

Større forvaltningsreformer og udfordringerne ved instrumentel læring

En analyse af evalueringen af NAV-reformen

Breidahl, Karen Nielsen; Furubo, Jan-Eric; Halvorsen, Anne; Hansen, Morten Balle

Published in:
Tidsskrift for Velferdsforskning

DOI (link to publication from Publisher):
[10.18261/issn.2464-3076-2016-04-02](https://doi.org/10.18261/issn.2464-3076-2016-04-02)

Creative Commons License
CC BY-NC 4.0

Publication date:
2016

Document Version
Også kaldet Forlagets PDF

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Breidahl, K. N., Furubo, J.-E., Halvorsen, A., & Hansen, M. B. (2016). Større forvaltningsreformer og udfordringerne ved instrumentel læring: En analyse af evalueringen af NAV-reformen. *Tidsskrift for Velferdsforskning*, 19(4), 305-320. Artikel 2016/4. <https://doi.org/10.18261/issn.2464-3076-2016-04-02>

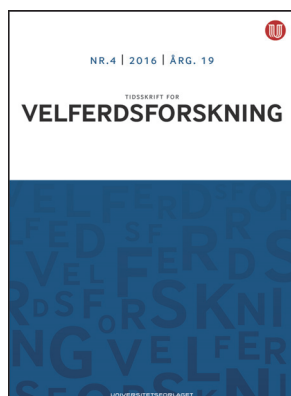
General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



Tidsskrift for velferdsforskning
Årgang 19, Nr. 4-2016, s. 305–320
ISSN online: 2464-3076

DOI: 10.18261/issn.2464-3076-2016-04-02

FAGFELLEVDERT ARTIKKEL

Større forvaltningsreformer og udfordringerne ved instrumentel læring:

En analyse af evalueringen af Nav-reformen

Karen N. Breidahl

Adjunkt, ph.d., Institut for Statskundskab, Aalborg Universitet, knb@dps.aau.dk

Jan-Eric Furubo

Revisionsråd, Riksrevisionen, jan-eric.furubo@riksrevisionen.se

Anne Halvorsen

Førsteamanuensis, Ph.d., Institutt for sosiologi og sosialt arbeid, Universitetet i Agder, anne.halvorsen@uia.no

Morten Balle Hansen

Professor, Aalborg Universitet, mbh@dps.aau.dk

SAMMENDRAG:

I 2005 blev Nav-reformen, en af de største forvaltningsreformer i Norge i nyere tid, vedtaget af det norske Storting. Det blev samtidig besluttet, at der parallelt med reformen skulle gennemføres en evaluering. Den havde flere målsætninger bl.a. at understøtte og korrigere implementeringen af Nav-reformen. Artiklen forholder sig kritisk til den udbredte ambition om, at evalueringer kan anvendes i praktiske handlesituationer og indkredser de særlige udfordringer, der kan være forbundet med instrumentel evalueringsanvendelse i forbindelse med større forvaltningsreformer. Der tages afsæt i en teoretisk analyseramme, der både inddrager indsigter fra evaluering- og policy-læringslitteraturen.

Stikkord

Evalueringsanvendelse, Nav-reformen, meta-evaluering, velfærdsreformer, policy-læring, evalueringskultur

ABSTRACT:

Large-Scale Public-Sector Reforms and the Challenges Related to Instrumental Policy Learning: An Analysis of the Evaluation of the NAV Reform

The NAV-reform was adopted in 2015 by a broad majority of the Norwegian Storting [Parliament]. At the same time it was decided to evaluate the reform. It was based on multiple purposes of which one was to follow the process in order to suggest adjustments and assist a successful implementation. However, afterwards it was evident that there was a huge gap between


UNIVERSITETSFORLAGET

 **Idunn**
Nordiske tidsskrifter på nett

This article is downloaded from www.idunn.no. © 2016 Author(s). This is an Open Access article distributed under the terms of the Creative Commons CC-BY-NC 4.0 License (<http://creativecommons.org/licenses/by-nc/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for non-commercial use, provided the original author and source are credited.

the formal ambitions for instrumental use of the evaluation and the actual utilization. In this article, we critically assess this topic and shed light on the question of which challenges are involved in instrumental use of evaluations when evaluating large-scale public sector reforms. This discussion is based on theoretical insights from research on evaluation and policy learning.

Keywords

evaluation use, NAV-reform, meta-evaluation, welfare reforms, policy learning, evaluation culture

INDLEDNING

Evalueringer i den offentlige sektor har de seneste 30 år vundet stor udbredelse og særligt har den instrumentelle tro på, at evalueringer kan anvendes til at levere konkrete handleanvisninger været fremherskende blandt administrative- og politiske beslutningstagere. Det er imidlertid gang på gang blevet dokumenteret, at evalueringer sjældent er gearret til at leve op til disse ambitioner. Evalueringsforskningen rummer forskellige syn på denne problematik, herunder hvorvidt fraværet af instrumentel anvendelse bør anses som problematisk eller nærmere betrykkende. På den ene side har vi de evalueringsforskere (særligt amerikanske) der, med afsæt i idealet om instrumentel rationalitet og et objektivi perspektiv på problemløsning, siden 1970'erne har anset det som ønskværdigt at fremme anvendelsen af viden fra evalueringer i praktiske handlingssituationer (Alkin og Taut 2003; Johnson, Greenesid, Toal, King, Lawrenz og Volkov 2009; Leviton og Hughes 1981; Mark og Henry 2004; Patton 1997; Patton mfl. 1975). På den anden side har vi en mere kritisk position, som anser idealet om instrumentel evalueringssanvendelse som urealistisk, da politiske beslutningsprocesser er præget af ustabilitet, uklare mål samt omskiftelige præferencer. Det anses endvidere som betrykkende, at instanser uden for det politiske rum ikke har patent på den objektivi korrekte viden (Albæk, 2001:43; se også Dahler-Larsen og Larsen 2001; Furubo 1994; Schwandt 2015).

Med afsæt i et dybdegående casestudie af evalueringen af Nav-reformen (EVANAV) har denne artikel til formål at bidrage til spørgsmålet: *Hvilke udfordringer er der forbundet med instrumentel evalueringssanvendelse i forbindelse med større forvaltningsreformer?*

Nav-reformen kan betegnes som en større velfærdsstatslig forvaltningsreform kendetegnet ved kompleksitet, flertydige målsætninger og foranderlige omgivelser, og casen repræsenterer et eksempel på, at instrumentel evalueringssanvendelse har svære vilkår. Størstedelen af litteraturen på dette område fokuserer på evalueringssanvendelse i forbindelse med konkrete indsatser og programmer (Breidahl, Gjelstrup, Hansen og Hansen 2016) og artiklen bidrager dermed med nye indsigter.

Nav-reformen bestod i en omfattende reform af arbejds- og velfærdsadministrationen og er blevet betegnet som af de største og mest omfattende

politisk-administrative reformer i Norge i nyere tid. Tre forskellige tjenester – tidligere Trygdeetaten, Aetat og den kommunale social tjeneste – blev samlet i en forvaltning. Ambitionen var en styrkelse af arbejdslinjen (flere i arbejde og færre på overførelsesindkomster), bedre brugerorientering og større effektivitet (St. meld. nr. 14 (2002–2003)). Reformen blev vedtaget af det norske Storting i 2005 og i den forbindelse blev det også besluttet, at der parallelt med reformen skulle gennemføres en omfangsrig evaluering. Ambitionerne, som de blev formuleret i Stortingsmeldingen, var store, da der var:

... et ønske om å benytte forskning som redskap for å fremskaffe relevant kunnskap i arbeidet med oppfølging og videreutvikling av Nav-reformen. Ambitionen var endvidere, at evalueringen skulle bidrage med kontinuerlige impulser og kritiske blikk utenfra under reformgjennomføringen med henblik på å trekke lærdommer underveis (St.prp. nr. 46 2004–2005:79).

I forbindelse med udarbejdelsen af et såkaldt mål- og rammedokument i 2006 blev disse ambitioner omsat til fire konkrete målsætninger: 1) Dokumentation, 2) korrektion, 3) indsigter om organisatorisk integration og 4) langsigtet kundskabsudvikling (Arbejds- og inkluderingsdepartementet 2006:3–4).

Særligt 'korrektion' og 'indsigter om organisatorisk integration' knytter an til ambitionen om at levere konkrete handleanvisninger samt give input til de politisk-administrative beslutningsprocesser med henblik på at understøtte en succesfuld realisering og implementering af reformen. Idealet om instrumentel anvendelse stod imidlertid langt fra alene, da målsætningerne om 'dokumentation' samt 'langsigtet kundskabsudvikling' også figurerede som centrale. I og med at der i denne artikel er særligt fokus på (den manglende) realisering(en) af den instrumentelle målsætning, bidrager casen af EVANAV også til belysning af en bredere forskningsdebat. Hertil kommer, at særligt denne målsætning blev genstand for diskussion efter evalueringens afslutning, da der var udbredt enighed om, at EVANAV ikke formåede at leve op til denne ambition (Halvorsen, Breidahl, Furubo og Hansen 2016). I analysen vender vi tilbage til målsætningernes indbyrdes relation.

De følgende afsnit gennemgår den teoretiske analyseramme og det empiriske materiale. Derefter følger de empiriske analyser. Det sidste afsnit opsamler og diskuterer resultaterne i et bredere komparativt perspektiv.

TEORETISK ANALYSERAMME

Den teoretiske analyseramme artiklen anvender til indkredsning af udfordringerne forbundet med praktisk evalueringsanvendelse i forbindelse med større forvaltningsreformer hviler på en antagelse om, at det ikke er tilstrækkeligt at tage højde for udformningen af den konkrete evaluering (organise-

ringen), dens gennemførelse samt resultaterne og karakteren af den leverede viden. De politiske omgivelser har også betydning – og det i særdeleshed når vi har at gøre med større forvaltningsreformer præget af omskiftelighed, kompleksitet og flertydige målsætninger. Artiklens teoretiske analyseramme (operationaliseret i figur 1) inddrager derfor både indsigter fra evalueringslitteraturen samt policy-læringslitteraturen. I forhold til førstnævnte trækkes der både på indsigter fra den klassiske ('snævre') samt den mere kritiske position jf. indledningen. Indsigter fra policy-læringslitteraturen er medtaget, da den bl.a. giver et indblik i, hvad der betinger politisk læring samt betydningen af ustabile politiske omgivelser for vidensanvendelse (Bennett og Howlett 1992; Breidahl 2012; Hall 1993; Sabatier og Jenkins-Smith 1993).

Figur 1: Analyseramme

- | |
|--|
| 1. De politiske vilkår for policy læring |
| 2. De administrative vilkår for policy læring |
| 3. Evalueringsorganisering og graden af forskningsfrihed |
| 4. Evalueringsoutput: Resultater og formidling |

Policy-læringsteoretikere har udfordret politologiens fremherskende tro på, at politik alene er en afspejling af interessevaretagelse og fordeling af magtressourcer: Politiske og administrative beslutningstagere anvender også viden og lærer af fortiden. Der hersker imidlertid langt fra enighed om, hvad der skal forstås ved policy-læring, herunder hvem der lærer, hvad der læres samt den bredere betydning af policy-læring (Bennett og Howlett 1992; Hall 1993; Sabatier og Jenkins-Smith 1993). Denne artikel tager afsæt i Peter Halls definition fra 1993, hvor policy-læring defineres som et overlagt forsøg på at tilpasse policymål og teknikker, som reaktion på tidligere erfaringer eller ny information (Hall 1993:278). Dermed afgrænser vi os fra individuel og organisatorisk læring. Hall sonderer mellem første-, anden og tredje-gradsforandringer og her er der fokus på de to førstnævnte, da der er tale om mindre justeringer af den allerede førte politik. EVANAVs instrumentelle målsætning var netop at lære af fortiden for på den baggrund at give input til de politisk-administrative beslutningsprocesser med henblik på at understøtte en succesfuld realisering og implementering af reformen. Ifølge Hall er der med disse typer af policy-forandringer gode forudsætninger for policy-læring internt i statsapparatet. Vi vil i forlængelse heraf argumentere for, at det også fordrer tilstedeværelsen af favorable mulighedsbetingelser. Antagelsen er derfor, at policy-læringen har bedre udfoldelsesmuligheder under nogle forhold end under andre (Breidahl 2012:124–135). Mulighedsbetingelser henviser både til de politiske rammevilkår og statsapparatets administrative kapacitet til at igangsætte, indsamle og omsætte viden. En indkredsning af mulighedsbetingelserne for policy-læring er afsættet for de to første punkter i analyserammen (figur 1).

1) De politiske læringsvilkår henviser til bevæggrunden(e) for en given evaluerings igangsættelse. Hvad var det politiske formål med evalueringen

(var der et entydigt formål?), var interesserne på forhånd veldefinerede og herskede der politisk konsensus op til vedtagelsen af reformen og i forbindelse med den efterfølgende implementering? (Breidahl 2012; Sabatier og Jenkins-Smith 1993). Hvis interesserne er for fastlåste kan det f.eks. have betydning for betingelserne for evalueringens udformning. De politiske læringsvilkår henviser også til, hvorvidt der var politisk efterspørgsel efter viden til anvendelse i praktiske handlingssituationer. Viden frembragt i evalueringer repræsenterer kun én ud af en række videnskilder, som er nødvendig for politisk beslutningstagning (Schwandt 2015:5). Blev andre videnskilder anvendt, og i så fald, hvad var forholdet mellem EVANAV og andre videnskilder? Belysningen af de politiske læringsvilkår åbner også op for indsigter fra de mere kritiske dele af evalueringsslitteraturen, jf. indledningen herunder pointerne om, at beslutningsprocesser kan være præget af alt andet end rationalitet og veldefinerede interesser.

2) Policy-læringens mulighedsbetingelser afhænger også af statsapparatets interesse i at lære, igangsætte viden og efterfølgende anvende den samt muligheden for at trække på administrative ressourcer som f.eks. information, analyser og ekspertise (Skocpol 1985:12). Antagelsen er dermed, at fraværet af disse forhold kan udfordre mulighedsbetingelserne for vidensanvendelse. Hvorvidt disse forudsætninger var til stede i forbindelse med EVANAV afdækkes i analysen af de administrative vilkår for policy-læring.

3) Det tredje punkt i analyserammen vedrører EVANAVs organisering, valget af evaluatore samt opdragsgivers rolle i evalueringen: Hvem sættes til at gennemføre evalueringen og hvem har det overordnede ansvar? Hvad er graden af forskningsfrihed og hvor stramt defineret var opdraget? Som nævnt indledningsvist har særligt den instrumentelle del af evalueringsslitteraturen i en årrække kredset om, under hvilke omstændigheder evalueringer har gode forudsætninger for at blive anvendt i praktiske handlingssituationer. På baggrund heraf er en række barrierer (samt konkrete handleanvisninger) blevet identificeret. Bl.a. er der blevet peget på vigtigheden af, at evalueringen er organiseret og designet i overensstemmelse med de overordnede formål samt af at målrette den løbende feedback og dermed raffinnere kommunikationsgangene (Alkin og Taut 2003; Johnson mfl. 2009; Leviton og Hughes 1981; Mark og Henry 2004; Patton 1997; Patton mfl. 1975).

4) Det sidste punkt i analyserammen knytter an til spørgsmålet om, hvilke resultater og anbefalinger evalueringen leverede samt formidlingen heraf.

Forskningsdesign og empirisk grundlag

Artiklen tager afsæt i et dybdegående teorifortolkende casestudie af EVANAV. En case der strækker sig over en 15 års periode: Fra de første overvejelser blev taget tilbage i 2001 til evalueringens formelle afslutning i sommeren 2014. EVANAV som case kan bidrage til at give et indblik i de udford-

ringer, der kan være forbundet med evalueringsanvendelse i forbindelse med større forvaltningsreformer. I konklusionen diskuteres resultaternes generaliseringspotentiale nærmere. Den teoretiske analyseramme angivet i figur 1 anvendes som fortolkningsramme med henblik på at belyse udfordringerne samt få en mere nuanceret og dybdegående forståelse heraf. Formålet er derfor ikke at teste de teoretiske positioners forklaringskraft. Derimod er der lagt op til en dybdegående analyse af den faktiske evaluering og hændelsesforløbet op til og omkring evalueringen. Artiklen baserer sig derfor på et stort empirisk materiale indsamlet i forbindelse med en meta-evaluering af EVANAV gennemført og afsluttet i foråret 2014 (Breidahl, Halvorsen, Furubo og Hansen 2014). Et bredt udsnit af dokumenter blev indsamlet, herunder evalueringsoplæg, rapporter, lovforslag, mødereferater etc. Med henblik på at afdække interessenteres synspunkter samt processen omkring evalueringens gennemførelse trækker de empiriske analyser endvidere på kvalitative interviews med 13 personer (indsamlet i foråret 2014), der repræsenterer en række af evalueringens centrale aktører og interessenter:

Den administrative ledelse i Arbejds- og Sosialdepartementet (Departementschefen), departementets (opdragsgivers) medlem i styringsgruppen, Kundskabsdirektøren fra AV-direktoratet, en repræsentant fra kommunernes interesseorganisation, repræsentanter fra Forskningsrådet og evalueringens styringsgruppe (formanden og en forskningsrepræsentant) samt fem forskere (evaluatorer).

Interviewmaterialet fokuserer på temaer som den politiske kontekst, evalueringens organisering, processen, konklusionerne, anvendelse etc. De giver et nuanceret, men samtidig også et subjektivt og retrospektivt indblik i hændelsesforløbet. Endvidere skal der tages forbehold for, at ikke alle interessenter er repræsenteret, herunder politikere fra stortinget, interesseorganisationer og fagforeninger. Det vil der så vidt blive taget højde for gennem analysen.

Interviewene er blevet transskriberet og efterfølgende kodet i NVivo. I forbindelse med evalueringens afslutning blev der afholdt et seminar, hvor resultaterne blev lagt frem til diskussion og nærmere drøftelse blandt flere af de centrale interessenter.

Endelig baserer det empiriske materiale sig på en systematisk analyse af publikationerne udgivet i regi af EVANAV frem til slutningen af maj 2014 (i alt ca. 87 publikationer). De er blevet læst og kodet med henblik på at afdække en række aspekter, herunder bl.a. deres fokus, målgruppe, anvendte metoder og resultater.

ANALYSE

De politiske og administrative vilkår for læring og vidensanvendelse

Nav-reformen blev vedtaget af Det norske storting i 2005 (St.prp. nr. 46, 2004–2005), men de første initiativer kan spores tilbage til 2001, hvor Stortinget besluttede at udrede spørgsmålet om en reformering af velfærdsadministrationen. Der var imidlertid ikke opbakning til den stortingsmelding, som regeringen efterfølgende leverede (Stortingsmelding nr. 14, 2002–2003). Heller ikke det forslag et efterfølgende ekspertudvalg leverede i 2004 (Rattsøudvalget) fik politisk opbakning. Det gjorde derimod et forslag fra en senere regering, som kom til at lægge grundstenen for vedtagelsen af Nav-reformen: Et forslag der hverken baserede sig på forslagene fra den tidligere Stortingsmelding eller ekspertudvalget. Nav-reformen opnåede bred tilslutning i Stortinget. Blot Fremskridtspartiet var fortalere for en alternativ model (Andreassen og Aars 2015:40–42).

Som nævnt blev det i forbindelse med vedtagelsen af reformen i stortinget besluttet, at der parallelt med reformen skulle gennemføres en evaluering. Hensigtserklæringer (som de kom til udtryk i en stortingsmelding) var forholdsvis løst formuleret. Dog blev der lagt op til, at evalueringen skulle være forskningsbaseret og organiseres eksternt i regi af Forskningsrådet (som led i en samlet FoU-satsning). Endvidere blev følgende hensigtserklæringer oplyst (St.prp. nr. 46 2004–2005:79):

For å sikre kontinuerlige impulser og kritiske blikk utenfra under reformgjennomføringen – og for å trekke lærdom av erfaringer underveis – vil det utvikles en samlet FoU-satsing hvor bl.a. følgende elementer er aktuelle:

- Systematiske utviklingsprosjekter der en gjennom ekstra innsats stimulerer til utprøving av metoder og arbeidsmåter
- Et systematisk opplegg for erfaringsinnhenting-, analyse- og spredning
- En idébank som legger til rette for direkte kontakt og læring mellom lokale instanser
- Fagutvikling knyttet til utvikling av en ny førstelinje, med bistand fra fagmiljøer
- Et helhetlig evalueringsopplegg som ikke bare undersøker effekter, men som også fotfølger reformen underveis, og som makter å løfte fram eksempler på vellykkede lokale løsninger.

Der var, som sagt, bred politisk opbakning til Nav-reformen og der er ikke noget der tyder på, at EVANAV i denne indledende fase var genstand for politiske magtkampe. Det er en konklusion, der understøttes af indsigter fra interviewmaterialet. Således blandede politikerne i Stortinget sig ikke nævneværdigt i spørgsmålet om evalueringen og dens nærmere udformning. Det blev, ifølge den daværende departementschef i Arbejds- og Socialdepartementet, overladt til embedsværket. Endvidere peger hun på, at der under de politiske forhandlinger ikke blev sat spørgsmålstegn ved, om reformen

skulle evalueres eller ej. Det blev anset som en selvfølgelighed, da reformen var så omfattende som tilfældet var.

Ovenstående giver derfor både et indtryk af fordelagtige og mindre fordelagtige betingelser for instrumentel evalueringsanvendelse, jf. den teoretiske analyseramme. Der herskede ikke de store politiske uenigheder om reformen eller om udformningen af evalueringen og det var ikke en proces, der var omgærdet af udprægede magtpolitiske eller taktiske hensyn. Her skal der tages forbehold for, at medlemmer fra Stortinget ikke er repræsenteret i interviewmaterialet. En afdækning af hændelsesforløbet understøtter dog denne tolkning (Reegård 2009).

Det er dog mindre fordelagtigt for betingelserne for instrumentel evalueringsanvendelse, at der i udgangspunktet ikke var meget fokus på at få præciseret evalueringens målsætning(er). F.eks. blev det ikke drøftet nærmere, om en ekstern forskningsbaseret evaluering placeret i regi af Forskningsrådet var det bedste middel til at belyse evalueringens målsætninger. Det blev i Stortingsmeldingen anset som en selvfølge, at evalueringen skulle organiseres på den måde. Hvis der havde været tale om en lineær, rationel proces kunne man have forventet, at selve organiseringen var blevet drøftet nærmere, herunder om en forskningsbaseret evaluering var mest hensigtsmæssig til at levere praktiske handleanvisninger sammenlignet med andre måder at organisere en evaluering på (f.eks. internt).

I forbindelse med udarbejdelsen af det såkaldte Mål- og rammedokument i vinteren og foråret 2006 blev evalueringens målsætninger konkretiseret. I den forbindelse peger flere af interviewene på, at det var tydeligt, at der fra starten af herskede divergerende opfattelser af og forventninger til EVANAVs output og anvendelsespotentialer. I denne fase var det departementet (embedsværket) som stod centralt, da dokumentet blev udarbejdet af en arbejdsgruppe med deltagelse af departementet, Forskningsrådet, flere forskere og repræsentanter fra det senere AV-direktorat, og dokumentet blev sendt ud i en høringsrunde til de berørte instanser, herunder KS. Som følge heraf blev fire resultatområder formuleret, som nærmere definerede formålet med evalueringen og den tiltænkte anvendelse.

Figur 2: Evalueringens fire resultatområder	
Dokumentation	Evalueringen skal give en systematisk oversigt over, hvad der er gjort og gøres i planlægnings- og implementeringsfasen
Korrektion	Evalueringen bør give svar på, hvordan implementeringsproblemer opstår, hvordan indsigter skabes, og spredes som ny kundskab, der giver anledning til justeringer og videreudvikling.
Indsigt om organisatorisk integration	Reformen forudsætter integration af forskellige organisationer og evalueringen bør give indsigt i modeller for vellykket organisatorisk integration.
Langsigtet kundskabsudvikling	Evalueringens resultater bør integreres i langsigtet samfundsvidenskabelig kundskabsudvikling

Kilde: Arbejds- og inkluderingsdepartementet 2006:3–4.

Som det fremgår af figur 2 blev det i dokumentet specificeret, at evalueringen både havde til formål at dokumentere og kontrollere, bidrage med instrumentel læring og korrektion samt bidrage til langsigtet kundskabsudvikling. De fire formål blev imidlertid ikke rangordnet eller prioriteret nærmere i dette mål- og rammedokument eller efterfølgende. Flere og – delvist modstridende – målsætninger var derfor i spil på et tidligt tidspunkt, hvilket viste sig at få afgørende betydning for den videre proces.

Det afspejler sig i interviewmaterialet, da der blandt de involverede aktører var forskellige forventninger til evalueringen, og det er i lyset af den manglende rangordning af målsætningerne, at de efterfølgende vurderinger af evalueringens faktiske anvendelse skal ses. For selvom der blandt de involverede interessenter var udbredt enighed om, at evalueringen ikke havde bidraget til instrumentel læring, herunder givet anledning til justeringer og videreudvikling af reformens implementering, så var der langt fra enighed om, hvorvidt det var problematisk eller ej (Halvorsen mfl. 2016). Det er også tilfældet internt i departementet: Mens departementets repræsentant i styringsgruppen i udgangspunktet havde en klar forventning om, at evalueringen kunne bidrage med løbende input til korrigerende af reformens implementering, anså departementschefen fra samme departement det hverken som realistisk og ønskværdigt. Således peger den administrative ledelse i departementet på, at det reelt var den målsætningen om langsigtet kundskabsudvikling, der blev vægtet højest. Endvidere udtaler hun, at de mere instrumentelle hensigter, som tydeligt kom til udtryk i Stortingsmeldingen, bundede i et hensyn til det politiske niveau, hvor det var en vigtig målsætning overfor omverdenen, at legitimere beslutningen om at anvende et forholdsvist stort beløb på en evaluering. Endvidere betoner hun, hvordan evalueringer har potentiale til at levere ét ud af mange vigtige bidrag til den politiske proces. I stedet vælger hun at fremhæve det positive i, at evalueringen har bidraget til demokratisk debat og samfundsoplysning og – ikke mindst – langsigtet kundskabsudvikling. Formanden for styringsgruppen og flere af de forskere vi interviewede erklærer sig enige i, at evalueringen primært har bidraget til demokratisk debat og samfundsoplysning og de understreger også vigtigheden af at den egentlige anvendelse har fungeret meget mere bredt, da resultaterne løbende er blevet formidlet på mange andre måder, bl.a. via afholdelse af oplæg etc. Også vores egen empiriske materiale vidner om, at en stor mængde af resultater blev formidlet og publiceret (over 87 publikationer) og forskerne deltog løbende i den offentlige debat, på konferencer, på forskersamlingerne, møder med brugerne etc.

Heroverfor stod kundskabsdirektøren fra AV-direktoratet og det daværende Arbejds- og inkluderingsdepartementets repræsentant i styringsgruppen, som i udgangspunktet havde en anden forventning til evalueringens output og anvendelsespotentiale. De er enige i, at evalueringen i høj grad har bidraget til debat og samfundsoplysning (hvilket forhåbentligt har kunnet bidrage til langsigtet kundskabsudvikling). De havde imidlertid også en forventning om at den skulle benyttes instrumentelt i konkrete handlingssituationer til at korrigere implementeringen af reformen samt opera-

tivt i forbindelse med den løbende styring. De betoner derfor det nedslående og problematiske i, at evalueringen ikke bidrog med input til implementeringsprocessen og ikke havde potentiale til at anvise konkrete anvisninger til justeringer. Departementets repræsentant i evalueringens styringsgruppe påpeger dog, at departementet i udgangspunktet havde for store forventninger til evalueringen.

EVANAV var derfor allerede i udgangspunktet præget af uklare målsætninger og divergerende forventninger, hvilket udfordrer det rationelle paradigmes tro på, at der lægger en veldefineret målsætning til grund for en given evaluering (Dahler-Larsen og Larsen 2001:212). Det rejser spørgsmålet om, hvorvidt de politiske og administrative beslutningstagere overhovedet havde et ønske om at opnå instrumentel læring? Knyttede evalueringens instrumentelle målsætninger op til formålsrationel ageren (effektivitet) eller en ambition om at udstråle legitimitet? Forløbet omkring EVANAV kan på den ene side karakteriseres ved at være formålsrationel, hvor den blev igangsat med henblik på at opnå et bestemt formål. Samtidig synes forløbet også at vidne om symbolik og et forsøg på at opnå legitimitet fra omverdenen. Sidstnævnte understøttes af departementschefens udtalelser om, at vedtagelsen af evalueringen blev betragtet som en selvfølge ligesom at dens nærmere udformning – at den skulle organiseres eksternt i regi af Forskningsrådet – ikke blev diskuteret nærmere fra politisk side. Noget tyder derfor på, at der i udgangspunktet ikke var politisk efterspørgsel efter viden til anvendelse i praktiske handlingssituationer og at evalueringens målsætninger gennemgående har været tvetydige. Derudover peger flere af interviewpersonerne på (forskerne såvel som repræsentanter fra departementet), at den politiske opmærksomhed på evalueringen forsvandt mere og mere gennem årene. Sidstnævnte skal ses i lyset af, at reformen og evalueringen heraf strakte sig over en lang tidsperiode (7 år).

Udtalelserne i interviewmaterialet vidner endvidere om, at det administrative modtagerapparat i forbindelse med EVANAV ikke har været gearret til at håndtere de mange informationer. Derfor var der en række konkrete forhold vedrørende EVANAV, som ikke er entydigt befordrerne for policylæring. Således er det et gennemgående tema i flere af interviewene, at departementet ikke havde tilstrækkelig med kapacitet til at forholde sig til den omfangsrige viden, som de mange evalueringspublikationer bidrog med. I starten af processen blev der etableret en læsegruppe i regi af departementet. Ambitionen var, ifølge medlemmet fra departementet i styringsgruppen, at man den vej igennem kunne få indsigter om processen. Men det viste sig, at dette forum var for uformelt og ikke sikrede en tilstrækkelig «absorbering» af viden, hvormed viden fra evalueringen ikke blev omsat og formidlet systematisk opad i systemet.

EVANAV var langt fra den eneste kilde til viden. Således blev der igangsat mange andre typer af udredninger og undersøgelser i forbindelse med Nav-reformen, som bidrog med løbende kundskab. Bl.a. blev der nedsat en ekspertgruppe, det såkaldte Hagen-udvalg, i februar 2010 som politisk svar på Rigsrevisionens kritik af Nav-reformen (se bl.a. Hagen, Sand, Kvåle, Borge

og Persson-Grivas 2010). Flere af interviewpersonerne, bl.a. medlemmet fra opdragsgiver (departementet) i styringsgruppen samt kundskabsdirektøren hos AV-direktoratet, fremfører, at resultaterne fra Hagen-udvalget i højere grad blev anvendt instrumentelt til at korrigere reformerne end tilfældet var med resultaterne fra Nav-evalueringen. Endvidere pointeres det, at formålet med nedsættelsen af dette udvalg netop var, at resultaterne skulle kunne anvendes instrumentelt til bl.a. at korrigere undervejs. Til sammenligning havde EVANAV et meget mere bredt opdrag og var derfor mindre gearret til at levere konkrete handleanvisninger. Endelig peger formanden for styringsgruppen på, at der løbende var opstået et akut vidensbehov, som man af gode grunde ikke kunne forudse, da man udarbejdede mål- og rammedokumentet. Det understreger pointen om, at beslutningsprocesser er præget af alt andet end veldefinerede mål og præferencer, hvilket udfordrer troen på instrumentel rationalitet og veldefinerede mål og præferencer.

EVALUERINGENS ORGANISERING OG RESULTATER

Som nævnt indledningsvist har dele af evalueringslitteraturen i en årrække forsket i, hvordan den evidensbaserede vidensanvendelse til brug i praktiske handlesituationer kan fremmes: Bl.a. har forskere peget på vigtigheden af at organisere og designe evalueringen i overensstemmelse med de overordnede formål (Alkin og Taut 2003; Johnson mfl. 2009; Leviton og Hughes 1981; Mark og Henry 2004; Patton 1997; Patton mfl. 1975). Som det er blevet demonstreret i analysen ovenfor er det nemmere sagt end gjort, når vi har at gøre med større forvaltningsreformer, der strækker sig over en lang tidsperiode og hvor præferencerne er uklare og omgivelserne foranderlige. Det blev, som nævnt, i udgangspunktet besluttet, at EVANAV skulle organiseres som en ekstern evaluering i regi af Forskningsrådet og som følge heraf fik eksterne evaluatore (i dette tilfælde forskere) til opgave at gennemføre evalueringen. Det daværende Arbejds- og inkluderingsdepartement havde det overordnede ansvar for evalueringen i kraft af sit sektoransvar, mens Norges forskningsråd havde ansvaret for evalueringens organisering og gennemførelse samt ansvaret for at varetage den faglige og administrative kontakt med forskningsmiljøerne. Det var bl.a. Forskningsrådets opgave at oprette en styringsgruppe, gennemføre udbudsrunder på baggrund af Mål- og rammedokumentet og sikre den forskningsmæssige og metodiske kvalitet (Samarbejdsaftale, 2006). Bruttonrammen for opdraget var 50 millioner kroner. Efter visse justeringer og suppleringer undervejs blev rammen til modulerne fastsat til ca. 45 millioner kroner, eksklusiv forskersamlinger og formidling. En række norske forskningsmiljøer bød ind på evalueringsopgaven og to konsortier, ledet af hhv. Arbejdsforskningsinstituttet (AFI) og Røkkansenteret, fik til opgave at udforme et oplæg for en samlet evaluering. Sidstnævnte blev den ansvarshavende institution for evalueringen og den blev organiseret omkring syv moduler.

Formelt blev der i samarbejdsaftalen, indgået mellem departementet og Forskningsrådet, lagt op til en stærkere styringsdialog mellem departementet og Forskningsrådet end, der ellers er kutyme for i forskningsprojekter finansieret af Forskningsrådet (Samarbejdsaftale 2006). Derfor fik EVANAV også et separat opdrag, der blev holdt ude af den ordinære programstruktur. Alligevel peger flere af interviewpersonerne på, at EVANAV ikke adskilte sig meget fra ordinære forskningsprojekter finansieret af Forskningsrådet, hvilket netop fik betydning for evalueringens manglende anvendelse i praktiske handlingssituationer. Således sammenkæder flere af interviewpersonerne, bl.a. departementschefen fra det daværende Arbejds- og socialdepartement, fraværet af praktisk anvendelse af evalueringen med den valgte organiseringsform for evalueringen. Også flere af forskerne peger i interviewene på, at de ikke opfattede evalueringen, som et opdrag fra arbejdsdepartementet, men som et projekt i regi af forskningsrådet. Og det havde uden tvivl betydning for, at outputtet fik den karakter som det gjorde. Således blev det forhold, at en ekstern evaluering i regi af Forskningsrådet lægger op til udstrakt forskningsfrihed, en udfordring for dens potentiale for at blive anvendt instrumentelt. Derimod var organiseringen med til at understøtte langsigtet kundskabsudvikling samt bidrage til demokratisk debat.

Den udstrakte forskningsfrihed afspejler sig også i de mange publikationer, der blev udgivet i regi af evalueringen (det sidste punkt i analyserammen). Selve resultaterne blev dokumenteret i ca. 87 publikationer, hvoraf størstedelen var rettet mod en norsk forskningsoffentlighed. Vores analyse af publikationernes indhold viser, at størstedelen af studierne fokuserede på implementeringen af reformen mens en mindre andel havde fokus på reformens målopfyldelse. Samlet set resulterede evalueringen i en påpegning af en række interessante generelle perspektiver, bl.a. at det for mange kommuner var alt for omfattende at integrere tre velfærdstjenester under et tag, at der var en række udfordringer forbundet med realiseringen af partnerskaber mellem det statslige og lokale niveau, at der opstod en række dilemmaer mellem at ansætte personer med generalist versus specialistkompetencer i Nav-kontorerne og at reformen næppe har fået flere i arbejde eller ført til øget brugerorientering og effektivisering. Blot for at nævne et lille udpluk af de mange forskelligartede resultater. Hertil kommer, at størstedelen af publikationerne var tilbageholdende med at levere anbefalinger og praktiske handlingsanvisninger. Det stemmer fint overens med interviewpersonernes påpegning af, at evalueringen kun i et begrænset omfang formåede at bidrage med input til implementeringsprocessen og forslag til justeringer. I stedet er evalueringens væsentligste hovedresultat et bidrag til langsigtet videnopbygning og til demokratisk debat både i kraft af de mange publikationer og den løbende formidling af evalueringens fund. Det kom bl.a. til udtryk ved mediedebat, oplæg for borgerne, brugere og beslutningstagere etc. (se Breidahl mfl. (2014) for en nærmere gennemgang og analyse af publikationerne).

Særligt kundskabsdirektøren i AV-direktoratet er kritisk overfor dette og pegede i interviewet på, at evalueringen i sin helhed i for høj grad var domineret af såkaldt 'universitetslignende' forskning og manglede indsigter om,

hvordan Nav-kontorerne bedre kan styres samt samspillet mellem Nav og det omgivende samfund, herunder arbejdsmarkedet. Med andre ord var den valgte organiseringsform med til at spænde ben for, at evalueringen løbende kunne give input og komme med tilbagemeldinger til arbejdet med implementeringen af Nav-reformen. Her kommer Hagen-udvalget ind i billedet igen, da den netop ikke byggede på «universitetslignende forskning» men gjorde brug af registerdata samt administrative data, hvorimod mange af evalueringspublikationerne i højere grad gjorde brug af kvalitative data. Derfor fik den, ifølge embedsmænd fra departementet samt kundskabsdirektøren i AV-direktoratet, en central rolle at spille i forhold til tilpasningen af reformen i implementeringsfasen. Et eksempel på, at denne evalueringsudstrakte forskningsfrihed på nogle fronter har været en udfordring for den rationelle paradigmes tro på instrumentel viden.

PERSPEKTIVERENDE KONKLUSION

Den stærke tro på at evalueringer kan anvendes til at effektivisere og forbedre samfundet er utallige gange blevet udfordret og det er en problemstilling, som evalueringslitteraturen har kredset om i en årrække. Formålet med denne artikel har været at bidrage til dette spørgsmål nærmere bestemt udfordringerne forbundet med instrumentel evalueringsanvendelse.

I modsætning til størstedelen af litteraturen, der har haft fokus på denne anvendelsesproblematik i forbindelse med evalueringer af afgrænsede programmer, har afsættet for denne artikel været evalueringsanvendelse i forbindelse med større forvaltningsreformer. Et empirisk genstandsfelt som et dybdegående casestudie af evalueringen af Nav-reformen (EVANAV) har givet os mulighed for at studere nærmere.

Den empiriske analyse vidnede om, at konteksten for policy-læring er af afgørende betydning, herunder det komplekse samspil mellem vidensanvendelse på den ene side og de politiske samt administrative mulighedsbetingelser på den anden. Fraværet af instrumentel anvendelse knytter dermed an til nogle mere grundlæggende problemstillinger forbundet med politik samt politiske problemstillinger og kan ikke alene henføres til udformningen af den pågældende evaluering.

Selvom at konteksten for policy-læring i forbindelse med EVANAV på nogle fronter var fordelagtig (bl.a. var reformen og evalueringen ikke præget af nævneværdige politiske magtkampe og på forhånd definerede interesser), så var den også præget af ustabile præferencer og foranderlige omgivelser. Det satte sit spor gennem hele processen, da det fra starten var uklart, hvad der var formålet med evalueringen. Idealet om instrumentel rationalitet hviler på en antagelse om lineære processer på problemløsning – man vælger den bedste mulige løsning til at opnå et givent mål. EVANAV vidnede imidlertid om, at præferencer er omskiftelige og at det er svært at spå om fremtiden. Således lagde man sig fast på en bestemt model for organisationen – midlet – inden at målet var fastlagt. Hertil kom, at der igennem hele processen herskede flere og divergerende forventninger til, hvad evalue-

ringen skulle anvendes til. Idealet om instrumentel policy-læring blev derfor allerede i udgangspunktet udfordret: Såvel internt i departementet som i det centrale Mål- og rammedokument. Processen omkring evalueringen af Nav-reformen vidner således om, at der ikke var tale om en opdragsgiver med et veldefineret mål, der valgte det mest hensigtsmæssige middel til at forfølge dette mål. At det er tilfældet er måske ikke så overraskende taget Nav-reformens kompleksitet, foranderlighed, et langt tidsperspektiv og ustabile omgivelser i betragtning. Det er resultater, der understøtter de mere kritiske dele af evalueringsanvendelseslitteraturen.

Den empiriske analyse rejste også tvivl om, hvorvidt departements modtagerapparat var gearet til at håndtere og omsætte resultaterne fra evalueringen. Således blev det med tiden mere eller mindre tilfældigt, hvem der læste evalueringens resultater. Der blev nedsat en uformel læsegruppe internt i departementet, men interessen for at læse resultaterne forsvandt gennem årene og resultaterne blev ikke omsat og bearbejdet. I stedet blev der anvendt andre videnskilder til at korrigere implementeringsprocessen.

Endelig synes også valget af Forskningsrådet som tovholder at være en vigtig kilde til forståelse af, hvorfor EVANAV i højere grad bidrog til langsigtet kundskabsudvikling end til korrigerende af implementeringen. Således understøttede den valgte organiseringsform, herunder armslængde princippet og den hertil hørende udstrakte forskningsfrihed, målsætningen om langsigtet kundskabsudvikling.

Resultaterne giver en indsigt i nogle af de særlige udfordringer, der er på spil, når vi har at gøre med større forvaltningsreformer præget af kompleksitet og foranderlighed, og kan dermed ikke generaliseres til udfordringerne forbundet med instrumentel evalueringsanvendelse mere generelt. Evalueringer af større forvaltningsreformer kan imidlertid også udformes og gennemføres meget forskelligt og noget tyder på, at udfordringerne måske var særligt udtalte i forbindelse med EVANAV, bl.a. fordi der i udgangspunktet var flere – delvist modstridende – målsætninger i spil og fordi der var tale om en evaluering organiseret eksternt i regi af forskningsrådet. Spørgsmålet er derfor, om en anden type af organisering (f.eks. hvor de sektorielle organer er en del af det permanente evalueringssystem (som tilfældet er i Sverige)) i højere grad kunne have bidraget til instrumentel policy-læring? Samtidig viser en komparativ analyse, at den instrumentelle brug af evalueringer i Danmark i forbindelse med evalueringen af kommunalreformen (også en større forvaltningsreform) havde bedre vilkår end i forbindelse med EVANAV. Måske fordi der var tale om en intern evaluering, hvor centrale interessenter fik evalueringsansvar og dermed en platform til at udøve interessevaretagelse. Dermed var den politiske interesse for evalueringen gennemgående tilstede, da der var politisk efterspørgsel efter viden til anvendelse i praktiske handlingssituationer. Samtidig var der også stor politisk uenighed om, hvorvidt reformen skulle evalueres (Breidahl mfl. 2016). Endnu et eksempel på, at konteksten for policy-læring er af afgørende betydning, herunder det komplekse samspil mellem vidensanvendelse på den ene side og de politiske samt administrative mulighedsbetingelser på den anden.

Som tidligere nævnt, blev EVANAV fra flere sider anerkendt for, at den bidrog til langsigtet kundskabsudvikling og til demokratisk debat i den bredere offentlighed. Målsætninger som ikke indgik som led i den danske evaluering af kommunalreform og det var da også meget begrænset, hvad den bidrog til på den front. Det rejser spørgsmålet om, hvorvidt vi har at gøre med en særlig oplysningsorienteret evalueringskultur i Norge.

LITTERATUR

- Albæk, E. (2001) Hvorfor nu al den evaluering? I P. Dahler-Larsen og H. K. Krogstrup (red.), *Tendenser i evaluering*, Odense: Syddansk Universitetsforlag.
- Alkin, M.C. og S.M. Taut (2003) Unbundling evaluation use, *Studies in Educational Evaluation*, 29:1–12. DOI:[https://doi.org/10.1016/S0191-491X\(03\)90001-0](https://doi.org/10.1016/S0191-491X(03)90001-0).
- Andreassen, T.A. og J. Aars (2015) *Den store reformen: Da NAV ble til*, Oslo: Universitetsforlaget.
- Arbeids- og inkluderingsdepartementet (2006) *Mål- og rammedokument for evaluering av NAV-reformen*, Oslo.
- Bennett, C.J. og M. Howlett (1992) The lessons of learning: Reconciling theories of policy learning and policy change, *Policy Sciences*, 25(3):275–294. DOI:10.1007/BF00138786.
- Breidahl, K.N. (2012) *Når staten lærer: En historisk og komparativ analyse af statslig policy læring og betydningen heraf for udviklingen i den arbejdsmarkedsrettede del af indvandrerpolitikken i Sverige, Norge og Danmark fra 1970 til 2011*, Ph.d. afhandling, Aalborg universitet.
- Breidahl, K.N., A. Halvorsen, J.-E. Furubo og M.B. Hansen (2014) *Opmærksomhedspunkter ved evaluering af politiske reformer: Rapport fra metaevalueringen af evalueringen af NAV-reformen i Norge*, Norges forskningsråd.
- Breidahl, K.N., G. Gjelstrup, H.F. Hansen, M.B. Hansen (2016) Evaluation of large-scale public sector reforms: A comparative analysis, *American Journal of Evaluation*, 1–20.
- Dahler-Larsen, P. og F. Larsen (2001) Anvendelse af evaluering: Historien om et begreb, der udvider sig, i P. Dahler-Larsen og H.K. Krogstrup (red.), *Tendenser i evaluering*, Odense: Syddansk Universitetsforlag: 211–220.
- Furubo, J.E. (1994) Learning from evaluations: The Swedish experience, i F. Leeuw, R. Rist, og R. Sonnichsen (red.), *Can governments learn? Comparative perspectives on evaluation and organizational learning*, New Brunswick, NJ: Transaction Publishers.
- Hagen, T.P., I.J. Sand, G. Kvåle, L.E. Borge og A. Persson-Grivas (2010) *NAV's organisasjon og virkemåte*. Delrapport 1 fra Ekspertgruppa som vurderer oppgave- og ansvarsdelingen i NAV, 30. april, og Tiltak for å bedre NAVs virkemåte. Sluttrapport fra ekspertgruppa, 24. juni.

- Hall, P.A. (1993) Policy paradigms, social learning and the state: the case of economic policy making in Britain, *Comparative Politics*, 25(3):275–96. DOI: 10.2307/422246.
- Halvorsen, A., K.N. Breidahl, J.-E. Furubo og M.B. Hansen (2016) Institusjonelle felt og forståelser av evaluering – en analyse av evalueringen av NAV-reformen, *Nordiske Organisasjonsstudier*, 2.
- Johnson, K., L.O. Greenesid, S.A. Toal, J.A. King, F. Lawrenz, B. Volkov (2009) Research on evaluation use: a review of the empirical literature from 1986 to 2005, *American Journal of Evaluation*, 30(3):377–410. DOI: 10.1177/1098214009341660.
- Leviton, L.C. og E.F.X. Hughes (1981) Research on the utilization of evaluations: A review and synthesis, *Evaluation Review*, 5(4):525–549. DOI:10.1177/0193841X8100500405.
- Mark, M.M. og G.T. Henry (2004) The mechanisms and outcomes of evaluation influence, *Evaluation*, 10(1):35–57. DOI: 10.1177/1356389004042326.
- Patton, M.Q. mfl. (1975) *In search of impact: An analysis of the utilization of federal health evaluation research: Minnesota center for social research*, University of Minnesota.
- Patton, M.Q. (1997) *Utilization-focused evaluation: The new century text*, Newbury Park, CA: Sage.
- Reegård, K. (2009) Historien om NAV-reformens unnnfangelse, i B.R. Nuland, B.S. Tranøy og J. Christensen (red.), *Hjernen er alene. Institusjonalisering, kvalitet og relevans i norsk velferdsforskning*. Oslo: Universitetsforlaget.
- Sabatier, P. og H. Jenkins-Smith (red.) (1993) *Policy change and learning: An Advocacy Coalition Approach*, Boulder: Westview Press.
- Samarbeidsavtale (2006) mellom Arbeids- og Inkluderingsdepartementet (AID) og Norges Forskningsråd (NFR) om forskningsbasert evaluering av NAV-reformen.
- Schwandt, T.A. (2015) Iagttagelser af videnskabelig evidens i politi og praksis, *Økonomi & Politik*, 88(1):5–12.
- Skocpol, T. (1985) Bringing the state back in: strategies of analysis in current research, i P. Evans, D. Rueschemeyer og T. Skocpol (red.), *Bringing the state back*, Cambridge: Cambridge University Press.
- St. meld. nr. 14 (2002–2003) *Samordning av Aetat, trygdeetaten og sosialtjenesten: Tilråding fra Sosialdepartementet av 13. desember 2002, godkjent i statsråd samme dag: Regjeringen Bondevik II*
- St.prp. nr. 46 (2004–2005) *Ny arbeids- og velferdsforvaltning*, Oslo: Arbeids- og Sosialdepartementet.