, at data og undersøgelser ikke giver belæg for udvalgets meget klare konklusioner om, at der ikke er grund til bekymring om forholdene i centraladministrationen, og at embedsmandsnormerne kan konstateres efterlevet i betydelig udstrækning.

I tabel 2 opsummeres resultaterne af gennemgangen af de enkelte undersøgelser og de konstaterede brud på væsentlige forskningskriterier. Endvidere angiver vi, om gennemgangen giver grundlag for reviderede konklusioner, og om nye undersøgelser kan anbefales

I forhold til undersøgelsen af de kritiske sager stødt vi ind i problemer med undersøgelsesvaliditet, målevaliditet, replicerbarhed og transparens-kriterierne.

Table 2: Konklusioner, brud på forsikringskriterier, reviderede konklusioner og anbefalinger

Undersøgelse	Udvalgets konklusion	Brud på forskningskriterier	Revideret konklusion	Ny undersøgelse?
I. Kritiske sager				
Antal	Ingen stigning	Målevaliditet	Ingen grundlag for konklusion om udviklingen i antal sager	Yderst vanskeligt/umuligt at identificere valid indicator
Fællesnævner	Ingen fællesnævner som indikerer underliggende problem	Replicerbarhed og transparens; undersøgelsesvaliditet	Ikke undersøgt om politiske motiver/pres er fælles problem	Undersøgelsen bør gentages
Konsekvenser	Ikke altid alvorlige konsekvenser, ingen udvikling	Replicerbarhed og transparens; målevaliditet	Ikke grundlag for konklusion om udviklingen i sværhedsgrad af konsekvenser	Resursekrævende og vanskeligt at lave en valid vurdering af sværhedsgrad
II. Embedsmandsnormer				
Kendskab til normer	Et mindretal har tilstrækkeligt kendskab til normer	Undersøgelisesvaliditet	Ikke grundlag for konklusion om normkendskab	Undersøgelse af politisk pres i praksis
Normer versus lydhørhed	Forskel på vægtning af de tre normer; tilbageholdenhed med at tilside sætte lovlighedskrav	Målevaliditet	1. Høj andel embedsmænd vil medvirke til ulovlige handlinger. 2. Svar på forskellige typer vignetter kan ikke sammenlignes. 3. Der kan ikke konkluderes noget klart på baggrund af vignetter om sandhed og faglighed	Undersøgelsen bør gentages

I den første delundersøgelse om antallet af sager siden 1980 er målevaliditetsproblemerne så grundlæggende for problematikken, at vi ikke mener, at en ny "optælling" umiddelbart kan løse problemet. Den anden del af undersøgelsen om eventuelle fællestræk ved sagerne kunne potentielt set give interessante indsigter, men dette ville for det første kræve en gennemgang af alle offentligt kendte sager (og ikke kun dem, som er blevet genstand for en offentlig undersøgelse) og for det andet forudsætte inddragelse af "politisk pres" som en mulig fællesnævner sammen med andre relevante årsager til brud på embedsmandsnormerne (som fx "berigelse", "faglig inkompetence", "uklar kompetencefordeling" mv.). Ved en eventuel gentagelse af den sidste del af undersøgelsen om sagernes alvorlighed bør analysen foretages ud fra de samme principper for alle sager i perioden.

Hvad angår undersøgelsen af embedsmændenes kendskab til normerne og anvendelsen i praksis, kan man ikke drage de slutninger, som udvalget gjorde på det foreliggende grundlag. Vi har først og fremmest konstateret alvorlige metodiske problemer med udformning af spørgsmål og vignetter i spørgeskemaet.

I den første del af spørgeskemaundersøgelsen om embedsmandsnormer var det særligt spørgsmålenes udformning, der skabte problemer. Målevaliditeten er så svag, at man ikke kan konkludere noget om respondenternes normkendskab. Der er således ikke belæg for at sige, at kun et mindretal af embedsmændene har tilstrækkeligt kendskab til normerne. Det kræver en ny undersøgelse.

Den mest oplagte konklusion på den anden del af spørgeskemaundersøgelsen forekommer i modsætning til udvalgets udlægning at være, at en stor andel af embedsmændene tilsyneladende er villig til at medvirke til ulovlige handlinger. Og det er vel at mærke i en hypotetisk situation med et fiktivt eksempel, hvor respondenterne ikke helt konkret oplever at være udsat for pres. Om den høje andel medvirkende til lovbrud skyldes manglende kendskab til normerne, eller at embedsmændene oplever ikke at kunne sige fra i praksis, er det ikke muligt at sige noget om, men det vil være oplagt at dykke ned i den problemstilling i en videre undersøgelse.

Det er tydeligt, at Bo Smith-udvalget har sat fokus på nogle væsentlige problemstillinger, men også at det vil kræve nye undersøgelser at få valid viden om de centrale spørgsmål, som rejses. I den forbindelse er det efter vores opfattelse vigtigt, at man også fokuserer på, hvordan forholdene udspiller sig i praksis og ikke kun i teorien. Oplever embedsmændene i centraladministrationen at blive udfordret på overholdelse af embedsmandsnormerne i varetagelsen af deres opgaver? Det kan undre, at der er et totalt fravær af spørgsmål i spørgeskemaet, som har til formål at afdække, om embedsmændene rent faktisk oplever for-

skellige typer pres for at tilsidesætte normerne. Det ville ligeledes være oplagt at undersøge, hvorvidt de oplever at kunne sige fra.

Vi håber med denne artikel at have udfordret den opfattelse, at vi med Bo Smith-udvalgets undersøgelser har fået valide svar på spørgsmålet om efterlevelse af embedsmandsnormer i centraladministrationen. Der er brug for at sætte fokus på dette vigtige emne og et oplagt behov for nye undersøgelser.

Noter

1. Bo Smith-udvalget blev nedsat og finansieret af DJØF i foråret 2014. DJØF udpegede 12 personer, der tilsammen repræsenterede tung erfaring fra politik, embedsværk, medier og forskningsverden.
2. Af vigtige bidrag kan nævnes: Nordskov Nielsen-udvalgets rapport *Fagligt-etiske principper i offentlig administration* (DJØF's fagligt etiske arbejdsgruppe 1993), Betænkning nr. 1315 (1996), Christensen (1997), Betænkningerne 1354 (1998), 1443 (2004) og 1537 (2013).
3. Da Jens Peter Christensen i sin disputats om ministeransvar fra 1997 foretog en opgørelse af formelle uafhængige undersøgelser af politisk-administrative skandalesager 1980-1993 og underkastede disse sager en nærmere analyse, var det med henblik på at udlede de retlige normer for embedsværkets samspil med ministrene og i forlængelse heraf undersøge, om der var sket en øget "retliggørelse" af politiske ansvarsforhold. Denne undersøgelse er således ikke sammenlignelig med Bo Smith-udvalgets undersøgelse af undersøgelsessager, der har til hensigt at måle omfanget af skandalesager i forvaltningen og deres grad af alvorlighed.
4. Den tredje og sidste undersøgelse – om embedsværkets tidsanvendelse og prioritering af opgaver – handler ikke direkte om embedsmandsnormer og vil derfor ikke blive behandlet her.
5. Da Nordskov Nielsen-udvalget i forbindelse med sit arbejde i begyndelsen af 1990'erne afholdt en række høringer med deltagelse af en række DJØF-grupper mv. (fx fuldmægtige fra centraladministrationen, chefer i centraladministrationen, forhenværende departementschefer, fhv. ministre, nuværende og forhenværende ombudsmænd og socialchefer), valgte man at sammenfatte høringsdeltagernes erfaringer og vurderinger i betænkningens kapitel III, dog således at der i anonymiseret form blev foretaget en sammenfatning af høringsdeltagernes udtalelser om typer af etiske problemer i deres daglige arbejde, hvilket var et af fokuspunkterne for Nordskov Nielsen-udvalgets arbejde
6. Udvalget har af naturlige årsager ikke inddraget Tibet-sagen og Carl Holst-sagen i deres optælling af problemsager, da sagerne først kom til offentlighedens kendskab efter offentliggørelsen af udvalgets rapport.

7. Den markante konklusion, at ”der ikke er kommet flere sager”, spillede også en fremtrædende rolle på pressemødet, hvor Bo Smith-udvalget fremlagde sit arbejde, og blev bredt refereret i pressedækningen (se fx <http://www.alinget.dk/artikel/clus-hjort-vi-har-en-staerk-embedskultur>).
8. Det kunne herimod indvendes, at vi ikke burde bebrejde udvalget for dets konklusioner, så længe modstående synspunkter og forbehold fremgår af selve rapportens brødtekst. Kritikere kunne pege på, at formidlede rapporter altid indeholder forsimplinger i de konkluderende afsnit. Når vi ikke desto mindre fastholder kritikken, er det for det første, fordi konklusionen forekommer vildledende i lyset af det empiriske materiale, den bygger på. For det andet er det i praksis kun de overordnede konklusioner, som lever videre i den offentlige debat og informerer beslutningstagere, og det har derfor vidtrækkende konsekvenser, hvis der ikke er dækning for dem.
9. Til sammenligning forsøger Jens Peter Christensen (1996: 259 ff) at kategorisere de undersøgte sager 1980-1993 med udgangspunkt i, ”hvilke forhold der i de respektive undersøgelser er blevet lagt de involverede embedsmænd til last”, altså ”karakteren af de begåede normbrud”, og opererer på den baggrund med kategorier, der knytter sig tættere til de relevante embedsmandsnormer.
10. Præciserende kan bemærkes, at Michael Gøtze angiver, at svarmulighed 4 bør vælges i begge situationer, mens Carsten Henrichsen angiver svarmulighed 4 og 3 som korrekte svar på de to vignetter

Litteratur

- Adcock, Robert og David Collier (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review* 95 (3): 529-546.
- Andersen, Lotte Bøgh (2012). Forskningskriterier, pp. 97-114 i Lotte Bøgh Andersen, Kasper Møller Hansen og Robert Klemmensen (red.), *Metoder i statskundskab*. København: Hans Reitzels Forlag.
- Betænkning nr. 1315 (1996). Betænkning om undersøgelsesorganer. København: Justitsministeriet.
- Betænkning nr. 1354 (1998). *Forholdet mellem ministre og embedsmænd*. København: Finansministeriet.
- Betænkning nr. 1443 (2004). *Embedsmænds rådgivning og bistand*. Finansministeriet.
- Betænkning nr. 1537 (2013). *Ministrenes særlige rådgivere – et serviceeftersyn*. København: Finansministeriet.
- Bo Smith-udvalget (2015). *Embedsmanden i det moderne folkestyre*. København: Jurist- og Økonomforbundets Forlag

- Christensen, Jens Peter (1996). Skandalesager og embedsmandsansvar. *Politica* 28 (3): 255-270.
- Christensen, Jens Peter (1997). *Ministeransvar* København: Jurist- og Økonomforbundets Forlag.
- Christensen, Jørgen Grønnegård, Julie Grunnet Wang og Niels Opstrup (2016). Embedsmandsdyderne og det ministerielle hierarki. *Politica* 48 (4): 431-454.
- Converse, Jean M og Stanley Presser (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Newbury Park, CA: Sage Publications.
- DJØF's fagligt etiske arbejdsgruppe (1993). *Fagligt-etiske principper i offentlig administration*. København: Jurist- og Økonomforbundets Forlag.
- Fisher, Robert J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research* 20 (2): 303-315.
- Gerring, John (2012). *Social Science Methodology: A Unified Framework*. Cambridge: Cambridge University Press.
- Henrichsen, Carsten (2015). Ministeransvar og embedsmænds ansvar. *Tidsskriftet Futuriblerne* 43 (2-3): 26-32.
- Johnson, Janet, H.T. Reynolds og Jason Mycoff (2008). *Political Science Research Methods* (6. udg.). Washington: CQ Press.
- Justitsministeriet (2014). *Beretning afgivet af Skattesagskommissionen*. Albertslund: Rosendahls – Schultz Distribution.
- Koch, Pernille Boye og Tim Knudsen (2014). *Ansaret, der forsvandt – om magten, ministrene og embedsværket*. Frederiksberg: Samfundslitteratur.
- Krumpal, Ivar (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant* 47: 2025–2047.
- Loft, Peter og Jørgen Rosted (2016). *Hvem har ansaret? Revner og sprækker i det danske embedsmandssystem*. København: Gyldendal.
- Olsen, Henning (2006). *Guide til gode spørgeskemaer*. København: Socialforskningsinstituttet.
- Smith, Bo og Jørgen Grønnegård Christensen (2016). Embedsmænd som dydens vogtere eller demokratiets tjenere. *Politica* 48 (4): 373-391.
- Talevski, Oliver (2002). *Undersøgelseskommissioner, embedsmandsansaret, Folketingets rolle*. København: Jurist- og Økonomforbundets Forlag.
- Tynell, Jesper (2014). *Mørkelygten. Embedsmænd fortæller om politisk tilskæring af tal, jura og fakta*. Frederiksberg: Samfundslitteratur.