

Oppmerksomhetspunkter ved evaluering av politiske reformer

Rapport fra metaevalueringen av evalueringen
av NAV-reformen i Norge



Oppmerksomhetspunkter ved evaluering av politiske reformer

**Rapport fra metaevalueringen av evalueringen
av NAV-reformen i Norge**

Metaevalueringens ekspertpanel
juli 2014

© Norges forskningsråd 2014

Norges forskningsråd
Postboks 2700 St. Hanshaugen
0131 OSLO
Telefon: 22 03 70 00
Telefaks: 22 03 70 01
bibliotek@forskningsradet.no
www.forskningsradet.no/

Publikasjonen kan bestilles via internett:
www.forskningsradet.no/publikasjoner

eller grønt nummer telefaks: 800 83 001

Grafisk design omslag: Design et cetera
Foto/ill. omslagsside: Shutterstock
Trykk: 07 Gruppen/Forskningsrådet
Opplag: 200

Oslo, september 2014

ISBN 978-82-12-03357-3 (trykksak)
ISBN 978-82-12-03358-0 (pdf)

Sammendrag

Metaevalueringer vurderer nytten av en eller flere evalueringer og må ikke forveksles med metaanalyser (kapittel 1). Formålet med denne metaevalueringen har vært å lære av erfaringene fra evalueringen av NAV-reformen i Norge med henblikk på å utlede noen generelle prinsipper for organisering av evalueringer av store politiske reformer (kapittel 2).

Store politiske reformer er kjennetegnet ved at de berører mange mennesker, har ofte flere målsetninger som vanskelig lar seg forene, utvikler og forandrer seg over tid, og er irreversible i den forstand at de ikke kan føres tilbake til utgangspunktet (kapittel 2 og 5).

Etttersom evalueringer av politiske reformer typisk involverer paradokser, kompromisser og dilemmaer med kontekstavhengig betydning, har ekspertpanelet valgt å formulere analyseresultatene som åtte oppmerksomhetspunkter for beslutningstakere (se kapittel 5 og nedenfor).

NAV-reformen er en stor administrativ reform med en rekke arbeidsmarkeds- og velferdsmessige implikasjoner. Kapittel 3 gir en kort analyse av innholdet og utviklingen i disse implikasjonene.

Rapportens to hovedkapitler er kapittel 4 og 5. I kapittel 4 presenteres resultatene fra den empiriske analysen som metaevalueringen bygger på. Her presenteres og sammenfattes NAV-evalueringens interessenter sine vurderinger og oppfatninger av evalueringen i en rekke spesifikke spørsmål til den konkrete evalueringen. I kapittel 5 har ekspertpanelet, på grunnlag av de empiriske analysene og en redegjørelse for vår begrepsmessige tilnærming, utledet åtte generelle oppmerksomhetspunkter til bruk i framtidige evalueringer av politiske reformer. Disse presenteres nedenfor i overskriftsform. Se kapittel 5 for en utdypende beskrivelse.

Oppmerksomhetspunkt 1: Hva er fordelene og ulempene ved en forskningsbasert evaluering i regi av Forskningsrådet sammenlignet med alternative modeller?

Oppmerksomhetspunkt 2: Hvilke aktører bør trekkes inn ved formuleringen av evalueringsoppdraget?

Oppmerksomhetspunkt 3: Hvilke medlemmer bør inngå i styringsgruppen?

Oppmerksomhetspunkt 4: Hvilke prinsipper bør kjennetegne relasjonen mellom styringsgruppe og forskere?

Oppmerksomhetspunkt 5: Hvordan håndteres flere og vanskelig forenelige målsetninger for evalueringen?

Oppmerksomhetspunkt 6: Hvordan innarbeides fleksibilitet i evalueringsprosessen?

Oppmerksomhetspunkt 7: Hvor ambisiøs bør man være med tanke på å integrere forskjellige forskningsmoduler i en evaluering?

Oppmerksomhetspunkt 8: Bør mulighetene til å evaluere være styrende når reformer skal vedtas og implementeres?

Forord

I januar 2014 nedsatte Norges forskningsråd et uavhengig skandinavisk ekspertpanel til å utarbeide en metaevaluering av evalueringen av NAV-reformen. Formålet har vært å trekke fram vesentlige erfaringer til bruk ved organisering av framtidige evalueringer av politiske reformer.

Medlemmene av ekspertpanelet var: Morten Balle Hansen, professor, Aalborg Universitet, Danmark (leder); Jan-Eric Furubo, revisjonsråd ved Riksrevisjonen i Sverige (medlem) og Anne Halvorsen, instituttleder, Universitetet i Agder, Norge (medlem). Karen Nielsen Breidahl, adjunkt, ph.d., Aalborg Universitet, Danmark, har vært både forskningsmedarbeider, sekretær og medforfatter av rapporten. Forfatterskapet til rapporten deles likt mellom de fire.

Seniorrådgiver Steinar Kristiansen fra Norges forskningsråd har også deltatt som en svært aktiv observatør på samtlige møter, og har vært til uvurderlig hjelp under hele metaevalueringen.

Rapporten består av et kort sammendrag og fem kapitler. Kapittel 1 og 2 presenterer metaevalueringens problemstillinger, samt begrepsmessige og metodiske tilnærming. Kapittel 3 gir en kort beskrivelse av NAV-reformens utvikling, mens kapittel 4 presenterer resultatene av de empiriske analysene av utfordringer knyttet til evalueringen av NAV-reformen.

I kapittel 5 har vi har valgt å sammenfatte de vesentligste funnene fra våre undersøkelser og analyser som åtte generelle oppmerksomhetspunkter til bruk i tilretteleggingen av framtidige evalueringer av politiske reformer.

Rapporten kan forhåpentligvis bidra til å forbedre den norske stats allerede ganske velutviklede evner til å lære av sine erfaringer med politiske reformer.

God lesing!

Morten Balle Hansen
Leder for metaevalueringens ekspertpanel

Innholdsfortegnelse

Sammendrag	3
Forord.....	4
Innholdsfortegnelse	5
Kapittel 1. Introduksjon til metaevalueringen.....	7
Kapittel 2. Problemstillinger, begrepsmessig tilnærming og undersøkelsesdesign	9
2.1 Problemstillinger og begrepsmessig tilnærming	9
2.2 Metaevalueringens forskningsdesign og metoder	12
2.3 Oppsummerende beskrivelse av evalueringens formelle organisering	12
Kapittel 3. NAV-reformen i korte trekk.....	13
Generelt om reformen.....	13
2001	13
2002	13
2003	14
2004	14
2005	14
2006	15
2007	15
2008	15
2009	16
2010	16
2011	17
2013	17
Kapittel 4. Analyse av det empiriske materialet.....	18
4.1. Evalueringens overordnede organisering.....	18
4.2. Utlysning og oppstart av evalueringen	19
4.2.1. Utformingen av Mål- og rammedokumentet	20
4.2.2. Utlysning og utvelgelse	21
4.2.3. Utfallet av utlysningsfasen – den formelle arbeidsdelingen	22
4.3 Gjennomføringsfasen.....	23
4.3.1 Evalueringens drift og styring	23
4.3.2 Samspillet mellom prosess- og effektdelen.....	25
4.4. Resultat: Den muntlige og skriftlige formidlingen av evalueringen	26
4.5 Synspunkter på evalueringens anvendelsesproblemer	26
4.5.1 Samspillet med annen kunnskapsproduksjon	28

4.6	Evalueringsspublikasjonenes innhold	29
4.7	Oppsummering av sentrale synspunkter	33
Kapittel 5. Metaevalueringens oppmerksomhetspunkter		35
5.1.	Utgangspunkt for ekspertpanelets vurderinger	35
5.2	Oppmerksomhetspunkter.....	38
Litteratur og anvendte dokumenter		46
Bilag 1: Liste over intervjuobjekter		49
Bilag 2: Mandatet for metaevalueringen.....		50
Bilag 3: Mål- og rammedokumentet.....		54

Kapittel 1. Introduksjon til metaevalueringen

I 2005 vedtok det norske Stortinget en av de større politiske reformene i Norge i nyere tid – NAV-reformen. I Mål- og rammedokumentet beskrives NAV-reformen både som en betydelig reorganisering av forvaltningen og som en sosialpolitisk reform (Mål- og rammedokumentet, 2006: Bilag 3).

Samtidig ble det bestemt at det skulle gjennomføres en stor forskningsbasert evaluering av reformen – Evalueringen av NAV-reformen. Evalueringen har vært organisert av Norges forskningsråd og ble avsluttet i 2014 med en evalueringskonferanse den 12. juni og en bok med hovedresultatene fra evalueringen som vil bli utgitt i løpet av 2014.

I forbindelse med avslutningen av evalueringen av NAV-reformen besluttet Norges forskningsråd å nedsette et ekspertpanel med sikte på å gjennomføre en metaevaluering med utgangspunkt i spørsmålet: *Hva kan læres av evalueringen av NAV-reformen?* (Mandatet for Metaevalueringen, 2014: Bilag 2). Meningen med metaevalueringen er å komme med forslag til hvilken lærdom man kan trekke for fremtiden av evalueringen av NAV-reformen. Tanken er at framtidige forskningsbaserte evalueringer av store velferdsstatsreformer skal kunne dra nytte av de erfaringer som er gjort i forbindelse med evalueringen av NAV-reformen.

En metaevaluering er rett og slett en evaluering av en eller flere evalueringer, hvor nytten og verdien av en eller flere gitte evalueringer vurderes (Scriven, 1969)¹. Den må ikke forveksles med en metaanalyse, som er en form for kvantitativ syntese av mange studier med en felles problemstilling. Metaevalueringer kan ta mange forskjellige former, men er vanligvis av betydelig mindre omfang enn de evalueringene de vurderer (Stufflebeam, 2001), noe som også gjelder i vårt tilfelle.

Metaevalueringen som denne rapporten omhandler, har vært organisert som et ekspertpanel. Panelet har avholdt fem heldagsmøter i perioden fra 3. februar til 26. mai og et avsluttende dialogseminar 17. juni med sentrale interessenter i evalueringen, og har løpende kommentert utkast til delelementer i rapporten. Ekspertgruppens arbeid ble avsluttet primo juli 2014 med elektronisk oversendelse av den avsluttende rapport i pdf- og Word-versjon til Norges forskningsråd.

Som støtte i arbeidet har ekspertgruppen knyttet til seg et forskningsprosjekt ved Aalborg universitet som har generert data om evalueringens organisering, forløp og innhold. Datagenereringen har funnet sted i form av dokumentanalyser og intervjuer med sentrale aktører i evalueringen i perioden fra februar til mai 2014.

Ekspertpanelet har bestått av følgende medlemmer: Morten Balle Hansen, professor, Institutt for statskunnskap, Aalborg universitet, Danmark (leder); Jan-Eric Furubo, revisjonsråd ved Riksrevisjonen i Sverige (medlem) og Anne Halvorsen, instituttleder, Institutt for sosiologi og sosialt arbeid, Universitetet i Agder, Norge (medlem). I tillegg har seniorrådgiver Steinar Kristiansen fra Norges forskningsråd deltatt som observatør på samtlige møter og vært behjelpelig med avtaler og framskaffelse av relevant informasjon. Sist, men ikke minst, har adjunkt ph.d. Karen Nielsen Breidahl fra Institutt for

¹ For en mer omfattende definisjon og behandling av metaevalueringsbegrepet, se Stufflebeam, D. L. 2001. "The metaevaluation imperative." *American Journal of Evaluation* 22:183-209.

statskunnskap, Aalborg universitet, Danmark, i kraft av sin rolle som forskningsmedarbeider og sekretær for ekspertpanelet, deltatt i samtlige møter, stått for datagenereringen og analysene av de empiriske dataene, og skrevet store deler av rapporten.

Rapporten representerer forfatterens egne analyser og vurderinger og ikke institusjonene der de er ansatt.

Rapportens oppbygning er som følger: Kapittel 2 gir en kort presentasjon av metaevalueringens mandat og ekspertpanelets fortolkning av og empiriske tilnærming til oppgaven (kapittel 2). Deretter gis det en kort redegjørelse for selve NAV-reformens innhold og forløp (kapittel 3). Rapportens to hovedkapitler er kapittel 4 og 5. I kapittel 4 presenteres resultatene fra den empiriske analysen som metaevalueringen bygger på. Her er det NAV-evalueringens interessenters vurderinger av og synspunkter på evalueringen som presenteres. I kapittel 5 har ekspertpanelet, på grunnlag av de empiriske analysene og en redegjørelse for begrepsmessige tilnærming, utledet åtte generelle oppmerksomhetspunkter til bruk i framtidige evalueringer av politiske reformer. Det er tre vedlegg til rapporten: Vedlegg 1 består av en liste over intervjuobjekter, vedlegg 2 inneholder mandatet for metaevalueringen fra 2014, og vedlegg 3 inneholder Mål- og rammedokumentet for evalueringen av NAV-reformen fra 2006 utarbeidet av det daværende Arbeids- og inkluderingsdepartementet (Forskningsrådets oppdragsgiver).

Kapittel 2. Problemstillinger, begrepsmessig tilnærming og undersøkelsesdesign

Dette kapitlet gir først en presentasjon av metaevalueringens problemstillinger og begrepsmessige tilnærming, og fortsetter med en kort redegjørelse for det anvendte undersøkelsesdesignet.

2.1 Problemstillinger og begrepsmessig tilnærming

Metaevalueringens problemstillinger og begrepsmessige tilnærming tar utgangspunkt i det mandatet Norges forskningsråd formulerte for ekspertpanelets arbeid, samt i ekspertpanelets tillegg til og presiseringer av dette mandatet (se vedlegg 2). Som framhevet i innledningen, er formålet med metaevalueringen framtidsrettet læring. Dette er formulert som følger i mandatet:

"Evaluere prosesser som er gjennomført i planlegging, gjennomføring og avslutning av NAV-evalueringen. Metaevalueringen skal gi grunnlag for myndigheter/oppdragsgivere og forskere for å lære av prosessene" (Bilag 2: 1).

Det er således evalueringens organisering og prosesser som står i fokus for metaevalueringens analyser. Det har ikke vært metaevalueringens oppgave å vurdere den forskningsmessige kvaliteten på de gjennomførte analysene.

Formålet er ytterligere presisert i mandatets forventninger til rapporten:

"... Rapporten skal også inneholde forslag til prinsipper for god organisering av Forskningsrådets evaluering av politiske reformer" (Bilag 2: 2).

Som det framgår av kapittel 4 og 5, må prinsipper for god organisering av evalueringer imidlertid forholde seg til en rekke kompromisser og dilemmaer mellom forskjellige legitime hensyn. Ekspertpanelet har derfor valgt å formulere sine konklusjoner som en rekke generelle oppmerksomhetspunkter til bruk ved organiseringen av framtidige evalueringer av politiske reformer.

Ekspertpanelet har videre tolket mandatet for metaevalueringen som kombinasjonen av en måloppnåelseevaluering med blick for vesentlige sideeffekter og en interessentevaluering (Vedung, 1997). Ved en måloppnåelseevaluering vurderes en intervensjon ut fra de mål oppdragsgiveren har formulert. Utover dette kan det være vesentlige positive eller negative effekter som ikke har vært formulert som formelle mål. Ved en interessentevaluering vurderes en intervensjon ut fra forskjellige interessentgruppers normative kriterier.

I mandatet er interessentene begrenset til følgende sentrale aktører i evalueringen:

"oppdragsgiver/AID, Forskningsrådet, styringsgruppen, forskerne".

I tillegg til disse aktørene har ekspertpanelet valgt å inkludere Arbeids- og velferdsdirektoratet (AV-direktoratet) og Kommunenes interesseorganisasjon (KS) i interessentevalueringen.

Prinsipper for god organisering av evalueringer av politiske reformer vil, utover alminnelige kriterier for god forvaltningspraksis, avhenge av hva som er formålet med evalueringen. Ved metaevalueringen av NAV-reformens evaluering har ekspertpanelet tatt utgangspunkt i de mål for evalueringen som ble formulert i Mål- og rammedokumentet (Mål- og rammedokumentet, 2006: Bilag 3) da evalueringen ble lagt ut på anbud. Der framgår det at det norske Stortinget i 2004–2005 vedtok følgende om evalueringen av NAV-reformen:

"Et helhetlig evalueringssopplegg som ikke bare undersøker effekter, men som også fotfølger reformen underveis, og som makter å løfte fram eksempler på vellykkede lokale løsninger" (Bilag 3: 1).

På bakgrunn av dette formulerte det daværende Arbeids- og inkluderingsdepartementet i 2006 følgende overordnede formål med evalueringen:

"Det grunnleggende formål med denne evalueringen er å gi svar på om NAV-reformen når de overordnede målene; bedre arbeidsretting, bedre brukerretting og en mer effektiv forvaltning. Evalueringen omfatter både en forskningsbasert evaluering av resultater/effekter i forhold til målene og en prosessevaluering for bl.a. å dokumentere hvordan reformen er blitt gjennomført." (Bilag 3: 2).

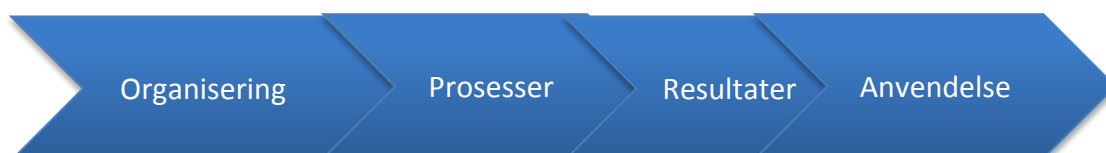
I forlengelsen av det overordnede formålet har Arbeids- og inkluderingsdepartementet formulert fire resultatområder (side 3–4): 1) *Dokumentasjon* – evalueringen skal gi en systematisk oversikt over hva som er gjort og gjøres i planleggings- og implementeringsfasen, 2) *Korreksjon* – evalueringen bør gi svar på hvordan implementeringsproblemer oppstår og hvordan innsikter skapes og spres som ny kunnskap som muliggjør justeringer og videreutvikling, 3) *Innsikt om organisatorisk integrasjon* – reformen forutsetter integrasjon av forskjellige organisasjoner, og evalueringen bør gi innsikt i modeller for vellykket organisatorisk integrasjon, og 4) *Langsiktig kunnskapsutvikling* – evalueringens resultater bør integreres i langsiktig samfunnsvitenskapelig kunnskapsutvikling.

De fire resultatområdene som metaevalueringen skal belyse, avspeiler de tre klassiske generelle formålene som er godt kjent fra evalueringslitteraturen: kontroll (punkt 1), forbedring (punkt 2–3) og langsiktig kunnskapsutvikling (punkt 4) (Vedung, 1997).

Det kan således være grunn til å tro at de erfaringer som kan utledes fra evalueringen av NAV-reformen, vil kunne danne grunnlag for formuleringen av noen meningsfulle generelle prinsipper for forskningsbaserte evalueringer av politiske reformer.

Noen mer spesifikke spørsmål kan også trekkes ut av mandatet for metaevalueringen. De springer ut av formålene som er beskrevet over, og kan deles inn i 1) spørsmål relatert til organiseringen av og prosesser relatert til evalueringen, og 2) spørsmål relatert til evalueringens innhold slik dette kommer til uttrykk i resultater og anvendelse. Figur 2.1 sammenfatter ekspertpanelets begrepsmessige tilnærming til metaevalueringens gjenstandsfelt.

Figur 2.1: Modell av metaevalueringens gjenstandsfelt: Evalueringen av NAV-reformen



Metaevalueringen omfatter således en beskrivende og vurderende analyse av NAV-evalueringens organisering, prosesser, resultater og anvendelse, med tanke på å utvikle velbegrunnede oppmerksomhetspunkter for forskningsbaserte evalueringer av politiske reformer.

Metaevalueringens mer detaljerte spørsmål springer ut fra mandatet for metaevalueringen, og kan sammenfattes med utgangspunkt i begrepene fra figur 2.1.

Evalueringens organisering er de strukturelle kjennetegn ved evalueringen som f.eks. styringsgruppens sammensetning og mandat, en samlet overordnet prosjektledelse, syv forskjellige forskningsmoduler og oppdragsgivers og Forskningsrådets avtalte roller i evalueringen.

Evalueringens prosesser er f.eks. forløpet av beslutnings- og kommunikasjonsprosesser i de forskjellige faser av evalueringen, herunder det konkrete samspillet mellom evalueringens forskjellige aktører. Betragtninger om prosesser og organisering reiser spørsmål som:

- *Hvordan bør styringsgruppen settes sammen?*
- *Hvilke aktører bør trekkes inn i formuleringen av evalueringsoppdraget?*
- *Hvordan bør samspillet mellom styringsgruppen og den forskningsbaserte evalueringen foregå?*
- *Har man funnet en god balanse mellom vitenskapelig styring og forskningsmessig frihet på den ene siden og politikernes/departementets legitime rolle før og under evalueringen på den andre?*
- *Hvilke forventninger har forskerne støtt på i prosessen, og hvordan har de håndtert dem?*
- *I hvilken grad har det forekommet forventninger som har skapt problemer i evalueringsprosessen?*

Evalueringens resultater er f.eks. de forskjellige publikasjoner og andre former for kommunikasjon av evalueringens resultater.

Evalueringens anvendelse er f.eks. den gjennomslagskraft evalueringen har hatt på NAV-reformens utvikling og den langsiktige kunnskapsutviklingen som er satt i gang. I tråd med det meste av evalueringslitteraturen (Dahler-Larsen & Krogstrup, 2003; Shulha & Cousins, 1997; Weiss, 1998) opererer man således med en bred forståelse av anvendelse som ikke bare omfatter instrumentell anvendelse, men også andre former for indirekte anvendelse, som f.eks. bakgrunnskunnskap, bidrag til den demokratiske debatt og generell langsiktig kunnskapsutvikling.

Betraktninger om resultat og anvendelse reiser spørsmål som:

- *Hvilke typer publikasjoner og andre former for resultater har evalueringen generert?*
- *Hvordan har evalueringens resultater blitt anvendt?*
- *Hvordan kommer evalueringens ulike formål til uttrykk i resultater og anvendelse?*

2.2 Metaevalueringens forskningsdesign og metoder

Metaevalueringens forskningsdesign og metoder må ses i lys av oppgaven som er gitt og ressursene som stilles til rådighet i form av tid og penger.

Som nevnt innledningsvis ble datagrunnlaget generert gjennom dokumentanalyser og intervjuer med sentrale aktører i evalueringen i perioden fra februar til mai 2014.

Med sikte på å avdekke interessentenes synspunkter og vurderinger av evalueringens organisering, prosesser, resultat og anvendelse er det gjennomført 12 kvalitative intervjuer med 13 personer som har vært sentrale aktører i evalueringen, og som representerer evalueringens sentrale interessenter: Arbeids- og sosialdepartementet, Arbeids- og velferdsdirektoratet (AV-direktoratet), Kommunenes interesseorganisasjon (KS), Forskningsrådet, styringsgruppen og en rekke utvalgte forskere (se listen over intervjuobjekter i vedlegg 1). Intervjuene ble transskribert og deretter kodet.

Med sikte på å få avdekket resultater slik det foreligger i publikasjonene, er det gjennomført en analyse av publikasjonene fra NAV-evalueringen som var utgitt fram til slutten av mai 2014 (i alt 87 publikasjoner). Disse er blitt lest og kodet med sikte på å avdekke innhold med relevans for metaevalueringen, bl.a. med fokus på publikasjonenes fokus, typer spørsmål og anvendte metoder.

Resultatene av de empiriske analysene som er gjennomført i forbindelse med metaevalueringen, er beskrevet i kapittel 4.

2.3 Oppsummerende beskrivelse av evalueringens formelle organisering

I rapporten vises det løpende til evalueringens formelle organisering. Hvordan den formelle organiseringen ser ut og prosessen fram mot endelig valg av organisering analyseres nærmere i kapittel 4. Det følgende oppsummerer kun de vesentligste trekkene.

Evalueringens formelle organisering

Oppstarts- og avslutningstidspunkt:

Selve evalueringsprosjektet ble formelt igangsatt 1. juli 2007, og 12. juni 2014 ble det holdt en avslutningskonferanse. Evalueringsoppdraget ble satt i gang 1. juni 2006.

Overordnet organisasjonsmodell:

Arbeids- og inkluderingsdepartementet (senere Arbeids- og sosialdepartementet) har hatt det overordnede ansvaret for evalueringen i kraft av sitt sektoransvar. Norges forskningsråd har hatt ansvaret for organisering og gjennomføring av evalueringen og for å ivareta den faglige og administrative kontakten med forskningsmiljøene som skulle utføre evalueringen.

Økonomisk omfang:

Bruttorammen for oppdraget var 50 millioner kroner. Etter visse justeringer og suppleringer underveis ble rammen for modulene fastsatt til ca. 45 millioner kroner, eksklusiv forskersamlinger og formidling.

Evaluatorene

To konsortier, ledet av hhv. Arbeidsforskningsinstituttet (AFI) og Rokkansenteret, fikk oppgaven med å utforme et opplegg for en samlet evaluering. Rokkansenteret har fungert som ansvarshavende institusjon, og prosjektlederen ble utpekt derfra.

Evalueringen har vært organisert omkring syv moduler, hver med en modulleder.

Kapittel 3. NAV-reformen i korte trekk

Kapitlet gjør kort rede for NAV-reformen i et tidsperspektiv, fra de tidlige diskusjonene om behovet for en ny organisering på 1970- og 1980-tallet fram til 2014.

Generelt om reformen

En samordning av en rekke offentlige tjenester har vært et viktig politisk tema under skiftende regjeringer siden 1970-årene, og på 1980-tallet ble det gjort forsøk med å få arbeids- og sosialkontor til å samarbeide, jf. frikommuneforsøkene. I starten knyttet diskusjonen seg primært til en samordning av helse-, sosial- og ytelsesdelene, men med tiden fikk målet om flere i arbeid og færre på sosiale ytelser større og større oppmerksomhet (Reegård, 2009; Andreassen & Aars, 2014).

Målene for en reform av velferdstjenestene ble presentert i St.meld. nr. 14 (2002–2003) *Samordning av Aetat, trygdeetaten og sosialtjenesten (SATS)*. De var a) å få flere i arbeid og i aktivitet og færre på overføringsinntekter (arbeidslinjen), b) forenkling for brukerne og bedre tilpassede tjenester (brukerretting), og c) en mer sammenhengende arbeids- og velferdsforvaltning (effektivisering). Ambisjonen bak sammenslåingen av tre tjenester – aetat, trygdeetaten og den kommunale sosialtjenesten – til én, var både at det skulle bli enklere for brukerne å forholde seg til hjelpeapparatet og at mulighetene for å tilpasse tjenestene til den enkeltes behov skulle bli større. Dette skulle bidra til at flere kom i arbeid framfor å være passive mottakere av overføringsinntekter. Reformen innebar at den såkalte arbeidslinjen ble understreket, dvs. at deltakelse i arbeidslivet ble sett på som et prinsipp for inkludering i samfunnet og sikring mot fattigdom. Den omfattet også endringer i en rekke tiltak og ordninger, og reformen beskrives derfor både som en *organisatorisk* reform og en *innholdsreform* (Andreassen & Fossetøl, 2011; Fimreite & Hagen, 2009, Mål- og rammedokument, 2006).

Resten av kapitlet gir en kronologisk beskrivelse av reformprosessen fram til 2014.

2001

I forbindelse med behandlingen av statsbudsjettet for 2002 vedtar Stortingets sosialkomite å be regjeringen om å utrede og komme med forslag til en felles etat for sosial-, helse- og arbeidsmarkedsetatene. Bondevik II-regjeringen (KrF, H og V) har overtatt makten etter valgseieren høsten 2001. Partene i arbeidslivet og regjeringen inngår intensjonsavtalen om inkludering i arbeidslivet, IA-avtalen. Målet er å redusere sykefraværet og hindre frafall og utstøtelse fra arbeidslivet.

2002

Regjeringen legger fram St.meld. nr. 14 (2002-2003) *Samordning av Aetat, trygdeetaten og sosialtjenesten* (SATS-meldingen) i desember. Den foreslåtte modellen innebærer én statlig etat for arbeids- og sosialrettede ytelser og én etat for pensjon og familieytelser. Videre foreslås et samarbeid mellom kommunene og staten om en felles førstelinje overfor personer som står i fare for å miste jobben eller har vansker med å skaffe seg jobb, men uten noen endring i arbeids- og oppgavefordelingen mellom kommune og stat.

Samme år blir det satt i gang treårige forsøk med samordning mellom Aetat, trygdeetaten og de kommunale sosialtjenestene, først i 13 og senere i 17 kommuner. Samordningsforsøkene varte til og med 2005, og i den påfølgende evalueringen av forsøkene ble de omtalt som en forberedelse til NAV-reformen (Møller, 2006).

2003

SATS-meldingen behandles i Stortinget i april. Flertallet (AP, SP, SV og Frp) oppfatter ikke det framlagte forslaget som godt nok når det gjelder forslaget om å slå sammen etatene til én felles etat. Meldingen blir derfor sendt tilbake til Regjeringen, jf. Innst.S. nr. 189 (2002–2003).

Rattsø-utvalget, et offentlig utvalg ledet av professor Jørn Rattsø, nedsettes i august med et mandat om å utrede "*ulike modeller for en felles velferdsetat bestående av dagens Aetat, trygdeetaten og sosialetaten*", jf. Innst.S.nr. 189 (2002–2003).

2004

Rattsø-utvalget legger fram sin rapport, NOU (2004:13) *Ny arbeids- og velferdsforvaltning* i juni. Rapporten presenterer fire modeller: a) én statlig etat for arbeid og én for pensjonsytelser, b) en statlig pensjonsetat og et kommunalt ansvar for arbeidsrettede tjenester, c) arbeid og pensjon samles i én statlig etat, og d) arbeid og pensjon blir et kommunalt ansvar. Utvalget anbefaler et statlig ansvar for å sikre arbeid og inntekt, men atskilt fra det statlige pensjonsansvaret, og en videreføring av ansvarsfordelingen mellom stat og kommune. Det anbefales at den statlige, arbeidsrettede førstelinjetjenesten blir samlokalisert med kommunenes sosialtjeneste. Innstillingen sendes ut på høring, og det er uenighet om hvordan den statlige delen bør organiseres, og til dels også om ansvarsfordelingen mellom kommune og stat (Hagen et al., 2010).

Fra oktober 2004 er ansvaret for arbeidsmarkeds-, trygde- og sosiale tjenester samlet i ett departement, Arbeids- og sosialdepartementet (ASD).

2005

I mars legger regjeringen (Bondevik II) fram et forslag om etablering av en ny Arbeids- og velferdsetat og organisering av førstelinjetjenesten i *partnerskap* mellom kommune og stat, dvs. NAV-kontor i kommunene, jf. St.prp. 46 (2004–2005) *Tilråding fra Arbeids- og sosialdepartementet om etablering av ny arbeids- og velferdsforvaltning*. Det skal etableres lokale samarbeidsavtaler med lovbestemte krav om at kommunene som et minimum overfører ansvaret for økonomisk sosialhjelp, rådgivning og individuelle planer til dette kontoret. Etter hvert overføres også boligsosialt arbeid og Kvalifiseringsprogrammet². Dermed får brukene *én dør* å forholde seg til. Etableringen av NAV-kontorene skal skje i løpet av perioden 2006–2009.

I mai behandler Stortinget proposisjonen, Innst.S. nr. 198 (2004–2005). Regjeringens forslag får tilslutning på alle punkter, og loven trer i kraft fra juli. Regjeringen foreslår etableringen av NAV-interim, Ot.prp. nr. 96 (2004–2005), som skal "planlegge og forberede gjennomføringen av statlige oppgaver innenfor en ny arbeids- og velferdsforvaltning". NAV-interim etableres i august. I skissen til de overordnede rammene for ansvars- og oppgavefordeling legges det til rette for organisering av spesialenheter i tillegg til lokale NAV-kontor, noe det også ble åpnet for i NAV-proposisjonen (Christensen, 2011, Andreassen & Aars, 2014, St.prp. 46 (2004–2005), pkt. 7.5.8).

I mai vedtar Stortinget prinsippene for en ny pensjonsordning (pensjonsforliket, NOU 2009:13). Det etableres et pensjonsprosjekt under NAV-Interim for å forberede

² Kvalifiseringsprogrammet: tilbud om opplæring, arbeidstrening og oppfølging for å komme i arbeid/aktivitet (nav.no)

pensjonsreformen (Christensen, 2011). I oktober får landet ny regjering (Stoltenberg II) bestående av AP, SP og SV. ASD blir avløst av Arbeids- og inkluderingsdepartementet (AID).

2006

Lovforslaget om en ny Arbeids- og velferdsforvaltning, Ot.prp. nr. 47 (2005–2006), legges fram i april og får tilslutning fra Stortinget. Loven trer formelt i kraft i juli. NAV-interim nedlegges. Det samme gjør Aetat og trygdeetaten, som samtidig slås sammen til Arbeids- og velferdsetaten (NAV³). AV-direktoratet overtar ansvaret for videre gjennomføring av den statlige delen av NAV-reformen. AID og KS inngår en rammeavtale om samarbeid om lokale NAV-kontor. I oktober etableres 25 piloter med lokale NAV-kontor basert på samarbeid mellom statlige og kommunale tjenester og med felles IKT-plattform. Målet er at alle kommuner skal ha etablert et NAV-kontor innen utgangen av 2009, men av praktiske grunner blir de siste 10 NAV-kontorene først etablert i 2010–11. Tilsammen blir det etablert 457 NAV-kontorer. Mange kommuner velger å legge flere oppgaver til NAV-kontorene enn det de er pålagt å gjøre.

I november legger regjeringen fram St.meld.nr. 9 (2006-2007) *Arbeid, velferd og inkludering* (innholdsreformen), med forslag til endringer i tiltak, ytelser og arbeidsmåter i NAV med sikte på å styrke inkluderingen i arbeidslivet.

2007

Kvalifiseringsprogrammet innføres⁴ fra november dette året, parallelt med etableringen av de lokale NAV-kontorene (Hagen et al., 2010). Kommunene har ansvaret for gjennomføring og drift av programmet, som omfatter både statlige og kommunale tiltak rettet mot brukere med vesentlig nedsatt arbeids- og inntektsevne som ikke mottar ytelser eller kun mottar begrensede ytelser. Folketrygdlovens nye regler for oppfølging av sykmeldte trer i kraft fra mars, jf. IA-avtalen og NAVs ansvar for oppfølging av sykmeldte. Disse reglene endres flere ganger, sist i mars 2014.

2008

Innholdsreformen, St.meld. nr. 9, behandles i Stortinget i mars. Hovedtyngden av de tilsammen 37 fylkesvise forvaltningsenhetene/avdelingene i forvaltningsenhetene etableres i løpet av første halvår (Hagen et al, 2010), og i desember opprettes fem spesialenheter for pensjon.

Kartlegginger gjennomført i 2006 og 2007 (Delta-prosjektene) danner grunnlaget for denne reorganiseringen, som innebærer en dreining mot større re-sentralisering og re-spesialisering av NAVs virksomhet. Oppgaver blir flyttet ut fra de lokale NAV-kontorene fordi det skal frigjøres ressurser til mer og bedre oppfølging av brukerne. Sluttbehandling og utbetaling av regelstyrte ytelser knyttet til familie og arbeid flyttes til forvaltningsenheter i fylkesleddet ("fylkeslinjen"), mens ansvaret for saksbehandling og utbetaling av pensjonsytelser flyttes til pensjonsenhetene, som sorterer direkte under AV-direktoratet (sammen med en del andre spesialenheter i "spesialenhetslinjen"). Brukerne har nå flere kanaler inn til NAV å forholde seg til, siden det ikke lenger er tilfelle at alle brukerhenvendelser går via de lokale NAV-kontorene (Hagen et al., 2010; Christensen, 2011; Christensen & Lægred, 2012).

³ En forkortelse for Ny arbeids- og velferdsforvaltning.

⁴ Med hjemmel i Lov om sosiale tjenester i NAV av 18. des. 2009 nr. 131.

I mars får AV-direktoratet et utvidet mandat som følge av at ansvaret for statlig oppfølging av økonomisk sosialhjelp og kvalifiseringsprogrammet overflyttes fra SHD⁵ (St.prp. nr. 51). På høsten starter arbeidet med å innføre nye metoder for oppfølging av brukerne, med behovs- og arbeidsevnevurdering og aktivitetsplaner (Hagen et al, 2010).

2009

En ny forskrift om arbeidsrettede tiltak, som ikke skiller mellom yrkeshemmede og andre arbeidssøkere, innføres i januar, og arbeidsrettet rehabilitering blir et nytt ordinært tiltak (St.prp. nr. 51). En ny lov om sosiale tjenester i arbeids- og velferdsforvaltningen vedtas i desember. Loven innebærer at kommunens ansvar for boligsosialt arbeid legges til NAV-kontorene.

Riksrevisjonen legger fram en tilleggsrapport om arbeids- og velferdsetaten i sin årlige revisjon av departementets virksomhet⁶ i oktober. NAV blir kritisert for en rekke feil og mangler, som manglende utbetalinger, lengre saksbehandlingstid, feil i utbetalinger, mangelfull oppfølging av brukere, og mangler i kontroll- og styringssystemer og i selve gjennomføringen av reformen. Riksrevisjonens konklusjoner vekker oppsikt og blir gjenstand for mye medieoppmærksomhet.

2010

Arbeidsavklaringspenger avløser tidsbegrenset uførestøtte, rehabiliterings- og attføringspenger fra og med mars. En ny lovbestemmelse som innføres i Arbeids- og velferdsforvaltningsloven i februar gjør arbeidsevnevurdering og aktivitetsplan til rettigheter for brukerne.

I januar er Riksrevisjonens rapport til høring i Stortingets kontroll- og konstitusjonskomite, og komiteens innstilling debatteres i Stortinget i mars. Det blir nedsatt en ekspertgruppe (Hagen-utvalget)⁷ med sekretariat i AID. Gruppen får som mandat å undersøke om arbeidsdelingen og samarbeidet mellom de kommunevise NAV-kontorene og forvaltningsenhetene fungerer etter hensikten. Det blir også opprettet en bredt sammensatt referansegruppe der fagforeninger, arbeidsgiverorganisasjoner og brukerorganisasjoner deltar.

Ekspertgruppen framlegger sin første rapport⁸ i april. Rapporten beskriver blant annet den gjeldende organiseringen av NAV-kontorene og forvaltningsenhetene (inkl. Pensjonsenhetene), arbeidsprosesser, arbeidsdeling, personalressurser og -forflytninger, IKT mv., pluss resultater og erfaringer med arbeidsdelingen. I juni leverer ekspertgruppen sin sluttrapport⁹, der de kommer med sine vurderinger og anbefalinger. I rapporten diskuteres en strategi for å løse NAVs problemer, og ekspertgruppen gir sine vurderinger og anbefalinger.¹⁰

⁵ Sosial- og helsedirektoratet

⁶ Riksrevisjonens rapport om revisjon av Arbeids- og velferdsetaten for budsjettåret 2008, tillegg 2 til Dokument 1 (2009-2010)

⁷ Prof. Terje P. Hagen, Bærum, UiO, Prof. Inger Johanne Sand, UiO, Førsteam. Gro Kvåle, UiA, Prof. Lars Erik Borge, NTNU og Ann Persson-Grivas, Forsäkringskassan, Stockholm.

⁸ *Delrapport 1 om NAVs organisering og virkemåte*

⁹ *Tiltak for å bedre NAVs virkemåte. Sluttrapport*

¹⁰ 1) Oppgavefokusering – rendyrke innsatsen på NAV-kontorene mot arbeid og aktivitet, 2) Effektivisering – forbedre ressursutnyttelsen og ressursfordelingen i arbeids- og velferdsetaten og i NAV-kontorene, 3) Diversifisering – kanalisere brukerne inn til NAV gjennom forskjellige kanaler alt etter deres behov, 4) Arbeidsdeling og forenkling – forbedre samarbeidet og saksflyten mellom NAV-kontorene og

2011

Som nevnt, kommer det siste lokale NAV-kontoret på plass dette året.

2013

I 2013 er NAV organisert i to styringslinjer; ytelseslinjen, som omfatter forskjellige spesialenheter og pensjons- og forvaltningsenhetene, og tjenestelinjen med bl.a. fylkeslederne og de kommunale NAV-kontorene, inkludert de landsomfattende kontaktsentrene (se organisasjonskartet på nav.no). Man er altså til dels tilbake til en organisering som har likhetstrekk med den tidligere organiseringen og tidlige reformforslag (Andreassen & Aars, 2014: Kap. 2).

Oppsummerende tidslinje:

2001

Stortinget ber regjeringen om å utrede en felles etat for sosial-, helse- og arbeidsmarkedstjenester.

2003

Stortingsmelding med forslag om en samordning av aetat, trygdeetaten og sosiale tjenester får ikke tilslutning i Stortinget, og sendes tilbake til regjeringen.

2005

Et nytt forslag til nyorganisering av arbeids- og velferdstjenester framlegges for Stortinget og vedtas. Tjenestene skal samles og brukene skal møte "én dør".

2006

De første pilotforsøkene med NAV-kontor i kommunene etableres.

2008

Fylkesvise forvaltningsenheter og fem nasjonale pensjonsenheter (spesialenheter) etableres, noe som innebærer en re-spesialisering og -sentralisering av NAV.

2009

Det er etablert NAV-kontor i nesten alle landets kommuner. Riksrevisjonen framlegger en rapport der NAV kritiseres for en rekke forhold.

2010

Riksrevisjonens rapport skaper debatt, og en ekspertgruppe blir nedsatt for å gjennomgå arbeidsdelingen i NAV-systemet.

2011

Etableringen av lokale NAV-kontor er gjennomført.

2013

Ny organisering med skille mellom en tjenestelinje og en ytelseslinje.

Kapittel 4. Analyse av det empiriske materialet

I dette kapitlet analyseres det innsamlede empiriske materialet, jf. kapittel 2, som metaevalueringens oppmerksomhetspunkter og tilhørende diskusjon tar utgangspunkt i.

Kapitlet er inndelt i syv avsnitt. Først presenteres evalueringens overordnede organisering, som ble besluttet så langt tilbake som i 2006 (avsnitt 4.1). Deretter vendes blikket mot oppstarten og utlysningen av evalueringen, før det gis en beskrivelse av den endelige rollefordelingen ved evalueringens oppstart våren 2007 (avsnitt 4.2). I avsnitt 4.3 analyseres gjennomføringsfasen, inkludert spørsmålene om drift og styring. Så følger en analyse av evalueringens publiseringer, formidling, mediedebatt osv. (avsnitt 4.4). I avsnitt 4.5 drøftes forskjellige synspunkter på evalueringenes senere og potensielle anvendelse, og i avsnitt 4.6 gis et innblikk i innholdet i de publikasjonene som har blitt utgitt i regi av evalueringen opp gjennom årene.

De sentrale synspunktene som har kommet fram gjennom intervjuene, blir framhevet underveis, og kapitlet avsluttes med en oppsummering (4.7) som danner utgangspunktet for den avsluttende diskusjonen av oppmerksomhetspunktene i kapittel 5. I forlengelsen av dette er det viktig å understreke at konklusjonene i dette kapitlet er trukket på bakgrunn av det empiriske materialet. Det er først i kapittel 5 at ekspertpanelet forholder seg selvstendig til utvalgte oppmerksomhetspunkter.

4.1. Evalueringens overordnede organisering

Stortingsproposisjon nr. 46 (2004–2005:79), der forslaget om NAV-reformen ble lagt fram, inneholdt også et forslag om at det skulle gjennomføres en evaluering av NAV-reformen parallelt med reformen. Denne evalueringen skulle ikke bare fokusere på reformens effekter, men også på implementeringsprosessen. Videre ble det lagt opp til at evalueringen løpende skulle komme med input og bidra med et kritisk blikk på gjennomføringen av reformen ut fra et ønske om å trekke inn erfaringer og bidra til løpende læring. Ifølge departementsråden var det på grunn av NAV-reformens størrelsesorden en selvfølge at den skulle evalueres.

Det ble på et tidlig tidspunkt bestemt at evalueringen skulle organiseres av Norges forskningsråd. Det daværende Arbeids- og inkluderingsdepartementet hadde det overordnede ansvaret for evalueringen i kraft av sitt sektoransvar, noe som innebar både finansieringsansvar og ansvar for utformingen av evalueringens innhold og innretning. Forskningsrådet hadde ansvaret for organiseringen og gjennomføringen av evalueringen, og for å ivareta den faglige og administrative kontakten med forskningsmiljøene som skulle utføre evalueringen. Nærmere bestemt var det bl.a. Forskningsrådets oppgave å opprette en styringsgruppe, gjennomføre anbudsrunder på bakgrunn av Mål- og rammedokumentet og sikre den forskningsmessige og metodiske kvaliteten. Samarbeidet mellom departementet og Forskningsrådet skilte seg fra de ordinære programstrukturene der departementet finansierer langsiktig forskning, noe som ble forklart med at evalueringen krevde en tettere styringsdialog mellom de to partene (Mål- og Rammedokumentet, 2006; Samarbeidsavtale, 2006).

Flere av intervjuobjektene peker i forlengelsen av dette på at evalueringen av NAV-reformen ikke har skilt seg nevneverdig fra ordinære forskningsprosjekter finansiert av Forskningsrådet. Formelt ble det likevel i samarbeidsavtalen mellom departementet og

Forskningsrådet lagt opp til en sterkere styringsdialog mellom departementet og Forskningsrådet enn det som ellers er vanlig i forskningsprosjekter finansiert av Forskningsrådet (Samarbeidsavtale, 2006). Derfor skulle evalueringen av NAV-reformen også ha et separat oppdrag som ble holdt utenfor den ordinære programstrukturen.

Valget av denne måten å organisere evalueringen på, særlig valget om å la evalueringen foregå i regi av Forskningsrådet, er noe mange av intervjuobjektene har synspunkter på. De fleste av intervjuobjektene synes at den valgte organiseringsformen, særlig valget av Forskningsrådet, var den mest hensiktsmessige, naturlige og legitime framgangsmåten. Dette synspunktet fremmes bl.a. av departementsråden, intervjuobjektet fra KS, forskermedlemmet i styringsgruppen, lederen av styringsgruppen og forskerne. Derfor pekes det heller ikke egentlig ut hensiktsmessige alternativer, og den gjennomgående forklaringen er at det neppe finnes alternativer for en evaluering av denne størrelsesorden. Særlig avvises det at det er realistisk å bruke konsulenter til å gjennomføre en så stor evaluering. I tillegg til størrelsen pekes det også på betydningen av at evaluatorene har spesiell (forsker-)kompetanse, noe som bl.a. framheves som viktig for å gjennomføre den kompliserte effektevalueringen. Endelig pekes det på betydningen av at en forskningsbasert evaluering kan bidra til å sikre langsiktig kunnskapsutvikling, noe som inngår som et av de fire resultatområdene i Mål- og rammedokumentet, jf. kapittel 2. Særlig framhever departementsråden i Arbeids- og sosialdepartementet dette synspunktet som helt sentralt.

Det framkom også noe kritikk mot den valgte organisasjonsformen på enkelte punkter. Bl.a. påpeker noen av representantene fra forskningssiden at man med fordel i større grad kunne ha supplert forskningsevalueringen med en mer anvendelsesorientert del (f.eks. utført av konsulentfirma), noe som kunne styrket ambisjonen om løpende og instrumentell anvendelse som også inngår som et av de fire resultatområdene, jf. kapittel 2. Den som er mest kritisk til den valgte organiseringsmåten er kunnskapsdirektøren i AV-direktoratet, som er grunnleggende kritisk innstilt fordi den valgte organiseringsformen bl.a. spenner ben for muligheten til at evalueringen kan gi løpende innspill og tilbakemeldinger til arbeidet med implementeringen av NAV-reformen innenfor AV-direktoratet. Dette er således en kritikk som i aller høyeste grad er knyttet til spørsmålet om evalueringsresultatene (instrumentelle) anvendelighet. Vi kommer tilbake til denne problemstillingen.

Som det vil framgå av analysen av det empiriske materialet under, har de forskjellige synspunktene på den valgte organiseringen av evalueringen betydning for hvordan det videre forløpet blir oppfattet av intervjuobjektene, inkludert oppstartsfasen, selve gjennomføringen av evalueringen, resultatene evalueringen har produsert, og ikke minst de problemstillingene som knytter seg til anvendelsen av evalueringens resultater.

4.2. Utlysning og oppstart av evalueringen

I mai 2006 gav Arbeids- og inkluderingsdepartementet Forskningsrådet i oppgave å evaluere NAV-reformen (brev av 19. mai med Mål- og rammedokument og forslag til samarbeidsavtale). Evalueringen av NAV-reformen ble utlyst i juni, altså før styringsgruppen ble opprettet, med søknadsfrist 31. august. Det ble satt en budsjettramme på ca. 50 millioner kroner over en periode på 6–8 år. Forskningsrådet godkjente oppdraget fra Arbeids- og inkluderingsdepartementet i august, og en styringsgruppe ble opprettet og møttes for første gang i september 2006. Forskningsrådet fikk inn fem søknader, som ble vurdert av et skandinavisk ekspertpanel (vanlig prosedyre hos

Forskningsrådet). Styringsgruppen behandlet søknadene i desember 2006. Resultatet ble at to konsortier, ledet av hhv. Arbeidsforskningsinstituttet og Rokkansenteret, fikk i oppgave å utforme et opplegg for en samlet evaluering. Rokkansenteret fikk prosjektlederansvaret. Denne fasen startet altså i juni 2006, og ble formelt avsluttet da den endelige prosjektbeskrivelsen var på plass i mars 2007.

Nedenfor behandles følgende temaer i forbindelse med utlysings- og oppstartsfasen: Utformingen av Mål- og rammedokumentet, utlysningen av evalueringen og utvelgelsen av søknader, og det endelige utfallet av denne første fasen.

4.2.1. Utformingen av Mål- og rammedokumentet

Mål- og rammedokumentet ble utformet i løpet av vinteren og våren 2006, og det ble lagt ned mye arbeid i dette dokumentet. Intervjuobjektene har mange synspunkter på prosessen omkring utformingen.

I forbindelse med utformingen av dokumentet ble det holdt flere planleggingsmøter der departementet, Forskningsrådet, flere forskere og representanter fra det senere AV-direktoratet deltok. Dokumentet ble også sendt ut på høring til de berørte instanser, inkludert KS. I forkant av dette ble Forskningsrådet trukket inn som følge av at Forskningsrådet var valgt som organisator. Ifølge departementets representant i styringsgruppen kom Forskningsrådet med innspill til utvelgelsen av forskerne, og som fungerte som konsulenter i arbeidet med dokumentet ved at de kom med muntlige og skriftlige innspill til evalueringens fokusområder og vinkling. Departementets representant understreker også at det var departementet som i siste instans utformet Mål- og rammedokumentet og som bestemte hva det skulle inneholde, mens forskerne primært hadde en rådgiverrolle og kom med utkast og synspunkter. Denne involveringen anses både som en fordel og en ulempe. Noen av intervjuobjektene oppfatter det som et mulig habilitetsproblem at noen av de involverte forskerne senere søkte på evalueringen og også fikk tildelt midler, siden de utpekte forskerne også var potensielle søkere. Ifølge departementets representant i styringsgruppen hadde dette valget den fordel at viktige innsikter ble trukket inn på et tidlig tidspunkt, og her må den norske konteksten tas i betraktning (forholdsvis små forskningsmiljøer osv.). Dessuten påpeker et av intervjuobjektene fra Forskningsrådet at man fra Forskningsrådets side var oppmerksom på at forskerne som ble involvert i planleggingsmøtene senere kunne bli potensielle søkere. Deres vurdering var likevel at dette ikke skulle diskvalifisere de nevnte forskerne. I så fall burde departementet ha opplyst forskerne om dette.

Disse synspunktene på hvem som har hatt og bør ha innflytelse på Mål- og rammedokumentet knytter videre an til det forhold at styringsgruppen, som først ble nedsatt etter at Mål- og rammedokumentet var formulert, ikke hadde noen innflytelse på dokumentets innhold og vinkling. Styringsgruppens leder er kritisk til dette, siden han anser Mål- og rammedokumentet som et sentralt arbeidsredskap for styringsgruppen. Bl.a. framhever han at det hadde vært mer hensiktsmessig å ha en større grad av dialog med oppdragsgiver om innholdet i mandatet (Mål- og rammedokumentet). Poenget er ikke at det er noe galt med Mål- og rammedokumentet knyttet til evalueringen av NAV-reformen, men at det kunne det i prinsippet ha vært – og i så fall ville det vært særlig viktig å involvere styringsgruppen.

Som det framgår av kapittel 2, inneholdt Mål- og rammedokumentet fire resultatområder som evalueringen skulle leve opp til: 1) Dokumentasjon, 2) Korreksjon, 3) Innsikt om

organisatorisk integrasjon, og 4) Langsiktig kunnskapsutvikling. Resultatområdene ble ikke rangert eller prioritert.

Intervjuobjektene uttrykker generelt ikke synspunkter på selve innholdet i Mål- og rammedokumentet, og de fleste gir uttrykk for at de ikke husker det spesifikke innholdet, herunder forventningene og kravene i dokumentet. Et av intervjuobjektene fra Forskningsrådet påpeker likevel dilemmaet i at det er snakk om en forskningsbasert evaluering, der resultatområdene og dokumentet som helhet ikke skal være for spesifikt og styrende. Det er ikke dermed sagt at dette innebærer noe egentlig problem, men det blir påpekt hvor viktig det er å være forsiktig i denne balansegangen. Særlig bør man, ifølge intervjuobjektet fra Forskningsrådet, være oppmerksom på om forskerne opplever dokumentet fra oppdragsgivers side som styrende. Kunnskapsdirektøren fra AV-direktoratet peker derimot på at departementet burde ha vært mer styrende og er misfornøyd med den valgte vinklingen av oppdraget. Vi kommer tilbake til denne problemstillingen.

4.2.2. Utlysning og utvelgelse

Etter at Mål- og rammedokumentet var ferdig, ble det utarbeidet en anbudstekst. Utlysningen av evalueringen (i juni 2006) og søknadsfristen i slutten av august 2006 betød at prosjektsøknadene måtte utarbeides i løpet av sommeren. Deretter nedsatte Forskningsrådet en skandinavisk ekspertgruppe for å evaluere søknadene og sette poeng (en fast "armlengde-rutine" i Forskningsrådet bruker for forskningsprosjekter). Ekspertgruppen kom med en innstilling til styringsgruppen. Forskningsrådet hadde ansvaret for utvelgelsen av evaluatorene.

Tidsperspektivet i søknadsrunden reiser en potensiell problemstilling om hvorvidt forskningsmiljøene hadde tilstrekkelig tid til å utforme søknader. Dessuten ble det lagt opp til at forskningsmiljøene skulle gå sammen om søknader som konsortier. Tidsperspektivet framheves dog ikke som problematisk i selve intervjuene, heller ikke fra forskernes side. Det er imidlertid viktig å merke seg at de intervjuede forskerne representerer dem som ble tildelt evalueringsoppgaven, og ikke dem som potensielt ikke rakk å søke. Vi kan derfor ikke uttale oss nærmere om hvorvidt den korte søknadsfristen var grunnen til at det bare var fem søkere til forskningsevalueringen.

Mange av intervjuobjektene, særlig forskerne, har synspunkter på konseptet der forskningsmiljøene skulle gå sammen og søke som samlede konsortier (i 2006 var dette et relativt nytt konsept). Særlig peker de intervjuede forskerne på problemene ved dette. Spesielt medførte det ulemper at styringsgruppen i vurderingsfasen ikke valgte et konsortium framfor et annet, men i stedet valgte å sette sammen et nytt. Bakgrunnen for dette var at styringsgruppen også valgte å se søknadene som en samlet helhet i forhold til oppdraget, dvs. avveie sammensetningen av evalueringstemaer og budsjettammen.

At det ble satt sammen nye konsortier betød at man måtte opparbeide nye samarbeidsrelasjoner og, i følge flere av de intervjuede forskerne, at det ble mindre penger til de enkelte søknadene. En del interne avklaringer var derfor påkrevet i evalueringens startfase. Bl.a. ble det innkalt til et forhandlingsmøte i januar 2007 for å utarbeide en revidert prosjektbeskrivelse og et revidert budsjett (Referat, SG møte 1/2007). Noen av de intervjuede forskerne peker i forlengelsen av dette på at denne prosessen fikk negative følger for den senere interne styringen og at samarbeidet mellom forskerne trolig ble svekket som følge av dette.

Internt i styringsgruppen var det ikke de store uenighetene om vurderingen av søknadene. Alle søknadene holdt høy kvalitet (fikk 4 eller 5 poeng av ekspertgruppen). I utvelgelsen ble det særlig lagt vekt på hvem som hadde kompetanse i å ivareta registerdata. To av styringsgruppens medlemmer var inhabile (Referat, SG møte 1/2007). Det framgår ikke nærmere av referatene fra styringsgruppemøtene hvorfor noen miljøer ble valgt framfor andre og hva som lå bak de mer detaljerte kriteriene for valget.

4.2.3. Utfallet av utlysingsfasen – den formelle arbeidsdelingen

I slutten av kapittel 2 ble det gitt en oversikt over evalueringens overordnede organisering og utforming. Dette avspeiler den endelige beslutningen, arbeidsdelingen og det faktum at styringsgruppen valgte å prioritere søknadene fra Rokkansenteret og Arbeidsforskningsinstituttet, og gav førstnevnte det samlede lederansvaret. Prosjektlederen ble utpekt fra dette forskningsmiljøet.

Etter den endelige utvelgelsen av søknader og de avsluttende forhandlingene var den formelle arbeidsdelingen på plass i mars 2007, og den medførte noen endringer i prosjektbeskrivelsen og budsjettet. Prosjektet fikk en samlet budsjetttramme på ca. 41 millioner kroner (mot 50 millioner i utlysningsteksten). Budsjettet ble redusert, bl.a. fordi det ble besluttet at formidlingsaktivitetene og prosjektsamling skulle holdes utenfor prosjektets budsjett og i stedet plasseres hos Forskningsrådet (Referat, SG møte 1/2007). Senere initiativer endte med at prosjektets samlede budsjett ble på ca. 45 millioner kroner. Det ble bestemt at det skulle opprettes syv moduler, hver med en modulleder.

Den formelle arbeidsdelingen ble at prosjektlederen skulle ha det overordnede faglige ansvaret for at evalueringen oppfylte sine målsetninger og for kontakten med styringsgruppen, mens modullederne hadde ansvar for framdrift og faglighet innenfor de enkelte modulene. Prosjektlederen¹¹ så det som sin oppgave å stå for kontakten mellom styringsgruppen og forskerne, og forstod således sin rolle mer som en koordinator enn en egentlig leder – den som koordinerte arbeidet mellom forskerne på den ene siden og Forskningsrådet og styringsgruppen på den andre. Dessuten fungerte hun som evalueringens ansikt utad.

Flere intervjuobjekter har påpekt at dette er en hensiktsmessig arbeidsdeling. Det ble likevel påpekt at denne arbeidsdelingen medfører en fare for båndtenkning og at det ble styrt for lite på tvers av modulene, men det ble også sagt at de regelmessige møtene og forskersamlingene kompenserte for dette problemet. Generelt hersker det stor tilfredshet med den måten prosjektledelsen fungerte på.

I Mål- og rammedokumentet ble det listet opp en rekke formaliteter omkring styringsgruppen, inkludert hvem den skulle bestå av og kriterier for utvelgelse, deres ansvar og rolle (Mål- og rammedokumentet, 2006: 13-14). Her går det fram at det var styringsgruppens oppgave å sikre forskernes integritet og evalueringens legitimitet, og styringsgruppens leder så det som sin oppgave å sikre prosessens integritet og evalueringprosessen (inkludert å få prosessen til å gå og å avslutte og rapportere til slutt), ikke å involvere seg i eller forholde seg til forskningens innhold. Således valgte han, etter eget utsagn, å ha en mening om evalueringprosessen og ikke om evalueringresultatene. Det hersker delte meninger om dette valget, noe som fremkommer som et sentralt

¹¹ I august 2013 ble prosjektlederen skiftet ut på grunn av karrierebytte.

synspunkt i intervjuene. Dette reiser spørsmål ved hvor sterkt engasjert styringsgruppen bør være i evalueringens innhold, herunder hvor mye de bør delta i forskningsseminarer, på hvilken måte osv. Engasjementet var begrenset, og dette reiser spørsmålet om hvilken rolle lederen for en styringsgruppe bør ha.

I tillegg til prosjektlederens kontakt med styringsgruppen, var det også kontakt mellom forskerne på den ene siden og styringsgruppen og departementet på den andre gjennom den årlige innsendingen av årsrapporter.

4.3 Gjennomføringsfasen

Selve evalueringsprosjektet ble formelt igangsatt 1. juli 2007, og 12. juni 2014 ble det holdt en avslutningskonferanse. Gjennomføringsfasen har således strukket seg over 7 år. I denne perioden ble det avholdt 2–3 styringsgruppemøter årlig der en rekke spørsmål har blitt drøftet løpende, herunder publikasjoner, formidling, nye tiltak etc. I perioden har det dessuten blitt holdt årlige forskersamlinger, brukerforum og andre typer formidlingstiltak, noe som framgår av årsrapportene. I det videre vil det rettes særlig fokus mot spørsmål knyttet til drift og styring.

Vi begynner med en kort drøfting av spørsmålet om tidspunktet for igangsetting av evalueringen. Ifølge prosjektlederne (både den første og den siste) ble evalueringen satt i gang på et noe tidlig tidspunkt og kunne med fordel vært utsatt. En utsettelse kunne ha styrket forskningsdesignet, inkludert samspillet mellom evalueringens prosessdel og effektdel, siden man på det tidspunktet evalueringen ble satt i gang visste for lite om iverksettelsesprosessen. I stedet kunne man, ifølge prosjektlederen, ha satt i gang en pilotevaluering.

4.3.1 Evalueringens drift og styring

Størstedelen av evalueringens drift har funnet sted i regi av styringsgruppemøtene, som har vært sentrum for drøfting av synspunkter på evalueringen og publikasjonene.

Ifølge intervjuobjektene har både evalueringens drift og styring ligget tett opp til et ordinært forskningsprogram i regi av Forskningsrådet. Dette har bl.a. kommet til uttrykk ved noen klare arbeidsdelinger, hvor Forskningsrådet har hatt ansvar for å sikre den forskningsmessige og metodiske kvaliteten (Samarbeidsavtale, 2006; 2–3).

De fleste intervjuobjektene gir uttrykk for at selve driften i gjennomføringsfasen har fungert svært godt, og det samme er tilfelle når det gjelder det konkrete samarbeidet internt i styringsgruppen og mellom styringsgruppen og forskerne. Flertallets vurdering er at dialogen mellom de involverte aktørene, herunder Forskningsrådet og departementet, har vært god og tillitsfull.

Det at NAV-evalueringen var lagt til Forskningsrådet bidro, ifølge flere av intervjuobjektene, til å understøtte en felles vilje til å oppnå noe. De fleste oppfatter Forskningsrådet som en meget velsmurt organisasjon som har gjort en god jobb med å understøtte prosessen, bl.a. fordi det har vært klare prosedyrer å forholde seg til. Det at det har vært fastsatt noen klare rammer på forhånd har lettet arbeidet for både lederen av styringsgruppen og prosjektlederen. Dessuten peker mange på at Forskningsrådets sekretariat fungerte godt, blant annet at det formelle fungerte uten problemer og at det fantes klare prosessregler. I og med forankringen i Forskningsrådet, var mange av prosedyrene og prosessene fastlagt på forhånd – rammene var gitt.

Det velfungerende samarbeidet blir også i høy grad tilskrevet en dyktig prosjektleder.

Medlemmene av styringsgruppen vurderer det også slik at samarbeidet har fungert godt. Bl.a. peker et av medlemmene i styringsgruppen på at den løpende informasjonen fra forskerne om tiltak og utvikling har vært god. Det trekkes likevel fram som problematisk fra flere hold at det har vært stor utskifting av medlemmer fra AV-direktoratet og KS, som derfor ikke har deltatt særlig mye i styringsgruppen. Ifølge et intervjuobjekt fra KS skyldes dette at KS' representanter i styringsgruppen (også de potensielle) er travle mennesker. Kunnskapsdirektøren fra AV-direktoratet peker på manglende interesse for og lyst til å engasjere seg i styringsgruppens arbeid, siden man raskt fant ut at man fra direktoratets side ikke kunne bruke resultatene fra NAV-evalueringen til særlig mye. Budskapet fra AV-direktoratet er generelt sett at man er langt fra begeistret for samarbeidet med styringsgruppen. Etter direktoratets mening var styringsgruppen i utgangspunktet feil sammensatt og preget av mangel på styring. Opplevelsen var dessuten at forskerne ikke lyttet til direktoratets innspill.

Forskerne er generelt tilfredse med driften og det konkrete samarbeidet, og prosjektlederen – både den nåværende og den forhenværende – framholder at de ikke har møtt de helt store utfordringene i forbindelse med prosjektlederrollen. Det at man har arrangert løpende forskersamlinger som alle de involverte forskerne ble invitert til, anses som positivt og som noe som har vært med å forebygge båstenkning innenfor hver av de sju modulene. Generelt var forskerne som ble intervjuet, svært positive til de årlige forskersamlingene, som også innbefattet besøk i utlandet, bl.a. i Tyskland (samlingene var dog mer en anledning til faglig samling enn til formidling). Styringsgruppen ble også invitert, noe som kunne bidra til en formidling av resultater. Fra forskningssiden har man vært skuffet over at bare noen få medlemmer av styringsgruppen valgte å delta.

Et viktig synspunkt som fremkommer av intervjumaterialet, gjelder styringsgruppens grad av styring. Det er en generell oppfatning, både blant styringsgruppens medlemmer og forskere, at styringsgruppen ikke har styrt særlig mye, men snarere fungert som en sparringspartner i en konsulterende og oppfølgende rolle. Det hersker delte meninger om hvorvidt dette er negativt eller positivt. Noen ser det som positivt, siden den manglende styringen kan ses som et uttrykk for at det ikke har oppstått nevneverdige problemer. Hvis mye hadde gått galt og forskerne ikke hadde levd opp til ansvaret sitt, hadde kanskje styringsgruppen hatt en annen rolle. Lederen av styringsgruppen påpeker at forskerne oppfylte sine forpliktelser, og derfor var det ikke behov for å styre dem særlig mye.

Noen av intervjuobjektene fra forskersiden gir likevel uttrykk for at de kunne ha ønsket seg mer styring fra styringsgruppens side. Videre framholdes det at forskningssiden nok har blitt gitt for stor innflytelse på dagsordenen for styringsgruppens møter. Ifølge Forskningsrådet ble både lederen av styringsgruppen og prosjektlederen trukket inn i arbeidet med å fastsette dagsorden.

En annen av forskerne ser det derimot slik at oppdragsgiver snarere har styrt styringsgruppemøtene for mye og på en uhensiktsmessig måte. Det er altså ikke enighet blant de intervjuede forskerne på dette punktet.

Et annet viktig synspunkt som kom fram på bakgrunn av intervjuene, er spørsmålet om oppdragsgivers posisjon i styringsgruppen: Skal oppdragsgiver være representert i styringsgruppen eller ikke? Departementets representant i styringsgruppen påpeker selv at det er en vanskelig rolle og at representasjonen var forbundet med både fordeler og

ulemper. Han understreker samtidig at han var seg sin rolle bevisst, herunder at det ikke på noen måte ble bedrevet aktiv styring i styringsgruppen. At både oppdragsgiver og evaluator er representert har den fordel at dialogen finner sted innenfor styringsgruppen heller enn utenfor. Det er samtidig viktig å påpeke at ingen av intervjuobjektene peker på konkrete problemstillinger som har oppstått i denne forbindelse. Det handler mer om en prinsipiell problemstilling. Ut fra mer pragmatiske hensyn framheves oppdragsgivers representasjon som meget hensiktsmessig, bl.a. av styringsgruppens leder og forskermedlemmet i styringsgruppen, blant annet fordi dette understøttet løpende dialog og god kontakt med departementet.

Forskningsfrihet?

Når det gjelder spørsmålet om forskernes frihet er blitt ivaretatt og opprettholdt, er det et gjennomgående synspunkt (også blant forskerne) at den har vært ivaretatt på beste måte, noe som blant annet tilskrives det faktum at evalueringen ikke har skilt seg nevneverdig fra ordinære forskningsprosjekter i regi av Forskningsrådet. Ikke overraskende framhever forskerne det som meget positivt at forskningsfriheten har hatt de beste vilkår.

Selv om de fleste av intervjuobjektene er enige om betydningen av å opprettholde forskningsfriheten, framsettes det også mer kritiske synspunkter. Departementets representant i styringsgruppen påpeker at andelen prosjekter basert på kvalitative metoder i prosessevalueringen kanskje var i største laget, noe som har hatt betydning for evalueringens resultaters direkte anvendelighet for framtiden. Vurderingen som følger av dette er ikke at forskningsfriheten skal reduseres, men heller at dette er noe som bør tas opp til vurdering. Som den tidligere prosjektlederen påpeker, var det gitt på forhånd i søknadene hvilke metoder man ønsket å benytte, og metodevalget kunne derfor neppe komme som en overraskelse.

Kunnskapsdirektøren i AV-direktoratet uttrykker at forskningsfriheten har vært for stor og departementets styring alt for begrenset, noe som har resultert i at evalueringens resultater ikke har vært særlig anvendelige for AV-direktoratet. De ville i større grad ønsket seg et oppdrag som gikk ut på å belyse hvordan organisasjonen kan utvikles i et framtidsperspektiv og hvordan NAV forholder seg til omverdenen, inkludert arbeidsgivere og utdanningssektoren. En annen konsekvens, som vi skal komme tilbake til senere i kapitlet, har vært at mye annen kunnskap enn selve evalueringen av NAV-reformen har blitt anvendt. I stedet fokuserte man på å intervju enkeltmedarbeidere på laveste nivå i organisasjonen. Det blir også påpekt at forskningsprosjektene som har inngått i evalueringen har belyst interessante og relevante perspektiver, men at dette snarere var unntaket enn regelen. Det hadde også betydning for at interessen for å involvere seg i styringsgruppen forsvant, og det samme gjaldt interessen for å invitere forskerne til å komme og holde presentasjoner osv. Synspunktet herfra er derfor at for mye forskningsfrihet er problematisk på flere fronter, og særlig for evalueringens anvendelighet.

4.3.2 Samspillet mellom prosess- og effektdelen

Som nevnt over, var det en av evalueringens sentrale ambisjoner at den skulle bygges opp omkring et samspill mellom prosess- og effektdelen. Denne ideen ble lansert i Mål- og rammedokumentet, og det var store forventninger knyttet til den fra flere hold. Ifølge de intervjuede forskerne fungerte ikke dette samspillet særlig godt, og noen går til og med så langt som til å hevde at samspillet mislyktes fullstendig. Det hersker imidlertid ikke enighet om hvorfor. Noen tilskriver det at premisset er grunnleggende urealiserbart. Andre, og dette representerer flertallet, har tro på premisset og på at det kan la seg gjøre,

men peker på mer konkrete problemstillinger: For tidlig oppstart, en komplisert reform, en "virkelighet" som forandret seg (herunder at reformen ble korrigert underveis) osv. Videre ble det pekt på problemene med å kombinere kvalitative data fra prosessdelen med effektdelen osv. Det hersker derfor ikke enighet om hvorvidt man ved et senere tilfelle kan lykkes med å kombinere en prosessdel med en effektdel.

4.4. Resultat: Den muntlige og skriftlige formidlingen av evalueringen

Formidlingen og offentliggjøringen av evalueringen har funnet sted løpende og på flere fronter. Derfor er det, ifølge lederen for styringsgruppen, viktig at formidlingen ikke bare vurderes ut fra antallet og typen publiseringer, men også ut fra forskernes deltakelse i den offentlige debatten og på konferanser og forskersamlingene, og ikke minst i møtene i Brukerforum. Det er særlig igjennom sistnevnte at resultatene har blitt løpende formidlet. Formidlings- og publiseringsfasen startet derfor så langt tilbake som i 2007, og er i skrivende stund fortsatt i gang. Vi kommer nærmere inn på innholdet i publikasjonene i avsnitt 4.6.

I 2006/2007 ble det besluttet at Forskningsrådet skulle påta seg en sentral og formell rolle knyttet til formidling av resultater, inkludert å arrangere møter i Brukerforum. Flere av intervjuobjektene peker på at møtene i Brukerforum fungerte svært godt, og at forskerne var dyktige til å formidle og diskutere kunnskap. I følge bl.a. lederen av styringsgruppen og forskerne selv la forskerne ned mye arbeid i denne delen av prosjektet, og her hadde man sjansen til løpende formidling av kunnskap og både foreløpige og endelige resultater.

Det hersker delte meninger om forskernes deltakelse i den offentlige debatten. De fleste mener at deltakelsen har vært passe. Enkelte, bl.a. oppdragsgiverens representant i styringsgruppen, påpeker at forskerne har vært for passive og godt kunne ha vært mer aktive når det gjelder å sette dagsorden. De kunne med fordel ha bidratt med innsikt til debatten, bl.a. når brukerne uttaler seg uten å ha den helt store innsikten. I stedet har debatten, i følge den samme representanten, vært dominert av rapporter fra interesseorganisasjoner i stedet for veldokumenterte forskningsresultater fra evalueringen av NAV-reformen.

De fleste av intervjuobjektene har ikke innvendinger mot eller nevneverdige synspunkter på den publiserte forskningen. Det gjennomgående synspunktet er at i og med det er snakk om en forskningsevaluering som legger stor vekt på langsiktig kunnskapsutvikling, så vil resultatet selvfølgelig være en rekke forskningsrelaterte publikasjoner.

4.5 Synspunkter på evalueringens anvendelsesproblemer

Spørsmålet om anvendelsen av evalueringens resultater står sentralt i stort sett alle intervjuene, og er et tema som allerede har blitt diskutert over: Kan og skal resultatene fra en evaluering som denne anvendes instrumentelt, og hvordan? Og hvis de ikke kan anvendes instrumentelt og direkte (som et av resultatområdene legger opp til), er det et problem?

Det er anvendelsesproblemer knyttet til at "evalueringens anvendelse" kan forstås både forholdsvis snevert (instrumentell anvendelse) og mer bredt (f.eks. kunnskapsutvikling, innspill med kunnskap til samfunnet mer generelt), jf. kapittel 2. Det er således viktig å reflektere over hvordan "anvendelse" egentlig skal forstås og å ta forskjellige

anvendelsesformer i betraktning. Dette spørsmålet er også knyttet til at det finnes forskjellige synspunkter på hva en (forskningsbasert) evaluering er.

Anvendelsesproblematikken er også knyttet til spørsmålet om hvorvidt evalueringen har levd opp til de fire resultatområdene som er beskrevet i Mål- og rammedokumentet, jf. kapittel 2. Det hersker bred enighet om at Mål- og rammedokumentets resultatområde 1, som gjelder korrigerende og læring, ikke har blitt innfridd. Evalueringen har bidratt med interessante perspektiver på et mer generelt plan, men de er ikke særlig anvendelige når det gjelder å komme med innspill til implementeringsprosessen og gjøre løpende justeringer. Noen av intervjuobjektene, bl.a. departementsråden fra Arbeids- og sosialdepartementet, knytter den manglende innfrielsen til den valgte organisasjonsformen som ikke fremmer instrumentell anvendelse i og med at den i større grad understøtter langsiktig kunnskapsutvikling. Andre, bl.a. oppdragsgivers representant i styringsgruppen, peker på tidsforskyvningens betydning, inkludert det faktum at det har tatt forholdsvis lang tid før resultatene ble klare. Dessuten peker han på at forskerne generelt har vært svært tilbakeholdne med å komme med anbefalinger underveis. Til slutt uttrykker han det synspunkt at departementet (kanskje) i utgangspunktet hadde for store ambisjoner når det gjelder hva en slik evaluering kan romme.

Generelt hersker det ikke enighet om hvorvidt den manglende instrumentelle anvendelsen bør anses som problematisk eller ikke. Bl.a. peker styringsgruppens leder på at den egentlige anvendelsen har fungert mye bredere, f.eks. gjennom presentasjoner gitt på møtene som har blitt arrangert i Brukerforum. Flere av forskerne er også inne på hvordan resultatene har blitt løpende formidlet på mange andre måter, bl.a. gjennom presentasjoner osv. Poenget er da at selv om den instrumentelle anvendelsen har vært begrenset, så har det funnet sted løpende læring. Det er imidlertid begrenset hvor mye egentlig korrigerende det har resultert i. Som vi skal komme tilbake til, er det imidlertid også viktig å nevne at det har vært satt i gang mange andre typer utredninger og undersøkelser som har bidratt til løpende kunnskap og dermed instrumentell anvendelse, herunder Hagen-utvalget, rapporten fra Riksrevisjonen osv.

Derimot hersker det bred enighet om at evalueringen levde opp til resultatområde 4 – langsiktig kunnskapsutvikling. Begge de intervjuede representantene fra departementet er enige på dette punktet, og det blir gjennomgående ansett som meget positivt at evalueringen har bidratt til oppbyggingen av et stort kunnskapsgrunnlag om gjennomføring av store velferdspolitiske reformer. Tidligere fantes det lite forskningsmessig interesse for velferdstjenester. Evalueringen pekte på noen problemer man allerede har kjennskap til, men ut fra dokumentasjonshensyn er det likevel nyttig at evalueringen har blitt utført.

Det at evalueringen anses som en suksess ut fra hensynet til langsiktig kunnskapsutvikling betyr bl.a. at ikke alle intervjuobjektene anser det for å være et problem at evalueringen ikke levd opp til resultatområde 2. Dette dreier seg bl.a. om departementsråden, lederen av styringsgruppen og forskermedlemmet i styringsgruppen. Begrunnelsen er at dette er en konsekvens av måten evalueringen ble organisert på. Det som er avgjørende er at evalueringen har bidratt til langsiktig kunnskapsutvikling mer generelt og at resultatene har blitt løpende formidlet, bl.a. gjennom møtene i Brukerforum.

Andre av de intervjuede representantene anser det som mer problematisk eller bemerkelsesverdig at evalueringen ikke kunne bidra med løpende læring og bidra til korrigerende. Dette synspunktet uttrykkes bl.a. av departementets representant i styringsgruppen, som påpeker at den opprinnelige ambisjonen var at de løpende resultatene skulle kunne brukes i den løpende styringen og mer operativt. Det at det i stor grad ble brukt kvalitative metoder framheves også som betydningsfullt for den manglende instrumentelle anvendelsen, bl.a. fordi det kan være tvil om hvorvidt resultatene bare representerer særtilfeller. Til slutt påpekes også betydningen av å ha et kritisk blikk på departementets mottakerapparat og på hvordan forskning og annen kunnskap "omsettes" oppover i systemet. Spørsmålet er videre om oppdragsgiver i utgangspunktet har hatt urealistiske forestillinger om hva forskningen kan bidra med i implementeringsprosessen.

Til slutt hevder kunnskapsdirektøren hos AV-direktoratet det synspunkt at evalueringen har vært grunnleggende uanvendelig. Dette tilskriver han evalueringens grunnleggende organisering, som bl.a. har betydd at den i for høy grad har vært dominert av universitetslignende forskning. I tillegg pekes det på for lite styring fra oppdragsgivers side, og på at evalueringen i større grad har anlagt et nedenfra-og-opp-perspektiv og gjennomført enkeltintervjuer med frontlinjen, mens man bl.a. mangler innsikter om hvordan NAV-kontorene kan styres bedre og om samspillet mellom NAV og samfunnet rundt, inkludert arbeidsmarkedet. På den annen side forteller flere av intervjuobjektene, inkludert flere av forskerne, at AV-direktoratet generelt ikke har vært interessert i å gjøre bruk av kunnskap fra NAV-evalueringen, og dette ble bekreftet i intervjuet med kunnskapsdirektøren fra AV-direktoratet. AV-direktoratet på sin side gir uttrykk for at forskerne ikke har vært interessert i deres innspill til evalueringen.

Disse synspunktene vitner om at det har eksistert ulike oppfatninger av og forventninger til hva NAV-evalueringen skulle brukes til: et instrumentelt eller et bredere bruksmål. Begge disse ambisjonene inngår i Mål- og rammedokumentet. Bl.a. påpeker den administrative ledelsen i departementet at det opprinnelig var den langsiktige kunnskapsutviklingen som ble tillagt mest vekt. En del tyder på at AV-direktoratet i utgangspunktet har hatt andre forventninger til evalueringen. Med andre ord har det ikke hersket enighet om formålet i utgangspunktet, og man har kanskje hatt for store eller for uklare forventninger. Spørsmålet er derfor hvor bevisst man i utgangspunktet har vært på disse uklare målsetningene.

Et annet viktig synspunkt som har framkommet gjennom intervjuene er knyttet til departementets mottakerapparat. Flere parter framholder at departementet bør bli mer profesjonelt og dyktigere til å absorbere kunnskap som blir bestilt, slik at den i større grad blir "tatt opp i systemet". Det ene intervjuobjektet fra Forskningsrådet påpekte også at man kunne bli bedre til å understøtte formidling og anvendelse internt. Det som etterspørres er med andre ord en større grad av profesjonalitet. I starten ble det etablert en lesegruppe i regi av departementet. Ambisjonen var, i følge departementets representant i styringsgruppen, at man på denne måten kunne få innsikt i prosessen. I etterpåklokskapens klare lys er det tydelig at dette forumet var for uformelt og ikke sikret tilstrekkelig "absorbering" av kunnskap.

4.5.1 Samspillet med annen kunnskapsproduksjon

Som politisk svar på Riksrevisjonens kritikk ble det oppnevnt en ekspertgruppe, Hagenutvalget, i februar 2010. Flere av intervjuobjektene, bl.a. oppdragsgivers representant i styringsgruppen og kunnskapsdirektøren hos NAV, framholder at resultatene fra Hagen-

utvalget i større grad enn resultatene fra NAV-evalueringen har blitt brukt instrumentelt. Samtidig blir det påpekt at det nettopp var Hagen-utvalgets formål at resultatene skulle kunne brukes instrumentelt, bl.a. til korrigerende underveis. Mange forslag fra denne ekspertgruppen har således blitt implementert. Enda en gang er det viktig å understreke at evalueringen av NAV-reformen hadde et litt annet oppdrag, siden formålet også var dokumentasjon og langsiktig kunnskapsutvikling.

Prosjektlederen peker selv på at det kanskje i større grad burde ha vært en tettere relasjon mellom NAV-evalueringen og annen kunnskapsproduksjon på området. Umiddelbart er det ingen ting som tyder på at forskerne har hatt noe problem med at de ikke har hatt monopol på kunnskap.

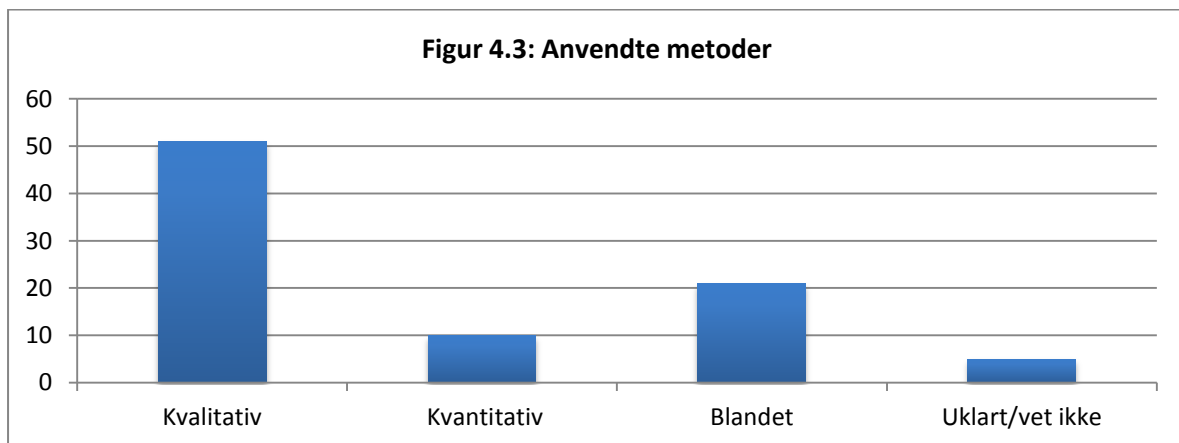
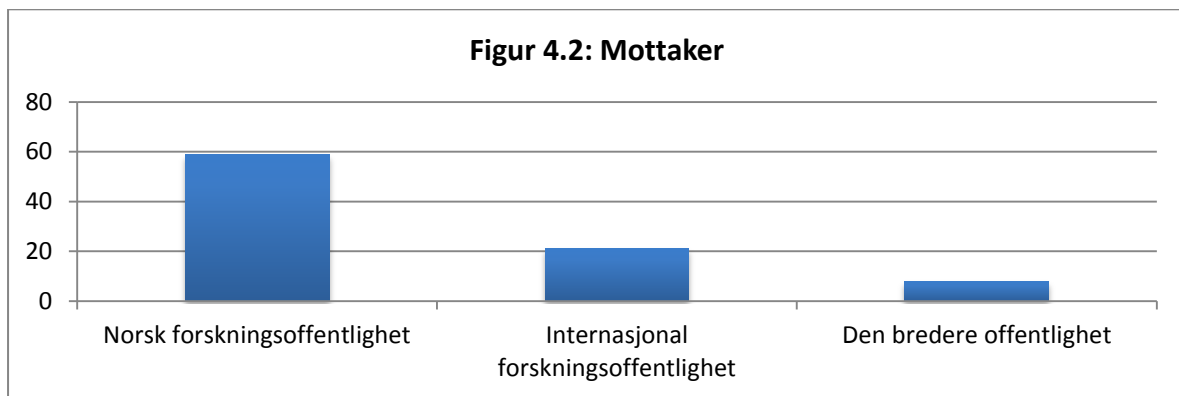
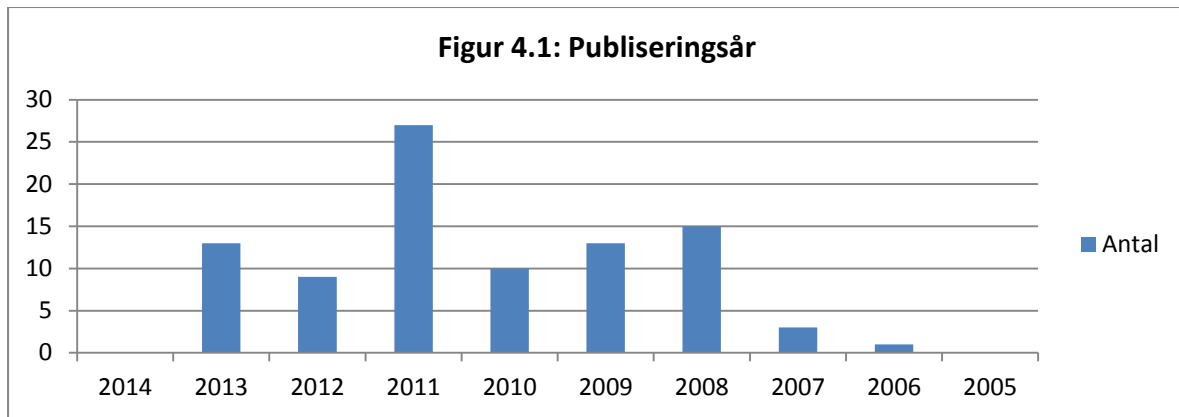
4.6 Evalueringspublikasjonenes innhold

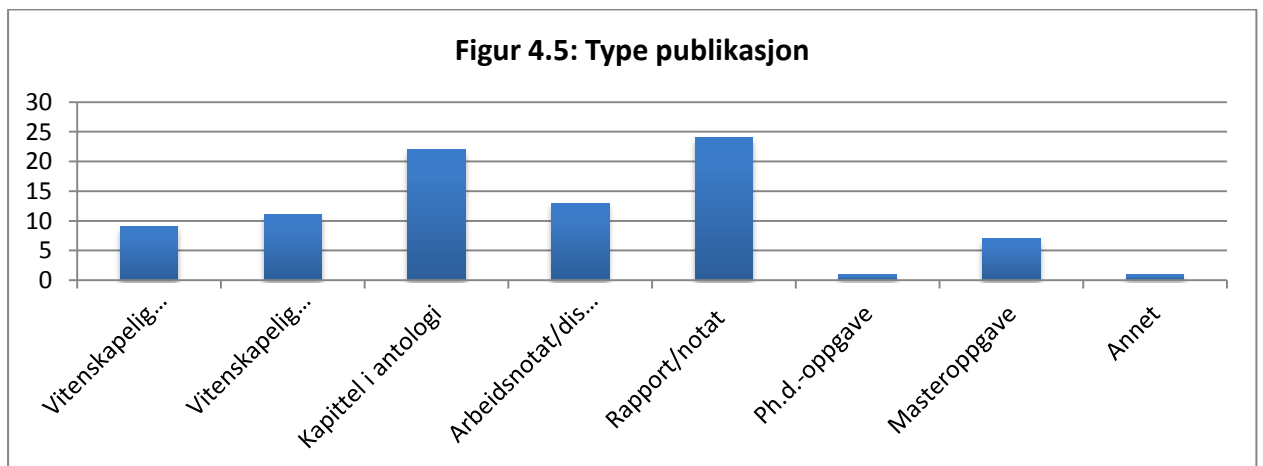
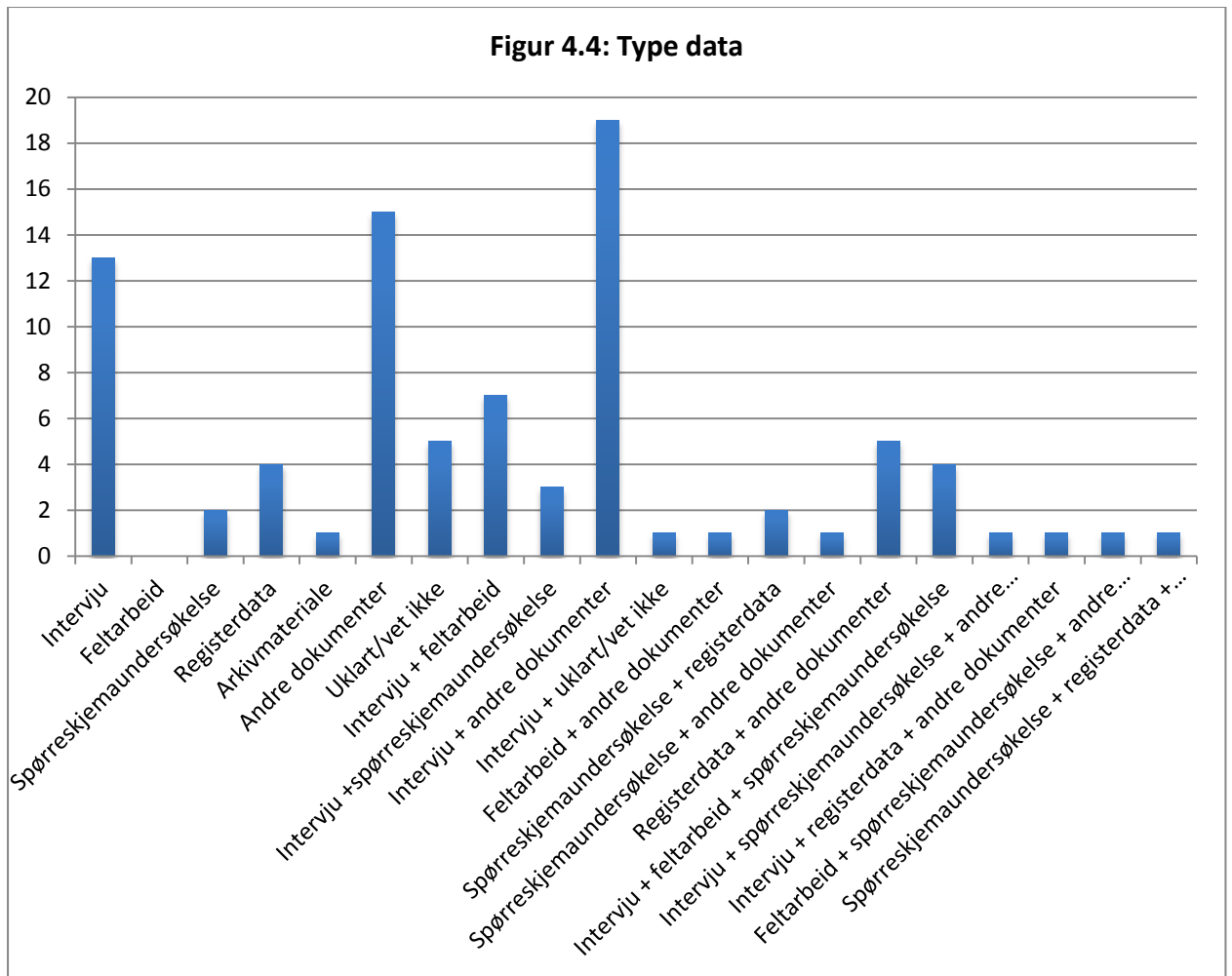
Nedenfor er utvalgte resultater fra kodingen av evalueringspublikasjonene illustrert i seks forskjellige figurer: 1) Publiseringsår, 2) Mottaker, 3) Anvendte metoder, 4) Typer av data anvendt, 5) Type publikasjon, og 6) Konklusjonenes karakter i publikasjonene – inneholder de anbefalinger eller ikke?

Figuren viser fordelinger av de 87 evalueringspublikasjonene som har blitt kodet, og dette dekker alle publikasjoner utgitt i regi av evalueringen av NAV-reformen fram til slutten av mai 2014. Fordelingen viser dermed fordelingen i tall, ikke i prosent.

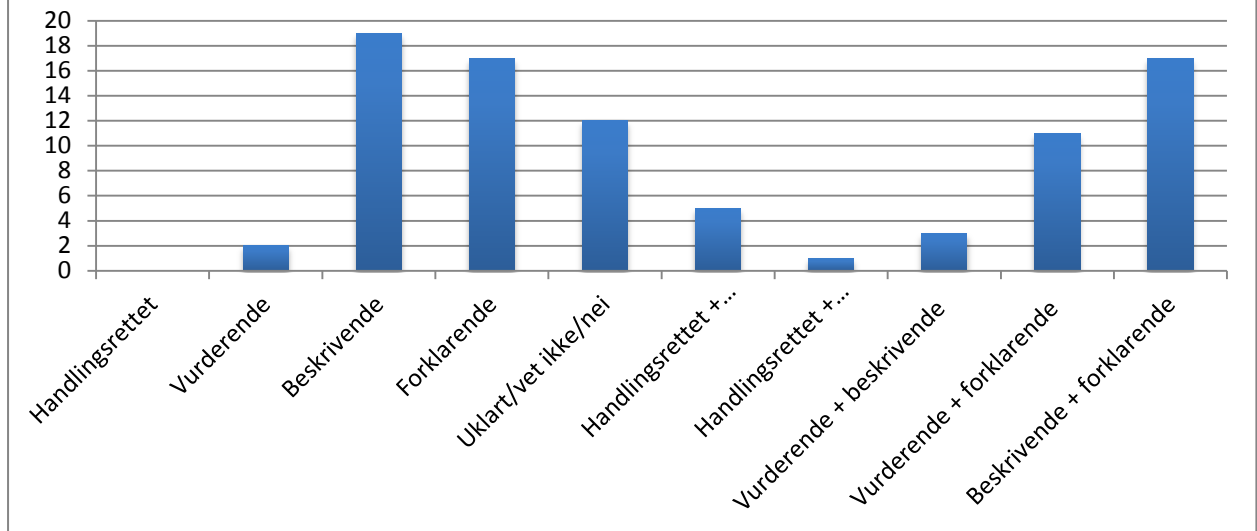
På bakgrunn av dette har vi valgt å framheve følgende punkter:

- 1) Som det framgår av figur 4.1, er publikasjonene utgitt løpende gjennom årene, med en stor overvekt av publikasjoner i 2011. Nesten 1/3 av det aktuelle materialet ble utgitt dette året.
- 2) Som det framgår av figur 4.2, er størstedelen av publikasjonene rettet mot den norske forskningsoffentligheten.
- 3) Som det framgår av figur 4.3, baserer størstedelen av publikasjonene (51 av 87) seg på kvalitative metoder, og en mindre andel på kvantitative metoder.
- 4) Som det framgår av figur 4.4, bruker flere av publikasjonene kombinasjoner av flere typer data, hvorav størstedelen bruker intervjumateriale og dokumenter.
- 5) Ingen av publikasjonene kommer med handlingsrettede anbefalinger i konklusjonen. De fleste av publikasjonenes konklusjoner har i stedet karakter av å være beskrivende eller forklarende.





Figur 4.6: Konklusjonenes karakter – anbefalinger?



4.7 Oppsummering av sentrale synspunkter

Den ovenstående analysen av evalueringen av NAV-reformen vitner om at det blant intervjuobjektene hersker forskjellige syn på hvordan evalueringen har gått og hvordan dilemmaer knyttet til styring, forskningsfrihet og evalueringsanvendelse har gjort seg gjeldende og bør håndteres. En rekke synspunkter har blitt framhevet underveis, og nedenfor oppsummeres og diskuteres disse nærmere. Noen av synspunktene er knyttet til justeringer av eksisterende rammevilkår, mens andre er knyttet til den valgte organisasjonsformen. Synspunktene vitner også om at det finnes ulike meninger om hva som forstås som en god evaluering. Felles for synspunktene som er framhevet nedenfor, er at de ikke bare gjelder evalueringen av NAV-reformen, men dessuten har en mer prinsipiell karakter som gjør at de også er av bredere interesse når lignende evalueringer skal organiseres og gjennomføres i framtiden.

► *Er det hensiktsmessig med en forskningsbasert evaluering forankret i Forskningsrådet?*

De fleste intervjuobjektene er enige om at den valgte organisasjonsformen, inkludert valget av en forskningsbasert evaluering forankret i Forskningsrådet, var grunnleggende fornuftig. Noen av intervjuobjektene, bl.a. forskerne, peker likevel på at et samspill mellom en forskningsbasert evaluering og et mer handlingsrettet evalueringsprogram kanskje kunne være hensiktsmessig. Til slutt har kunnskapsdirektøren fra AV-direktoratet uttrykt det synspunkt at den valgte organisasjonsformen medfører en rekke ulemper, særlig sett i forhold til etterfølgende (instrumentelle) anvendelsesmuligheter. I et framtidsperspektiv reiser dette spørsmålet om hvorvidt det er mer hensiktsmessig med en egentlig handlingsorientert evaluering eller et samspill mellom en forskningsbasert evaluering og andre typer evalueringer, som kan være mer handlingsrettede. Med andre ord: Burde departementets evalueringsdagsorden ha vært en annen? Hva som er mest hensiktsmessig kommer an på det grunnleggende formålet med evalueringen – er det å skape langsiktig kunnskapsutvikling, eller er det å lære og korrigere løpende? Dette dilemmaet reiser det grunnleggende spørsmålet om hva alternativet er i forbindelse med en så stor evaluering.

► *Hvem bør ha innflytelse på utformingen av Mål- og rammedokumentet?*

Flere av intervjuobjektene har synspunkter på hvem som burde være med på å utarbeide og formulere Mål- og rammedokumentet. Som det gikk fram av analysen over, ble det holdt planleggingsmøter i forbindelse med utformingen av dokumentet, hvor bl.a. noen forskere og Forskningsrådet deltok. Dette reiser spørsmålet om hvorvidt de forskerne som kommer med innspill til et slikt oppdrag, senere skal ha muligheten til å søke. Det har også å gjøre med hvorvidt styringsgruppen bør ha større mulighet til å komme med innspill til Mål- og rammedokumentet og ha innflytelse på dets utforming.

► *Hvor spesifikt og styrende bør Mål- og rammedokumentet være?*

Hvor spesifikt og styrende skal et slikt dokument være når det er snakk om en forskningsbasert evaluering? Skal det være et egentlig oppdrag, eller skal det være bredere? I intervjumaterialet kom denne problematikken særlig til uttrykk i diskusjonen av de anvendte metodene og evalueringens fokus. Noen framholdt den utstrakte forskningsfriheten som positiv, mens andre, herunder AV-direktoratet, mente at det burde vært mye mer styring fra oppdragsgivers side.

► ***Bør oppdragsgiver være representert i styringsgruppen?***

Dette synspunktet er særlig knyttet til om oppdragsgiveren bør være representert i styringsgruppen eller ikke, og dermed opptre som en aktør i det som skal evalueres. Ut fra pragmatiske hensyn viser ikke intervjuobjektene til at det finnes store problemstillinger, men på prinsipielt grunnlag blir det påpekt at det kan være problematisk dersom oppdragsgiver er representert. Det blir altså pekt på en rekke både fordeler og ulemper.

► ***I hvor høy grad skal og bør forskningsfriheten ivaretas?***

Dette synspunktet er i høy grad knyttet til spørsmålet om Mål- og rammedokumentets utforming og om evalueringens anvendelighet. Det hersker bred enighet om at forskningsfriheten i høy grad er blitt ivaretatt. Men kan hensynet til forskningsfriheten få for stor plass? Og bør man være mer oppmerksom på at det kan være forskjell på en evaluator- og en forskerrolle?

► ***Bør evalueringer bygges opp omkring store konsortier?***

Det hersker delte meninger om dette. Skapte det flere problemer enn muligheter? Særlig forskerne peker på at denne konsortiumtilnærmingen har resultert i en rekke problemstillinger, og de er kritiske til om denne tilnærmingen var den mest hensiktsmessige.

► ***Samspillet mellom en prosess- og en effektdel?***

I intervjumaterialet framkommer det mange synspunkter på spørsmålet om hvorvidt det er realistisk å integrere en prosess- og en effektdel i forbindelse med offentlige evalueringer. Det hersker imidlertid enighet om at dette kan være en stor utfordring. Dette knytter seg også til spørsmålet om fleksibilitet og om hvordan en evaluering som denne skal igangsettes: Umiddelbart etter reformens igangsettelse, eller når implementeringen er godt i gang? Noen mener at en senere igangsettelse kunne styrket utformingen av evalueringen. Risikoen er at en kunne gå glipp av muligheten til å følge implementeringsprosessen og egenverdien av dette.

► ***Skal resultater fra forskningsevalueringer anvendes, hvordan og til hva?***

Som det framgår over, er dette et gjennomgående problem i forbindelse med evalueringen av NAV-reformen. Intervjumaterialet gjenspeiler store forskjeller mellom ulike oppfatninger av hva en forskningsbasert evaluering bør bidra til og om man i etterkant bør være skuffet eller ikke. Det hersker heller ikke enighet om hvor problemet ligger: Kan anvendelsesproblematikken føres tilbake til den valgte organiseringsformen, forskernes evne og vilje til løpende formidling, eller departementets mottakerapparat?

Kapittel 5. Metaevalueringens oppmerksomhetspunkter

I forrige kapittel gjennomførte vi en analyse av evalueringen av NAV-reformen med det eksplisitte formål å trekke ut noen *oppmerksomhetspunkter* til bruk ved framtidige evalueringer av større politiske reformer. Grunnlaget for analysen var de synspunktene vi ble konfrontert med i forbindelse med intervjuene samt en analyse av innholdet i publikasjonene fra evalueringen av NAV-reformen. Det har ikke vært ekspertpanelets oppgave å rose eller kritisere evalueringen, og således heller ikke å vurdere kvaliteten av evalueringens innhold. Vår oppgave har snarere vært å utlede erfaringer til bruk i tilretteleggingen av framtidige evalueringer av større politiske reformer.

Dette kapitlet består av to deler. I den første delen presenterer ekspertpanelet vår forståelse av noen viktige begreper i rapporten: Store politiske reformer, evalueringer og forskningsbaserte evalueringer. Denne forståelsen har dannet utgangspunktet for ekspertpanelets arbeid. I andre del av kapitlet løfter vi blikket fra den konkrete evalueringen av NAV-reformen og forsøker å trekke ut noen generelle oppmerksomhetspunkter og anbefalinger. I det forrige kapitlet var det aktørene i feltet som framla sine synspunkter på den konkrete evalueringen, men *i dette kapitlet er det ekspertpanelet* som på bakgrunn av den empiriske analysen og våre løpende diskusjoner i panelet, *framhever noen punkter og erfaringer*. Dette er punkter som ekspertpanelet mener at sentrale aktører bør være særlig oppmerksomme på når det gjelder tilretteleggingen av framtidige evalueringer av større politiske reformer. En tidligere versjon av oppmerksomhetspunktene ble presentert på det tidligere omtalte seminaret i Oslo 17. juni 2014. På bakgrunn av seminaret og de senere diskusjonene i ekspertpanelet har vi kommet fram til en revidert liste over oppmerksomhetspunkter som vi presenterer her.

5.1. Utgangspunkt for ekspertpanelets vurderinger

NAV-reformen er en administrativ reform, men som evalueringsobjekt kan den beskrives som en typisk politisk reform i og med at NAV-reformen rommer mange av de samme problemstillingene som vi kjenner fra diskusjonen om evalueringer av større politiske reformer. Det er likevel ikke selvsagt hva som menes med større politiske reformer.

Mange reformer besluttes eller formes i et grenseland mellom politikk og forvaltning, men også mellom institusjoner innenfor så vel som utenfor staten. Innenfor de enkelte stater blir det tatt beslutninger som har politisk betydning og setter rammer for politikernes handlingsmuligheter, men som ikke påvirkes av folkevalgte politikere. At reformer blir besluttet av politiske beslutningstakere behøver heller ikke å bety at de er gjenstand for omfattende politisk debatt. Bare det å få fastsatt hva som skal forstås med en politisk reform, kan bli formålet med diskusjonen.

Denne diskusjonen skal vi likevel se bort fra her. I stedet tar vi utgangspunkt i det vi antar er en ganske utbredt forestilling om hva en politisk reform er. Men hva menes med "store" politiske reformer? Det ligger i sakens natur at målestokken både er relativ og subjektiv. Ut fra et økonomisk perspektiv kan f.eks. reformens andel av statsbudsjettet gi grunnlag for å beskrive en reform som stor. Samtidig er det mulig å forestille seg reformer som er kostnadskrevende, men likevel ikke oppfattes som store. Til slutt kan organisatoriske og institusjonelle reformer som stort sett ikke koster noe, betraktes som store fordi de påvirker spillereglene og de strukturelle betingelsene.

I dette kapitlet refererer betegnelsen "større politiske reformer" til reformer som involverer en rekke forskjellige aktører som opptrer i ulike roller – som beslutningstakere og deltakere i implementeringen, men også som brukere. Målsettingene er ofte flertydige, og mange ganger er det uklart hva som beskrives som de faktiske mål. Dette kan bl.a. skyldes at de forutgående politiske kompromissene har hatt fokus på formuleringen av målene. Hvordan reformen skal presenteres når det gjelder mål og motiv, har vært mer kontroversielt enn reformens saklige innhold. De offisielle målsetningene gir således ikke et fullstendig bilde av reformens bakenforliggende reelle ambisjoner.

Reformen kan videre forventes å ha effekter både innenfor og utenfor målområdet, og dette kan også gjelde det som senere viser seg å være reformens viktigste effekter.

Store reformer bærer også ofte, men ikke alltid, preg av å være irreversible. Man kan nok justere reformen, men man kan ikke gå tilbake til det man gjorde før reformen uten meget store politiske så vel som økonomiske omkostninger. Det gjelder ikke minst relasjonene mellom de forskjellige institusjonene som inngår i reformen.

Det er derfor tydelig allerede fra begynnelsen av at en evaluering av større politiske reformer vil by på en rekke problemer. Dette kan knyttes til at den reformen som besluttes, ofte ikke er den som faktisk blir gjennomført. Reformens faktiske karakter formes i implementeringsprosessen, og det er derfor vanskelig å avgrense selve evalueringobjektet. For det andre er reformen bare en av mange faktorer som påvirker måloppnåelsen. En rekke andre statlige tiltak har også en effekt, i tillegg til andre nasjonale og internasjonale faktorer. Alle disse faktorene inngår i en kompleks interaksjon. For det tredje spiller tidsperspektivets dobbelthet inn: Det tar tid før effektene av en stor reform blir tydelige, men jo lengre tid som går, jo vanskeligere blir det å avdekke hvilke forandringer som skyldes reformen og hvilken betydning andre faktorer har. For det fjerde kan vi merke oss at det ikke er selvsagt hvordan reformens forskjellige mål og det innbyrdes forholdet mellom dem blir forstått. Avslutningsvis vil den konteksten som evalueringen er tenkt brukt i ofte skille seg fra den konteksten som evalueringen ble gjennomført i. Med andre ord: "Virkeligheten" har forandret seg underveis.

Disse og andre faktorer er med på å tydeliggjøre at det er meget begrenset hvor anvendelig kunnskap som høstes av evalueringer av større politiske reformer, er. Dette gjelder særlig spørsmål som "hvordan gikk det?", "hva ble resultatet?", "ble målene oppnådd?" osv.

I forbindelse med flere av oppmerksomhetspunktene kommer vi inn på spørsmål som gjelder evalueringens forskjellige formål og hvordan de blir påvirket av en bestemt form for organisering. Under kommer vi inn på spørsmålet om forskningsbaserte evalueringer og deres forhold til andre typer evalueringer.

Evalueringbegrepet kan beskrives som et "*contested concept*", for å bruke Gallies term (Gallie, 1956). Evalueringer kan i mange tilfeller begrenses i forhold til mange andre former for kunnskapsproduksjon, men det kan også være betydelig overlapping i måten begrepet brukes på. Beskrivende statistikk, som kan inngå som et viktig element i en evaluering, og som har betydning for hva vi ser som problemer i samfunnet, burde vanligvis ikke oppfattes som en evaluering. Heller ikke alle uttalelser om hva som

forklarer en bestemt utvikling, skal oppfattes som en evaluering. Dette gjelder også uttalelser om en reforms "merit and worth".

Vanligvis er det underforstått, eller mer eksplisitt anført, i definisjoner av evalueringer at vitenskapelige metoder skal anvendes. Evalueringen "has to be a carefully conducted assessment" (Vedung, 1997:3). Dette gjelder for så vidt også for forskning generelt. Forskjellen mellom evalueringer og annen forskning er at evalueringer alltid har et program eller en intervensjon som objekt – eller, for å bruke Mertons uttrykk, "purposive social action" (Merton, 1936). En evaluering er således en "investigative activity conducted in relation to something which reflects a particular kind of purpose and intention" (Furubo & Karlsson-Vestman, 2011: 4).

Beslutningen om å gjennomføre en reform bygger ofte på antakelser om forskjellige sosiale, biologiske, økonomiske og andre mekanismer. Slike mekanismer studeres i både samfunnsvitenskapelig og annen forskning. Det er mulig å studere en rekke slike mekanismer. Klassiske studier av f.eks. relasjonen mellom lydighet, autoritet og hvor villige folk er til å skade andre kan påvirke hvordan antimobbeprogrammer konstrueres, og studier av hvordan lærernes forventninger påvirker elevenes prestasjoner kan ha betydning for hvordan lærerutdanningen bygges opp. Denne typen forskning kan imidlertid ikke anses som en evaluering. Forskningsobjektet er de sosiale mekanismene som sådan, og ikke effektene av politisk besluttede handlinger som har som formål å påvirke disse mekanismene.

I tillegg til at evalueringer dreier seg om intervensjoner (eller programmer) i bred forstand, skiller de seg fra annen forskning ved at de gjerne har en identifiserbar oppdragsgiver som iverksetter den aktuelle evalueringen for å få svar på spørsmål angående intervensjonen. Som regel er spørsmålene oppdragsgiveren ønsker svar på om intervensjonen har virket som forventet eller hatt en effekt, men det kan også være at de gjelder forskjellige aspekter ved implementeringen. Evalueringer er ofte, men slett ikke alltid, basert på vitenskapelige metoder, og resultatene blir ikke nødvendigvis publisert og gjort tilgjengelig for andre enn oppdragsgiver. En forskningsbasert evaluering vil naturlig nok ha flere fellestrekk med forskning. I tillegg til å være basert på en vitenskapelig metode, er en forskningsbasert evaluering også kjennetegnet ved at forskningsfriheten blir ivaretatt, at den forholder seg til relevant teoretisk og empirisk forskning (se senere avsnitt), og at resultatene blir åpent formidlet og publisert, noe som ikke er noen selvfølge i forbindelse med andre evalueringer (Halvorsen, 2012). De siste to punktene knytter an til langsiktig kunnskapsoppbygning som en vesentlig målsetting ved forskningsbaserte evalueringer.

Forskningsbaserte evalueringer handler som oftest, men ikke nødvendigvis, om evalueringer av politiske tiltak eller intervensjoner. At det er snakk om en forskningsbasert evaluering betyr imidlertid ikke at andre former for evaluering i mindre grad baserer seg på etablerte vitenskapelige metoder. Evalueringer som gjennomføres av f.eks. statlige myndigheter, enkelte organisasjoner etc. kan i like stor grad anvende vitenskapelige metoder og utføres av personer med høy vitenskapelig kompetanse. Etter ekspertpanelets mening innebærer selve betegnelsen "forskningsbasert" også at evalueringer allerede fra starten av burde gjennomføres med det eksplisitte formål å bidra til å bygge kunnskap. På den måten kan forskningen bli bedre rustet til å svare på spørsmål om hvordan forskjellige samfunnsforhold (problemer?) kan påvirkes gjennom forskjellige politiske innsatser (Prewitt et al, 2012; Vaessen & Leeuw, 2010).

5.2 Oppmerksomhetspunkter

I det følgende gjennomgås de oppmerksomhetspunktene som ekspertpanelet har bestemt seg for å legge vekt på og dermed også framheve som de mest sentrale. Med dette mener vi noen kritiske punkter som de sentrale aktørene i evalueringen – oppdragsgiveren, den evaluerende part og de som blir evaluert – må være oppmerksomme på, forholde seg til og treffe beslutninger om. Oppmerksomhetspunkter knyttet til reformer og evalueringer av disse reformene springer ofte ut av en *balansegang* mellom motstridende, men legitime hensyn (Margetts, 6 & Hood, 2010; Pollitt & Bouckaert, 2011). Det kan være snakk om mer eller mindre rene *kompromisser*, der prioritering av ett hensyn nødvendigvis må bety at et annet blir nedprioritert. Det kan også være snakk om noen *grenser*, for eksempel hensynet til den frie forskning som ikke kan overskrides hvis en evaluering skal kunne sies å være forskningsbasert. Det kan også være snakk om direkte *dilemmaer*, der evalueringen vil bli kritisert uansett hvilke resultater den kommer fram til. Noen ganger viser *paradokser* – i betydningen tilsynelatende motsetninger – seg imidlertid seg å bare være tilsynelatende. Hvis de håndteres på en pragmatisk og løsningsorientert måte som tar i betraktning den konkrete konteksten evalueringen foregår i, kan de i praksis vise seg å være falske motsetninger.

Gjennomgangen er strukturert ut fra den modellen som dannet utgangspunktet for meta-evalueringen, og som ble presentert i figur 2.1 i kapittel 2: *Organisering, prosess, resultat og anvendelse*. I de fleste tilfeller har vi valgt å la oppmerksomhetspunktet stå uten noen egentlig anbefaling fra ekspertpanelet. Ofte vil den konkrete beslutningen om hvordan oppmerksomhetspunktet skal håndteres avhenge av den konkrete konteksten og en ledelsesmessig vurdering av avveininger mellom forskjellige hensyn. I tilfeller der vi har funnet det meningsfylt, har vi likevel valgt å formulere en direkte anbefaling knyttet til oppmerksomhetspunktet.

I formuleringen av ekspertpanelets oppmerksomhetspunkter er vi naturligvis inspirert av og dypt forankret i de synspunktene som er framkommet i metaevalueringens empiriske analyse, men vi har tolket og reformulert og dessuten lagt til og trukket fra en rekke punkter i forhold til de mange synspunktene som ble framført i kapittel 4.

Oppmerksomhetspunkt 1:

Hva er fordelene og ulempene ved en forskningsbasert evaluering i regi av Forskningsrådet sammenlignet med alternative modeller?

Norges forskningsråd har spilt en nøkkelrolle i evalueringen av NAV-reformen, og så vidt panelet forstår også av visse andre politiske reformer i Norge¹². Selv om det finnes få sammenlignbare studier av organiseringen av evalueringsvirksomheten i andre land, vurderer ekspertpanelet det slik at den norske organiseringen av reformevalueringer, i den form den har hatt i forbindelse med NAV-reformen, skiller seg fra praksis i de fleste andre land.

Når organiseringen av evalueringen tas opp, er det ikke tilstrekkelig å forholde seg til aktivitetene innenfor og utformingen av de enkelte evalueringene. Vi må også på et mer overordnet nivå forholde oss til rolle- og maktfordelingen som omgir evalueringprosessen (beslutningen om evaluering, dens omfang (ressurser), beslutningen

¹² De andre igangsatte reformevalueringene gjelder Samhandlingsreformen (helse), Mammografiprogrammet og Pensjonsreformen.

om hvem som skal gjennomføre evalueringen og hvordan den er tenkt anvendt, valg av evalueringsspørsmål og vurderingskriterier, valg av kriterier etc.). Organiseringen berører dermed et sentralt tema i evalueringsdiskusjonen, f.eks. når det gjelder intern og ekstern evaluering og hvilke faktorer som bidrar til eller motvirker forskjellige former for anvendelse.

I Sverige, som det er nærliggende å sammenligne med, finnes det noen delvis overlappende organisasjonsmodeller. De myndighetene som har ansvaret for offentlig aktivitet innenfor en bestemt virksomhet innenfor en sektor, har ofte også et vidtrekkende ansvar for evalueringen av aktiviteten. Parallelt med dette har det blitt opprettet en rekke særskilte evalueringsorganisasjoner, ofte med vitenskapelig kompetanse, innenfor forskjellige sektorer, for eksempel arbeidsmarkedspolitik, kulturpolitikk, kriminalpolitikk osv. Dette betyr at sektortankegangen beholdes, men at evalueringsansvaret skilles fra det operative ansvaret. I tillegg finnes det ikke-sektortilknyttede evalueringsorganisasjoner, f.eks. ekspertgruppen for studier i offentlig økonomi, som fordeler ressurser til forskjellige studier, ikke sjelden basert på tidligere forskning og evalueringer.

I Sverige, men også i Danmark, har forskningsrådene, som kan sammenlignes med Norges forskningsråd, ikke evalueringsoppgaver slik tilfellet er med Norges forskningsråd. I de nordiske landene har vi således forskjellige modeller for hvordan evalueringer kan organiseres.

Valget av organisasjonsmodell for evalueringer kan diskuteres ut fra flere perspektiver. Det er ikke gitt at en spesifikk modell gir bedre forutsetninger for at en evaluering vil bidra til langsiktig kunnskapsutvikling enn andre modeller. Det kan likevel ikke utelukkes at forskjellige modeller kan påvirke *hvilke deler* av forskningssamfunnet som involveres. I en modell som den svenske kan de særskilte sektororganene kanskje medvirke til at forskerne som identifiseres som evaluatorene innenfor et spesifikt område i særlig stor grad kommer til å involveres i forskjellige evalueringer. Den "norske modellen" kan muligens medvirke til at flere forskere enn dem som tidligere har arbeidet med evalueringer innenfor den aktuelle sektoren, lokkes til å svare på anbudsrunder av den typen som ble brukt i forbindelse med evalueringen av NAV-reformen. På denne måten kan en slik modell bidra til pluralisme i evalueringen.

Denne diskusjonen berører også spørsmålet om verdien av substanskunnskap i forbindelse med evalueringer. Til syvende og sist handler det om hvorvidt det finnes en særskilt evalueringskompetanse som en evaluator kan anvende på et hvilket som helst område. Med henvisning til Scrivens svært allmenne definisjon (Scrivens, 1991: 139), der han beskriver evaluanden som "something", er Dahler-Larsen i *The Evaluation Society* inne på den ytterste konsekvensen av dette: "Perhaps it is this very generalization and abstraction of the evaluand, and its liberation from any specific and substantial human activity, that now makes it possible to conceive of evaluation as a distinct cognitive activity in and of itself. If you wanted to evaluate music, you might call a good musician. Now, if you want to evaluate that something, there is a need for an evaluation specialist" (Dahler-Larsen, 2011: 6).

Valget av sektorielle organer for evaluering av statlige virksomheter innenfor spesifikke områder bygger derimot, mer eller mindre eksplisitt, på tanken om at substanskunnskap er viktig ved en evaluering. Modellen innebærer at evaluatorene som evaluerer f.eks.

arbeidsmarkedspolitiske virkemidler, har kunnskap om forskningen som foregår på mekanismene som påvirker bl.a. sysselsetting og arbeidsledighet og konsekvensene av forskjellige statlige intervensjoner. Forskere og andre evaluatorene knyttet til slike organer blir således ikke bare eksperter på evaluering, men også eksperter på selve saksområdet (institusjonelle strukturer, historiske forutsetninger, diskursen osv.). Fordelene ved dette virker åpenbare.

Imidlertid finnes det også ulemper ved en modell som den svenske. De sektorielle organene kan bli del av det (permanente) evalueringssystemet innenfor sektorene, og kan derfor bli alt for tilbøyelige til å akseptere de teoretiske antakelsene som forskjellige reformer bygger på og som de selv kan ha vært med på å formulere. Det kan dermed være rimelig å anta at slike tendenser kan danne utgangspunkt for utforming av ulike indikatorer på at en reform er vellykket (Leeuw & Furubo, 2008). Sannsynligheten for slike problemer kan reduseres med den norske forskningsrådsmodellen.

Ekspertpanelets ovenstående analyse av forskjellige modeller for organisering av evalueringer kan virke noe overfladisk. Vi håper imidlertid at den kan gi anledning til mer presise analyser av fordeler og ulemper ved forskjellige modeller enn det har vært mulig å foreta innenfor rammene av denne metaevalueringen. Vi mener likevel at Norge har valgt en, i et internasjonalt perspektiv, uvanlig og bemerkelsesverdig modell ved å bruke Forskningsrådet som organisator. Vi har inntrykk av at modellen har vært effektiv med hensyn til måloppnåelse på det av evalueringens resultatområder som legger vekt på å bidra til langsiktig kunnskapsutvikling.

Oppmerksomhetspunkt 2:

Hvilke aktører bør trekkes inn ved formuleringen av evalueringsoppdraget i forbindelse med evalueringen av store politiske reformer?

Når oppdragsgiveren formulerer evalueringsoppgaven ved store politiske reformer, kan det være en fordel å trekke inn forskjellige aktører. Dette reiser det prinsipielle spørsmålet om hvilke aktører som bør/skal formulere oppdraget, og dermed også hvem som skal ha muligheten til å påvirke styringen av selve evalueringen. Det kan dels dreie seg om å involvere forskere og eksperter, dels å involvere evalueringens interessenter som inngår i styringsgruppen. Det medfører både fordeler og ulemper å involvere begge typer aktører, og ekspertpanelet konkluderer derfor med at dette er et vesentlig oppmerksomhetspunkt. Ved evalueringen av NAV-reformen, for eksempel, ble styringsgruppen ikke involvert i formuleringen av evalueringsoppgaven.

Fordelen ved å involvere forskere og eksperter kan være at de kan bidra med råd og veiledning, noe som kan bidra til en presisering av oppdraget. Det som imidlertid kan være en utfordring er at i små land som Norge og de andre skandinaviske landene, der forskningsmiljøene er relativt små innenfor noen områder, kan det være de samme forskerne som deretter velger å søke på den samme evalueringen. De vil derfor, under ellers like forhold, ha en konkurransemessig fordel i anbudsprosessen i forhold til andre evalueringskonsortier. Ekspertpanelet mener det er snakk om en avveining mellom hensynet til en adekvat oppgaveformulering og hensynet til en rettferdig anbudsprosess på like vilkår. Ekspertpanelets konklusjon er at det er snakk om et kompromiss mellom to legitime hensyn, og at dette må håndteres ut fra ledelsesmessig skjønn i hvert enkelt tilfelle. Etter ekspertpanelets syn er det dog ikke hensiktsmessig om habilitetsproblemer blir et hinder for at den best mulige ekspertise trekkes inn ved formuleringen av

evalueringsoppgaven. Mer bruk av internasjonale eksperter, f.eks. fra de andre skandinaviske landene, kan være en del av løsningen.

Involvering av evalueringens interessenter, som inngår i styringsgruppen, i selve formuleringen av evalueringsoppgaven, kan bidra til å sikre eierskap til og sterkt engasjement i evalueringen fra styringsgruppens side. Involvering av interessentene i formuleringen av evalueringsoppdraget gir også anledning til å avklare interessentenes forventninger til og forståelse av hva en evaluering er og hva den krever av ressurser og involvering fra deres side.

En mulighet som ble foreslått på metaevalueringseminaret, og som gjelder involvering av potensielle aktører mer generelt, er å invitere alle potensielle evaluatorene til et åpent seminar der formuleringen av evalueringsoppdraget drøftes. I den grad man måtte velge å selektivt trekke inn nasjonale eksperter som også leverer tilbud på oppgaven, bør det tas høyde for denne skjevheten ved tilretteleggelsen av anbudsprosessen. Det bør gis rikelig tid slik at konsortier som ikke har vært involvert har muligheten til å formulere konkurransedyktige tilbud, og anbudsdokumentene bør utformes på en slik måte at de letter dette arbeidet.

***Oppmerksomhetspunkt 3:
Hvilke medlemmer bør inngå i styringsgruppen?***

Ekspertpanelet finner at viktige hensyn tilsier at de vesentligste interessentene er representert i styringsgruppen ved evalueringen av store politiske reformer, slik som ved evalueringen av NAV-reformen. Dette gjelder både representanter for oppdragsgiveren, for organisasjonene som skal implementere reformen, og for evaluatorene (forskerne). I tillegg kan relevante brukerorganisasjoner også være representert i styringsgruppen. En slik sammensetning kan bidra til å fremme kunnskapsdeling og anvendelse av resultater, og sikre at det finnes et forum der viktige uenigheter og praktiske problemer i forbindelse med evalueringen kan håndteres og avklares.

Det har vært argumentert for at oppdragsgiver ikke bør sitte i styringsgruppen, siden dette kan representere et brudd med armlengdeprinsippet. På samme måte kan det argumenteres for at andre interessenter, for eksempel implementeringsorganisasjonene, ikke burde sitte i styringsgruppen. Det kan skape problemer hvis det kan reises tvil ved legitimiteten som er knyttet til autonomi og høy forskningsmessig kvalitet.

Balansegangen mellom oppdragsgivers og andre gruppers legitime interesse og hensynet til den frie forskning bør imidlertid ofte kunne sikres gjennom å etablere prosedyrer for samarbeidet mellom styringsgruppen og forskerne.

Igjen er det snakk om en avveining som avhenger av kjennetegn ved den konkrete reformen som skal evalueres. Hvis det er snakk om en kraftig politisert reform med høy grad av samfunnsmessig mistillit til oppdragsgivers rolle, kan det være aktuelt å formalisere armlengdeprinsippet gjennom sammensetningen av styringsgruppen. Eventuelt kan oppdragsgiver ha observatørstatus i slike tilfeller.

Oppmerksomhetspunkt 4:

Hvilke prinsipper bør kjennetegne relasjonen mellom styringsgruppe og forskere?

I en forskningsbasert evaluering organisert av Forskningsrådet vil den langsiktige kunnskapsutviklingen ofte naturlig settes i høysetet, mens den løpende instrumentelle anvendelsen av foreløpige resultater som kunne være nyttig i forbindelse med tilpasning og implementering av reformen, ofte vil kreve tiltak.

Her kan styringsgruppen, sammen med prosjektlederen og forskningskoordinatorer, spille en vesentlig og legitim rolle som brobygger og formidler mellom forskning og implementeringsprosess. En aktiv styringsgruppe kan også legge til rette for at ressurser og kompetanse som kan være til nytte for evalueringen, blir stilt til disposisjon, noe som ofte kan ha stor positiv betydning for evalueringsprosessen og evalueringens kvalitet.

Det må imidlertid sikres at denne brobygger- og formidlingsrollen ikke bryter med armlengdeprinsippet og hensynet til høy forskningsmessig kvalitet. For å bevare de kvalitetene som kjennetegner forskningsbaserte evalueringer og sikre evalueringens legitimitet, er det noen grenser for innblanding som må overholdes.

Det er rimelig at oppdragsgiveren og medlemmene i styringsgruppen har en mulighet til å få informasjon om evalueringens analyser og konklusjoner før de offentliggjøres. På denne måten kan oppdragsgiver og implementeringsorganisasjoner ligge i forkant og raskt trekke inn relevante resultater i den løpende implementeringen av reformen.

Oppdragsgiver og implementeringsorganisasjoner vil også i enkelte tilfeller sitte inne med faktakunnskap som kan bidra til å bedre kvaliteten på publikasjoner fra evalueringen.

Men slik innblanding må aldri få karakter av å være noe mer enn velmente forslag. Det må ikke kunne stilles spørsmål ved forskernes uavhengighet. Dette er allmenngyldig og er etter ekspertpanelets mening særlig viktig ved evalueringer av store politiske reformer. I forbindelse med slike reformer vil det ofte være vesentlige interesser knyttet til å påvirke evalueringen i bestemte retninger. Det må aldri være tvil om forskningens integritet og uavhengighet av oppdragsgivers og andre aktørers eventuelle interesse av å skjule problemer ved implementeringsprosessen.

Ekspertpanelet merker seg at Forskningsrådet har bestemt at oppdragsgiver ikke skal være representert i styringsgruppen ved evalueringer som Forskningsrådet står for.

Oppmerksomhetspunkt 5:

Hvordan håndteres flere og vanskelig forenelige målsetninger for evalueringen?

Etter ekspertpanelets oppfatning har valget av Forskningsrådet til å organisere NAV-evalueringen betydd at oppdragsgiveren – bevisst eller ubevisst – har valgt å prioritere langsiktig kunnskapsutvikling som et vesentlig element i evalueringssoppgaven. Ved å ivareta og prioritere Forskningsrådets lange og velutviklede tradisjon for å sikre høy kvalitet, har man (formodentlig) tenkt at man sikrer både evalueringens kvalitet og legitimitet ved at forskningen foregår i henhold til armlengdeprinsippet fra oppdragsgiver. Samtidig inngår det også som et resultatområde i Mål- og rammedokumentet at formålet

med evalueringen er å bidra med å løpende gjøre evalueringens resultater anvendelige i forbindelse med implementeringen.

Etter vår mening burde de utfordringene som de forskjellige anvendelsesformålene skaper, ha vært gjenstand for en mer omfattende diskusjon allerede før valget av organisasjonsform. På NAV-evalueringens avsluttende konferanse 12. juni 2014 siterte styregruppens leder Peter Dahler-Larsen den franske sosiologen og antropologen Marcel Mauss som følger: "I vitenskapelige spørsmål kan man ikke gå for langsomt frem. I praktiske spørsmål er det ingen tid å miste". Selv om sitatet setter tingene på spissen, er det en tendens til at langsiktig kunnskapsutvikling setter andre og delvis motsatte krav til evalueringen enn behovet for løpende tilbakemelding til bruk i implementeringsprosessen.

Løpende anvendelse av evalueringresultatene kan finne sted på flere nivåer, fra det å stilling på et mer overordnet nivå til selve fordelingen av oppgaver og ansvar mellom ulike aktører, og til forandringer i de lokale kontorenes arbeidsmåte. De delene av evalueringen som gjennomføres med tanke på løpende tilbakemelding, må derfor målrettes mot disse ulike nivåene. En slik evaluering krever en tett relasjon mellom de som evaluerer og de forskjellige aktørene som skal gjøre bruk av evalueringresultatene (Hansen, 2005), og dette kan være en utfordring i forhold til autonomi og armlengdeprinsippet. Noen evalueringforskere som legger vekt på direkte anvendelse, f.eks. Patton (Patton, 1997), har derfor lagt vekt på at evalueringer snarere enn distanse og armlengdes avstand krever nærhet til dem som skal gjøre løpende bruk av evalueringen. Det hevdes videre at evalueringer med slike instrumentelle målsetninger må innebære at evaluatoren selv blir involvert i de virksomhetene som skal evalueres.

Vi mener derfor at det her er snakk om et paradoks, i betydningen tilsynelatende motsetning, når målsetningen både er løpende tilbakemelding og langsiktig kunnskapsutvikling. Det hersker delte meninger om hvorvidt det er snakk om en genuin konflikt mellom kravene til autonomi ved langsiktig kunnskapsutvikling og kravene til brukbar løpende tilbakemelding. Vi mener likevel at det finnes logistiske og praktiske problemer som kan gjøre det vanskelig å la en evaluering romme så forskjellige anvendelsesmålsetninger. Vi har forståelse for at det kan være hensiktsmessig, ut fra taktiske hensyn i planleggingsfasen, å hevde at forskjellige målsetninger skal inkluderes i evalueringen.

Evalueringen av NAV-reformen vitner om de problemene som oppstår når evalueringer med flere målsetninger (knyttet til anvendelse), skal gjennomføres. Selv om det kan la seg gjøre å utforme en evaluering som tar sikte på langsiktig kunnskapsutvikling på en slik måte at den også bidrar til løpende instrumentell anvendelse, så vil *denne* formen for anvendelse neppe være en forskningsbasert evalueringens sterkeste side. Vi mener derfor at beslutninger om valget av organisering for kommende evalueringer i fremtiden må omfatte overveielser av i hvilken utstrekning forskjellige organiseringsmåter fremmer forskjellige målsetninger.

Oppmerksomhetspunkt 6:

Hvordan innarbeides fleksibilitet i evalueringsprosessen?

I innledningen til dette kapitlet har vi beskrevet visse kjennetegn på store politiske reformer. Blant annet ble det påpekt at reformens faktiske karakter i høy grad formes i selve implementeringsprosessen. Dette er i høy grad blitt tydeliggjort ved NAV-reformen. Om det er den intenderte eller den faktiske reformen som evalueres, kommer derfor an på hvordan evalueringen besluttes og designes. Etter vår vurdering er det viktig at man forsøker å evaluere den reformen som faktisk gjennomføres.

Forskjellige ideer til dette kan hentes fra metaevalueringens empiriske undersøkelse av evalueringen av NAV-reformen. De viktigste er vel for det første å forholde seg mer avventende i evalueringens første fase med tanke på å observere hvordan reformen arter seg i praksis, og for det andre, i forbindelse med dette, å sette til side midler til uforutsette analyser som underveis viser seg å være særlig viktige. På denne måten får styringsgruppen, i samspill med prosjekt- og modullede, en mer vesentlig oppgave enn i tradisjonelle forskningsprosjekter med å utvikle evalueringsdesignet på grunnlag av hvordan reformen utvikler seg. Selve reformens foranderlighet i implementeringsprosessen må anses som et viktig spørsmål i evalueringene. Dette aspektet har det ikke blitt tatt tilstrekkelig høyde for i beslutningen om og utformingen av NAV-evalueringen. Dette er derfor noe som bør diskuteres grundigere i forbindelse med framtidige evalueringer.

Oppmerksomhetspunkt 7:

Hvor ambisiøs bør man være med tanke på å integrere forskjellige forskningsmoduler i en evaluering?

I metaevalueringens empiriske undersøkelse av evalueringen av NAV-reformen har flere forskere stilt spørsmålsteget ved hvor ambisiøs man skal være når det gjelder å koordinere og integrere forskjellige forskningsmoduler i et større samlet forskningsprosjekt, eller om man i større grad burde operere med en løsere prosjektkoordinering eller flere uavhengige forskningsprosjekter. Argumentet mot ambisiøs integrering er at man kan kaste bort mye tid og mange ressurser på å forsøke å integrere forskjellige forskningstradisjoner i tilfeller der det er meget vanskelig, om ikke umulig, å bygge bro mellom dem. Argumentet for ambisiøs integrasjon er at forskjellige forskningsprosjekter med en godt gjennomtenkt sammenheng kan styrke troverdigheten til forskningsdesignet som helhet.

Etter ekspertpanelets syn er dette et vesentlig oppmerksomhetspunkt som må avhenge av en vurdering av den enkelte evalueringsoppgave. Ekspertpanelet merker seg likevel at metaevalueringens empiriske analyser og NAV-evalueringens avsluttende konferanse tyder på at man her, til tross for en del frustrasjoner underveis, er kommet langt med koordineringsstrategien. Prosessevalueringene gir således noen troverdige forslag når det gjelder prosessene som kan forklare resultatene fra effektevalueringene. Forskjellige nye tendenser innenfor evalueringsteori peker også på hvordan prosessevalueringenes kartlegging av årsakssammenhenger ofte er helt avgjørende for å kunne si noe om hvorvidt effekter som avdekkes av effektmålinger skyldes reformen eller andre forhold, og, enda viktigere, hvilke kjennetegn ved reformen og dens kontekst som forårsaker effektene.

Oppmerksomhetspunkt 8:

Bør mulighetene til å evaluere være styrende når reformer skal besluttes og implementeres?

NAV-evalueringens utfordringer med å avdekke effekter aktualiserer diskusjonen om hvorvidt det bør tas hensyn til evalueringen i forbindelse med tilretteleggingen av politiske reformer. Bør reformer ideelt sett organiseres slik at det kan utføres "ordentlige" effektmålinger organisert som kvasi-eksperimenter med (ideelt sett) randomiserte kontrollgrupper og før- og ettermålinger? Dette er en gammel idé fra 60-tallet, og praksis siden tyder på at dette sjelden eller aldri lar seg gjøre for store politiske reformer. Synspunktet har også fått en renessanse gjennom diskusjonen av evidens, og også gjennom den tidligere amerikanske administrasjonens syn på disse spørsmålene.

Ekspertpanelet mener at disse standpunktene er problematiske ut fra to hensyn: For det første burde det i dag herske enighet, også mellom de ulike sidene i evidensdebatten, om at muligheten for å uttale seg med sikkerhet om effektene av store politiske reformer er meget begrenset. Forskjellige institutter (f.eks. Cochrane og Campbell) som arbeider med disse spørsmålene har vist at langt de fleste evalueringer, også evalueringer gjennomført av veletablerte forskere, ikke bidrar med slik kunnskap. Den andre siden av debatten har bl.a. problematisert selve spørsmålet om hvilken type kunnskap som skal regnes for evidens og den grunnleggende forståelsen av begrepet. De siste tiårs debatt gjør oss derfor mindre tilbøyelige til å tro at reformer bør organiseres ut fra evalueringers krav om å belyse reformens effekter. Vi betrakter derfor disse standpunktene som grunnleggende urealistiske, ikke bare i synet på hvordan politikk blir til, men også i synet på de mulighetene som kunnskapsproduksjon gir.

Standpunktet er også problematisk ut fra et normativt og demokratisk perspektiv, da det innebærer at politikernes krav skal underordnes kunnskapsproduksjonen i samfunnet og evaluators krav. Dette reiser spørsmålet om hvilken rolle effektmålinger bør ha og hvordan de bør organiseres. Dette er ikke et spørsmål som kan besvares på generelt grunnlag, men det er et vesentlig oppmerksomhetspunkt ved tilretteleggelsen av evalueringer av politiske reformer.

Litteratur og anvendte dokumenter

Andreassen, T. A. & J. Aars (2014): *NAV-reformen - iverksettelsesprosess og resultater* (arbeidstittel). Upublisert bokmanus til sluttpublikasjon for evalueringen.

Andreassen, T. A. & K. Fossetøl (2011): "Innledning – Organisasjonsendring som velferdsreform" i Andreassen, Tone Alm og Knut Fossetøl (red.): *NAV ved et veiskille. Organisasjonsendring som velferdsreform*. Oslo: Gyldendal akademiske.

Arbeidsdepartementet- og inkluderingsdepartementet (2006): *Mål og rammedokument for evaluering av NAV-reformen*. Oslo.

Christensen, T. (2011): *Etablering av forvaltningsenhetene og pensjonsenhetene i NAV – en gjennomtenkt reorganisering med positive effekter?* Rapport 2 – 2011. Uni Rokkansenteret.

Christensen, T. & P. Lægveid (2012): *Welfare reform and "wicked issues" – from coupling to de-coupling?* Working paper 2 – 2012. Uni Rokkansenteret.

Dahler-Larsen, P. (2011): *The Evaluation Society*. Stanford, CA: Stanford University Press.

Dahler-Larsen, P. & H. K. Krogstrup (2003): *Nye veje i evaluering*: Systime Academic.

Furubo, J.E. & O. Karlsson Vestman (2011): "Evaluation For Public Good or Professional Power" i Eliadis, P., Furubo, J.E. & Jacob, S. (red.): *Evaluation: Seeking Truth or Power?* New Brunswick, NJ: Transaction Publishers.

Gallie, W.B. (1956): "Essentially Contested Concepts", *Proceedings of the Aristotelian Society*, Vol. 56, pp.167-198.

Hagen, T. P.; I. J. Sand; G. Kvåle; L. E. Borge & A. Persson-Grivas (2010): *NAVs 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.

Halvorsen, A. (2012): "Forskningsbaserte evalueringer og andre evalueringer" i Halvorsen, A. E. Lier Madsen og N. Jentoft (red.): *Evalueringer. Tradisjoner, praksis, mangfold*. Bergen: Fagbokforlaget.

Hansen, M. B. (2005): *Anvendelse af analyser i beslutningsprocesser. Det beslutningsteoretiske grundlag for mini-MTV*. Center for Evaluering og Medicinsk Teknologivurdering. Sundhedsstyrelsen, pp. 1-19.

Innst.S. nr. 189 (2002-2003): *Innstilling fra sosialkomiteen om samordning av Aetat, trygdeetaten og sosialtjenesten*.

Innst.S. nr. 198 (2004-2005): *Innstilling fra sosialkomiteen om ny arbeids- og velferdsforvaltning*.

Leeuw, F.L. & J.E. Furubo (2008): "Evaluation systems. What are they and why study them", *Evaluation*, Vol. 14, No 2.

Margetts, H.; P. 6 & C. Hood (2010). *Paradoxes of modernization: Unintended Consequences of Public Policy Reform*: Oxford: Oxford University Press.

Merton, R. K. (1936). "The unanticipated consequences of purposive social action", *American Sociological Review*, Vol 1, No 6.

Møller, G. (2006): *Evaluering av samarbeidsforsøkene mellom Aetat, trygdeetaten og sosialtjenesten*. Sluttrapport. Telemarksforskning rapport 231.

NOU (2009, 13): *Brede pensjonsordninger*.

NOU (2004, 13): *En ny arbeids- og velferdsforvaltning: Om samordning av Aetats, trygdeetatens og sosialtjenestens oppgaver*.

Patton, M. Q. (1997): *Utilization-focused Evaluation*. The New Century Text. Thousand Oaks: SAGE.

Pollitt, C. & G. Bouckaert (2011). "Trade-Offs, balances, limits, dilemmas, contradictions and paradoxes." I C. Pollitt & G. Bouckaert (red.), *Public Management Reform. Comparative Analysis of New Public Management, Governance and the Neo-Weberian State*. Oxford: Oxford University Press, pp. 182-205.

Prewitt, K., T.A. Schwandt, & M.L. Straf (2012): *Using Evidence in Public Policy*. Washington, DC: National Research Council of the National Academies.

Ot.prp. nr. 96 (2004-2005): *Om lov om interimsorganisering av ny arbeids- og velferdsetat*.

Ot.prp. nr. 47 (2005-2006): *Om lov om arbeids- og velferdsforvaltningen (arbeids- og velferdsforvaltningsloven)*.

Reegård, Kaja (2009): "Historien om NAV-reformens unnfangelse" i B. R. Nuland, B. S. Tranøy & J. Christensen (red.), *Hjernen er alene. Institusjonalisering, kvalitet og relevans i norsk velferdsforskning*. Oslo: Universitetsforlaget.

Scriven, M. S. (1991): *Evaluation Thesaurus*. Newbury Park: SAGE Publications.

Scriven, M. S. (1969): "An introduction to meta-evaluation", *Educational Products Report*, 2, pp. 36-38.

Stufflebeam, D. L. (2001): "The Metaevaluation Imperative", *American Journal of Evaluation*, Vol. 22, No. 2, pp. 183-209.

Shulha, L. M. & J. B. Cousins (1997): "Evaluation use: Theory, research, and practice since 1986", *Evaluation Practice*, Vol. 18, pp. 195-208.

St.meld. nr. 14 (2002-2003): *Samordning av Aetat, trygdeetaten og sosialtjenesten*.

St.meld.nr. 9 (2006-2007): *Arbeid, velferd og inkludering*.

St.prp.nr. 46 (2004-2005): *Tilråding fra Arbeids- og sosialdepartementet om etablering av ny arbeids- og velferdsforvaltning*.

St.prp. nr. 51 (2008-2009): *Redegjørelse om situasjonen i arbeids- og velferdsforvaltningen og forslag om tilførsel av ressurser til Arbeids- og velferdsetaten*.

Vaessen, J. & F. L. Leeuw (red.) (2010): *Mind the gap. Perspectives on policy evaluation and the social sciences*. New Brunswick & London: Transaction Publishers.

Vedung, E. (1997): *Public Policy and Program Evaluation*. New Brunswick: Transaction Publishers.

Weiss, C. H. (1998): "Have we learned anything new about the use of evaluation?" *American Journal of Evaluation*, 19:21-33.

Øvrige dokumenter

Riksrevisjonens rapport om revisjon av Arbeids- og velferdsetaten for budsjettåret 2008, tillegg 2 til Dokument 1 (2009-2010)

Samarbeidsavtale mellom Arbeids – og Inkluderingsdepartementet (AID) og Norges forskningsråd (NFR) om forskningsbasert evaluering av NAV-reformen.

Referater fra styringsgruppemøtene

Bilag 1: Liste over intervjuobjekter

Medlemmer av styringsgruppen:

- 1) Peter Dahler-Larsen, professor Københavns Universitet (DK) og leder av styringsgruppen
- 2) Ann-Helen Bay, Institutt for Samfunnsforskning og forskermedlem i styringsgruppen
- 3) Odd Helge Askevold, departementet/oppdragsgivers medlem i styringsgruppen.

Forskerne:

- 4) Anne Lise Fimreite, forhenværende prosjektleder
- 5) Jacob Aars, nåværende prosjektleder og leder av modul 2
- 6) Tone Alm Andreassen, leder av modul 3
- 7) Knut Røed, leder av modul 5 (fram til 2011)
- 8) Tom Christensen, leder av modul 1

Den administrative ledelsen i Arbeids- og sosialdepartementet:

- 9) Ellen Seip, departementsråd

AV-direktoratet og Kommunenes interesseorganisasjon

- 10) Yngvar Åsholt, kunnskapsdirektør (AV-direktoratet)
- 11) Magne Hustad, seniorrådgiver (KS)

Forskningsrådet:

- 12) Eivind Hovden, avdelingsdirektør
- 13) Steinar Kristiansen, seniorrådgiver

Bilag 2: Mandatet for metaevalueringen

Meta-evaluering av evalueringen av NAV-reformen – mandat og organisering

Bakgrunn

Evalueringen av NAV-reformen startet i 2006 og avsluttes våren 2014.

Styringsgruppen har sluttet seg til sekretariatets forslag om å igangsette en meta-evaluering. Meta-evalueringen er en evaluering NAV-evalueringen (EVA-NAV), med særlig vekt på prosesser og gjennomføring. Utgangspunktet er spørsmålet om *hva kan læres av evalueringen av NAV-reformen*.

Dette spørsmålet kan ha en bredere interesse: I løpet av de siste ca 10 årene har interesse og bruk av ulike former for evalueringer økt mye. Samtidig har erfaringer blant oppdragsgivere på ulike nivåer i offentlig sektor ikke vært samlet og syntetisert. Både Direktoratet for økonomistyring (DFØ) og Norsk evalueringsforum er instanser for å vise og samle evalueringsaktivitetene, men foreløpig uten å utvikle en egen arena for læring. Vår kontakt med DFØ tyder imidlertid på interesse for nettopp det.

For å sikre gode evalueringsprosesser er det viktig å lære av evalueringer som er gjennomført. NAV-evalueringen kan være et case som viser hvilke læringsspørsmål som er aktuelle når offentlig sektor planlegger, bestiller og bruker evalueringer, knyttet til roller, organisering og prosesser.

I desember 2012 godkjente DSAH retningslinjer for slike evalueringsoppdrag til Forskningsrådet. Reformevalueringer inneholder spørsmål som er relevante for evalueringer av andre tiltak på ulike nivåer.

Meta-evalueringen avgrenses til Forskningsrådets evaluering - hvordan evalueringen er lagt opp og gjennomført. Styringsgruppen er tilfreds med at denne evalueringen organiseres av administrasjonen.

Mandat

Mandatet for meta-evalueringen inneholder følgende punkter:

- Bakgrunn og planlegging
- Formål: Evaluere prosesser som er gjennomført i planlegging, gjennomføring og avslutning av NAV-evalueringen. Meta-evalueringen skal gi grunnlag for myndigheter/oppdragsgivere og forskere for å lære av prosessene.
- Meta-evalueringen skal omfatte organisering, metoder og rollefordeling (oppdragsgiver/AD, Forskningsrådet, styringsgruppen, forskerne), utlysning og oppstart, drift, formidling og publisering
- Meta-evalueringen gjennomføres av et skandinavisk ekspertpanel med tre medlemmer. Det engasjeres en sekretær ved lederens institusjon.
- På NAV-evalueringens budsjett er det avsatt inntil 900 000 kroner til dekning av utgifter til honorarer, reise- og oppholdsutgifter.
- Meta-evalueringen gjennomføres i løpet av januar-juni 2014.

- Meta-evalueringen som case vurderes presentert på en konferanse hvor de prinsipielle spørsmålene har en sentral plass.
- Meta-evalueringen skal avsluttes med en rapport på inntil 50 sider. Rapporten skal også inneholde forslag til prinsipper for god organisering av Forskningsrådets evaluering av politiske reformer
- Kilder: Mål- og rammedokumentet/oppdraget fra Arbeidsdepartementet til Forskningsrådet, utvalgte andre dokumenter og informanter blant aktørene (oppdragsgiver/AD, forskerne, styringsgruppen og Forskningsrådet)

Hovedtemaer

Meta-evalueringen skal også belyse og drøfte noen hovedtemaer:

TEMA	Problemstilling	Kommentar
Politikkens legitime rolle	I hvilken grad har man klart å finne det gode balansepunktet mellom faglig (vitenskapelig) styring/ og forskningsmessig frihet på den ene siden, og politikken/oppdragsgivernes legitime rolle før og under evalueringen?	Et klassisk problem som har relevans for mye oppdragsforskning.
Aktør-interesser og betydningen av deltakelse og debatt	Hvordan har ulike aktørinteresser kommet til uttrykk i evalueringsprosessen? I hvilken grad har evalueringen klart å finne det gode balansepunktet mellom distanse og deltakelse i debatten?	Identifisere aktørene!
Roller og forventninger til forskerne i evalueringen	Hvilke forventninger har forskerne møtt i prosessen, og hvordan har de møtt dem? I hvilken grad har dette vært forventninger som har skapt problemer i evalueringsprosessen?	Ikke alle forventninger er like legitime! Det må også tas høyde for at forventninger fra ulikt hold kan være motstridende. Legge særlig vekt på «forventningsgapet» mellom behovet for å etablere en forståelse av reformens effekter og metodiske problemer med å isolere slike effekter.
Samspillet med annen kunnskapsproduksjon	Hvilke andre kunnskapskilder har spilt en rolle? (Hvilken rolle?) Hvordan har den forskningsbaserte evalueringen taklet det faktum at de ikke har monopol på kunnskap?	Noen ganger produseres intern styringsinformasjon og –kunnskap som fører til raskere korrigeringer underveis enn det som en større og langsiktig evaluering kan levere.
Internasjonale forhold vs den særegne norske politiske og kulturelle konteksten	Hvilke trekk ved evalueringsprosessen må forstås ut fra den spesielle norske konteksten? Hva kan ha overføringsverdi til internasjonale forhold?	

Forberedelse og igangsetting

Det har vært uformell kontakt med de tre foreslåtte personene (se eget notat). De er enige i at vi først holder et møte hvor vi kan avklare ambisjonsnivået, forståelse av oppdraget og arbeidsformen. På det grunnlaget kan vi komme fram til realistiske rammer for ressursbehovet – penger og tidsbruk. De mener at det er en god måte å starte på.

Etter møtet kan vi i mandatet legge til opplysninger om roller, arbeidsform og ressursrammer med tilhørende kontrakter.

Vedtak (14. januar 2014): Mandat og organisering av meta-evalueringen av NAV-evalueringen godkjennes. Forskningsrådet fastsetter et endelig mandat for oppdraget etter panelets første møte.

Forskningsrådet oppnevner et ekspertpanel for en meta-evaluering av EVA-NAV med følgende sammensetning:

Professor Morten Balle Hansen, Aalborgs universitet, leder
Førsteamanuensis Anne Halvorsen, Universitetet i Agder
Revisjonsrådgiver Jahn-Eric Furubo, Den svenske riksrevisionen

Forskningsrådet v/Steinar Kristiansen er observatør på panelets møter.

Ekspertpanelets tillegg og presiseringer

Spørsmålet er omfattende, og metaevalueringens tids- og ressursmessige rammer gjør det nødvendig å begrense og tilpasse ambisjonsnivået. Dette avklares under.

Sekretariatsfunksjonen organiseres som et forskningsprosjekt ved Aalborgs universitet. Karen Breidahl Nielsen, Aalborgs universitet, engasjeres som sekretær for panelet.

Det skal holdes 5–6 møter i ekspertpanelet underveis, og de to første møtene er berammet til 3. februar og 3. mars.

Data og empiri

Datagenereringen vil omfatte to typer data: intervjuer og dokumenter.

Det vil bli foretatt 7–10 kvalitative intervjuer, eventuelt supplert med noen telefonintervjuer.

Respondentene vil hentes fra de fire-fem primære interessentene i evalueringen: forskerne, Forskningsrådet, styringsgruppen, Arbeids- og sosialdepartementet og NAV-direktoratet.

Dokumentene som inngår i metaevalueringen er:

- Mål- og rammedokumentet/oppdraget fra Arbeidsdepartementet til Forskningsrådet
- Møtereferater fra styringsgruppen
- Årsrapporter fra evalueringsprosjektet og Forskningsrådet
- Utvalgte publikasjoner fra NAV-evalueringen. Ressursene er ikke tilstrekkelige til å få analysert samtlige publikasjoner, og den nærmere begrensningen vil bli besluttet på første møte.

De 7–10 kvalitative intervjuene vil bli gjennomført av Karen Breidahl Nielsen. Morten Balle Hansen og Anne Hjørdis Halvorsen deltar ved 1–2 intervjuer, enten sammen med Karen, eller alene om praktiske grunner skulle tilsi det. Ved behov suppleres det med et mindre antall telefonintervjuer.

Publikasjon

Både de tre ekspertene og sekretæren vil bli oppført som medforfattere av den endelige rapporten fra metaevalueringen.

Rapportens hovedforfatter vil være sekretæren Karen Breidahl Nielsen i tett samspill med ekspertpanelets leder Morten Balle Hansen, men Anne Hjørdis Halvorsen og Jan-Eric Furubo vil også ha en rolle som medforfattere av rapporten.

Rapporten fra metaevalueringen vil bli skrevet på dansk. De avsnittene som eventuelt skrives av Jan-Eric Furubo og Anne Hjørdis Halvorsen vil bli oversatt til dansk av Karen Breidahl, slik at språket framstår som ensartet.

Etter planen skal rapporten fra metaevalueringen være ferdig i manuskriptform medio juni, og den trykte rapporten skal være klar primo juli.

Metaevalueringens konklusjoner og anbefalinger presenteres på en konferanse i Oslo i midten av juni. Konferansen organiseres av Forskningsrådet.

Utover selve rapporten planlegges en vitenskapelig publikasjon i et evaluerings- eller Public Administration-tidsskrift. Dette vil bidra til å sikre at rapporten fra metaevalueringen og dens anbefalinger er solid forankret i en vitenskapelig tilnærming.

Bilag 3: Mål- og rammedokumentet

Mål og rammedokument for evaluering av NAV-reformen.

1. Bakgrunn

I St.prp. nr.46 (2004-2005) ble det lagt fram forslag til en ny arbeids- og velferdsforvaltning. Stortinget ga sin tilslutning til forslaget ved behandlingen 31. mai 2005. Trygdeetaten og Aetat nedlegges i sin nåværende form, og samles i en ny, statlig etat, Arbeids- og velferdsetaten. Det etableres en førstelinjetjeneste i alle landets kommuner hvor statsetaten og kommunene går sammen i et forpliktende avtalefestet partnerskap om å opprette felles lokale arbeids- og velferdskontor. Kontoret skal ivareta oppgaver for den statlige etaten og som et minimum kommunens oppgaver etter sosialtjenestelovens kapittel 5 (økonomisk sosialhjelp og veiledning/rådgivning i den forbindelse). Etaten og kommunen kan avtale at også andre av kommunens tjenester skal inngå i kontoret. Kommunene skal fortsatt ha en sentral rolle i arbeids- og velferdspolitikken og beholder sitt ansvar for oppgaver etter sosialtjenesteloven. Det lokale NAV-kontoret skal være det samlende kontaktstedet i dagens tredelte forvaltning slik at brukerne får ett sted å henvende seg.

Reformen gjennomføres for bedre å kunne oppnå tre overordnede mål:

- * **Arbeidsretting** Flere i arbeid og aktivitet, færre på stønad
- * **Brukerretting** Enklere for brukerne, og bedre tjenester tilpasset deres behov
- * **Effektivisering** En helhetlig, og effektiv arbeids- og velferdsforvaltning

Som et ledd i en samlet FoU-satsning, er det i St.prp.nr 46 (2004-2005) forutsatt at det parallelt med reformen skal gjennomføres

”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”(jf kap 7 side 79).

Stortinget sluttet seg til dette.

I dette notatet beskrives målsettinger, problemstillinger, mulige framgangsmåter for, og organiseringen av, et evalueringsprogram for NAV-reformen. Det legges til grunn at prosjekter skal starte opp så raskt som faglige og organisatoriske hensyn tillater. Det skal både settes i gang arbeid som følger reformen i hele gjennomføringsperioden (2006-2009) og videre fram til reformen har satt seg, og prosjekter som skal frambringe kunnskap underveis.

Programmet skal være en forskningsmessig evaluering av reformen på et overordnet plan. Parallelt vil det også etableres opplegg mellom arbeids- og velferdsetaten og kommunene for systematisk erfaringsutveksling og læring, særlig knyttet til organisering og utvikling av verktøy og arbeidsmetoder i de felles lokale NAV-kontorene. Disse prosessene skal holdes atskilte. Det er ikke til hinder for at prosjektene under dette programmet inngår samarbeid

med arbeids- og velferdsetaten/kommunene om egne forsøk- og utviklingsprosjekter eller bruker data fra pågående prosjekter i den samlede evalueringen, men arbeids- og velferdsetaten og kommunene vil ha et selvstendig ansvar for evaluering av prosjekter i egen regi.

2. Formål med evalueringen

Det grunnleggende formål med denne evalueringen er å gi svar på om NAV-reformen når de overordnede målene; bedre arbeidsretting, bedre brukerretting og en mer effektiv forvaltning. Evalueringen omfatter både en forskningsbasert evaluering av resultater/effekter i forhold til målene og en prosessevaluering for bl.a. å dokumentere hvordan reformen er blitt gjennomført.

Arbeids- og inkluderingsdepartementet har et overordnet sektoransvar, både på kort og lang sikt, for å definere egne kunnskapsbehov og finansiere forskning innen sitt ansvarsområde. Videre skal alle statlige virksomheter, iflg. Reglementet for økonomistyring i staten (§ 16), sørge for at det gjennomføres evalueringer for å få informasjon om effektivitet, måloppnåelse og resultater. Evalueringene skal belyse hensiktsmessigheten bl.a. av innsatsfaktorer, aktiviteter, produkter og tjenester samt kartlegge effekter i forhold til brukere og samfunn mv.

NAV-reformens store omfang, dens nyskapende institusjonalisering av et tidligere uprøvd forhold mellom stat og kommune, samt de ambisiøse velferds mål som knyttes til den, tilsier at reformen må gjøres til gjenstand for en særskilt evaluering.

NAV-reformen er både en betydelig reorganisering av forvaltningen og en sosialpolitisk reform. Stortinget, regjeringen, departement, underliggende etater, kommunene, brukere og allmennheten vil ha et løpende kunnskapsbehov om lærdommer som kan trekkes underveis og om målene er nådd når reformen er fullt ut gjennomført. Utviklingen i ulike gruppers deltakelse i arbeid, forbruk av trygdeytelser og omfang av sosial utestenging, følges nøye fra mange kanter. Hvorvidt NAV-reformen som *forvaltningsendring* faktisk leder til at det *arbeidsmarkedspolitiske* målet om en større samlet yrkesdeltakelse oppnås bedre i årene som kommer (enn om den gamle fler-etatsmodellen hadde vært videreført) er en viktig tematikk. Måloppnåelse i forhold til de overordnede målene som er nevnt foran er det sentrale i evalueringen av reformen, hvor det er viktig bl.a. å forstå de enkeltelementer og prosesser som fremmer eller hemmer denne måloppnåelsen. Disse elementene/prosessen kan være knyttet til for eksempel sentrale avtaler og styringsdialog. De kan også berøre forhold på lokalt organisasjonsnivå eller arbeidsmetoder i samhandlingen med brukerne.

Dette evalueringsprogrammet skal fokusere på NAV-reformen som en forvaltningsreform og på dens bidrag til å oppnå de overordnede mål. Det ligger utenfor rammen av denne evalueringen å ensidig belyse bruk av de ulike virkemidlene (tjenester, tiltak og ordninger) som arbeids- og velferdsforvaltningen disponerer, uten å relatere de til de overordnede målene og sentrale trekk ved reformen. Den nye forvaltningen etableres i utgangspunktet med de samme sosiale ordninger og virkemidler som trygdeetaten, Aetat og den kommunale sosialtjenesten rår over. Det bør likevel ikke trekkes noe skarpt skille mellom forvaltningspolitiske og sosialpolitiske virkninger.

For det første er reformens overordnede mål om å heve deltakelsen i arbeid og å utvikle mer brukerrettede tjenester, klare velferdspolitiske mål. En viktig ambisjon med reformen er at organisatorisk samordning og effektivisering og nye samarbeidsformer skal lede til bedre, mer effektivt og brukerrettet arbeid med positive virkninger på brukernes velferd. En belysning av

om dette faktisk oppnås, forutsetter faglige opplegg hvor organisatoriske og sosialfaglige perspektiver må kombineres.

For det andre skal reformen også fremme sosial inkludering. En vesentlig andel av sosialhjelpsmottakerne og enkelte andre brukergrupper av NAV-kontorene befinner seg i livssituasjoner hvor de står i fare for å bli sosialt ekskludert og marginalisert. Flere studier viser at en mindre del av brukerne, men samlet sett likevel et høyt antall, har ulike og sammensatte problemer i form av lav/mangelfull inntekt, mangelfull utdanning, rusmiddelbruk, svak helse, kriminalitet og sosial fungering. NAV-reformen skal bidra til å frigjøre ressurser overfor målgrupper som har behov for individuelt tilrettelagt tjenesteløp. Evalueringen bør derfor søke å belyse i hvilken grad reformen bidrar til å fremme sosial inkludering.

For det tredje kan det forventes gjennomført endringer i velferdsordningene, i løpet av omstillingsperioden, både i rettighetssystemer og i ressursbruk, og reformen gjennomføres samtidig med at en ny IA-avtale inngås mellom arbeidslivets parter. Det er en metodisk utfordring å skille mellom de effekter på brukerne som skyldes nye virkemidler (og som ville ha kommet også innen den gamle organisasjonsformen) og de som kan tilbakeføres til NAV-reformen. Kunnskap om virkninger av ulike sosialpolitiske virkemidler er derfor viktig for å kunne identifisere forhold i den observerte utviklingen som med empirisk troverdighet kan tilskrives forvaltningsmessige forhold.

NAV-reformen implementeres over relativt lang tid, og det må forventes at også andre, mer samfunnsmessige forhold vil influere på de observerte mønstrene for blant annet arbeidsmarkeds- og aktivitetsdeltakelse. De økonomiske konjunktorene vil kunne endres. Reformen kan i seg selv påvirke folks holdninger og nye alderskohorter kommer til som brukere. Om forventningene til bruker- og individretting av tjenestene også endres, blir det mer krevende å skille mellom slike faktorer og de mer direkte effektene av selve forvaltningsreformen. Denne utfordringen kan tenkes å være større fordi NAV-reformen er en så omfattende reform, men i prinsippet er ikke de vitenskapelige problemene annerledes enn ved enhver annen evaluering. Mange faktorer vil virke inn, men forvaltningsreformens bidrag til måloppnåelse skal være en helt sentral problemstilling. Det er viktig at evalueringen får fram prosjekter som metodisk og datamessig på en god måte kan belyse betydningen av forvaltningsendring som virkemiddel til å nå de overordnede målene som er satt for reformen, og særlig målet om flere i arbeid og aktivitet, og færre på stønad.

Den innsats som skal igangsettes har fire resultatområder:

1: **Dokumentasjon.** Den skal gi en systematisk oversikt av hva som konkret er gjort og som gjøres i planleggings- og gjennomføringsfasen fra ulike aktører, herunder departementets oppfølging av gjennomføringen og av NAVinterim, samarbeidet med andre involverte parter, bl.a. KS og SH-dir. Hva fungerte bra, hva kunne vært gjort bedre, og hvilke lærdommer kan trekkes? Slik kunnskap er nødvendig for å skape et fundament for styring, planlegging og budsjettering. Dette bør samtidig ses i sammenheng med utforming av et nytt styringssystem for den nye arbeids- og velferdsforvaltningen. Reformen er også ment å gi en effektivisering av arbeids- og velferdsforvaltningen. Evalueringen skal dokumentere om reformen har bidratt til effektiviseringsgevinster. En evaluering skal gi et dypere og bredere grunnlag for etablering av styringsparametre og for en stadig bedre utøvelse av politisk og strategisk styring av reformprosessen.

2: **Korreksjon.** Reformen gjennomføres i faser og over en lengre periode. Den er designet for å muliggjøre erfaring og læring underveis. I hvilken grad dette også skjer i praksis er et sentralt kunnskapsspørsmål. En evaluering kan og bør gi svar på hvordan implementeringsproblemer oppstår og tematiseres, og hvordan ulike aktørers erfaringer omdannes til generaliserte innsikter, og om denne nye kunnskapen spres (raskt) som læringseffekter som i sin tur gir ny input til justeringer og videreutvikling av styringssystemet underveis.

3: **Innsikt om organisatorisk integrasjon.** Den nye arbeids- og velferdsforvaltningen skal etableres gjennom en prosess hvor to selvstendige statlige etater samles i en ny statsetat og etablerer et forpliktende avtalebasert samarbeid med den kommunale sosialtjenesten. Den statlige etaten og kommunene skal ha stor grad av autonomi til å utvikle lokale løsninger, både internt i arbeids- og velferdskontoret og i grensenettet mot brukerne. Reformen åpner bl.a. for bruk av interkommunale løsninger og mulighet for at stat og kommune kan løse oppgaver for hverandre. En evaluering må fokusere på hvordan enkeltelementene i reformen fungerer i forhold til de sentrale målene, både med et styringsperspektiv ovenfra og et brukerperspektiv nedenfra. Det er viktig at det mangfold av løsninger som kan forventes, utnyttes til å etablere kunnskap som kan identifisere vellykkede så vel som mindre gode effekter, og på den måten gi et kunnskapsgrunnlag for iverksetting av relevante tiltak.

4: **Langsiktig kunnskapsutvikling.** Det norske velferdsmodellen er basert på at vi alle før eller siden er brukere av det offentlige rettighetssystem og hjelpeapparat. Gitt den store andelen av samfunnets ressurser og av befolkningen som til enhver tid er involvert, er det et viktig hensyn hele tiden å øke kunnskapen om hvordan institusjonell utforming, ressurshushold og individuell velferd er koblet sammen. Selv om dette ikke kan være et hovedfokus for evalueringen er det viktig at det datamaterialet som samles inn gjennom denne evalueringen, de teoretiske perspektiver som anlegges i de enkelte prosjekter, de forskningsmiljøer som mobiliseres og de studier som evalueringen frambringer, gjennomføres med et bevisst sideblikk til også å etablere et kompetanse- og kunnskapsgrunnlag for annen og videre forskning.

3. utfordringer knyttet til evaluering av reformen – presisering og mulige avgrensninger.

NAV-reformens tre hovedmål er i evalueringssammenheng av noe ulik karakter. Å få flere av dagens stønadsmottakere i arbeid og aktivitet er et overordnet mål for hele samfunnsutviklingen. På individnivå er den bærende ide er at den enkelte borger, i rollen som bruker, skal oppleve å stå ovenfor en organisasjon, og ikke måtte forholde seg til dennes interne organisering. På institusjonsnivå er antakelsen at reformen vil kunne bidra til en mer effektiv ressursutnyttelse. På samfunnsnivå er antakelsen at dersom NAV-reformen når sine mål som forvaltningsreform, vil det kunne bidra til økt sysselsetting og mindre passivt stønadsmottak i yrkesbefolkningen enn om reformen ikke var blitt gjennomført.

NAV-reformen består ikke av et enkelt identifiserbart tiltak eller en bestemt aktivitet, men er en sammensatt og kompleks reform med mange ulike elementer. Den er en organisatorisk reform som skal gjennomføres på ulike nivåer. Den innbefatter også kompetanseutvikling og vil medføre endrede arbeidsmåter, nye rutiner, ny metodikk og nye verktøy mv. Bildet kompliseres ytterligere ved at reformen skal gjennomføres over en relativt lang periode. En forskningsbasert evaluering vil på sin side måtte utgå fra det motsatte utgangspunkt; å

identifisere konkrete, avgrensede spørsmål som gjennom konfrontasjon med empirisk materiale, gir svar på om de mål man hadde er realisert eller (hvorfor) ikke.

NAV-reformens ambisjoner om å bringe bedre samsvar mellom mål, virkemidler og resultater gjennom en forvaltningsreform, kan derfor hensiktsmessig omformes til fem hovedtema. Hver av disse må få en grundig belysning. Den forskning som settes i gang må knyttes til de overordnede målene for reformen og belyse i hvilken grad/på hvilken måte de ulike elementene i reformen som gjøres til gjenstand for forskning kan bidra til å nå disse målene

I Koblinger mellom reformens mål og strukturell utforming

NAV-reformen innebærer at to store statlige organisasjoner skal samles i en etat, som skal etablere et forpliktende samarbeid med det eldste elementet vi har av kommunalt selvstyre, sosialtjenesten. Hensikten er å øke befolkningens tilknytning til arbeid, gjøre tjenesteytingen mer helhetlig og brukerrettet samt å oppnå organisatoriske effektivitetsgevinster. Om noen år må en forskningsbasert evaluering kunne si om NAV-reformen – i det store og hele - var vellykket. Solide innsikter bør da være etablert om og hvorfor det (ikke) gikk slik de sentrale aktørene håpet.

Denne hovedproblemstillingen kan angripes med ulike utgangspunkt, betinget av hvordan forskningen selv finner det hensiktsmessig å balansere forholdet mellom et styringsperspektiv og et brukerperspektiv. Noen eksempler på problemstillinger/temaer som kan belyses her:

* På hvilke måter, og i hvor sterk grad, kan de tre hovedmålene for reformen sies å ha blitt sammenføyet i en felles strategi i den strukturelle utformingen av den nye arbeids- og velferdsforvaltningen.?

* Hvilke eventuelle spenninger og potensielle inkonsistenser eksisterer mellom hovedmålene, og hvordan er slike forhold håndtert og (ikke) blitt nedfelt i organisasjonsutformingen?

* Hvordan har hvert av de tre hovedmålene blitt konkretisert og vektlagt som hensyn ved utformingen av totaliteten og ved intern strukturering og arbeidsdeling innen den nye forvaltningen.?

* I hvilken grad avhjelper NAV-reformen innelåsnings-/ ”kasteball effekter”, og problemer med kostnadsoverveltning? Kan reformen føre til at det oppstår nye ”kasteballs mekanismer” og overveltningsproblemer?

Det er ønskelig at prosjekter som tar utgangspunkt i slike spørsmål også innarbeider en komparativ dimensjon i evalueringen. NAV kan sammenliknes med andre relevante forvaltningsreformer og det vil være nyttig med et sideblikk til utviklingen i for eksempel Finland og Danmark. Også de tyske ’Harz-reformene’ og den britiske modellen for ’JobCentre-plus’ framstår som relevante i et komparativt perspektiv.

II Forholdet mellom de ulike organisatoriske komponentene i reformen

Den nye arbeids- og velferdsforvaltningen er en organisasjonsmodell som på den ene siden utvikler sin egen indre struktur, og på den andre siden etablerer en styringskjede fra demokratisk valgte myndigheter ned til den enkeltes borgers velferd. I sum, og i kombinasjon,

skal de ulike elementene resultere i *helhetlighet, effektivitet og brukerrettet* tjenesteyting. Disse målene bør derfor, om reformen lykkes, kunne observeres som framvoksende karakteristika ved NAV etter hvert som reformen implementeres. Den andre hovedproblemstillingen som ønskes belyst er hvordan den interne organisering og forholdet mellom de indre komponentene i hele reformen struktureres og i praksis fungerer, - og hvordan dette kan relateres til de observerte grader av måloppnåelse.

Denne tematikken kan gripes med utgangspunkt i tre dimensjoner. Den første angår forholdet mellom brukernær oppgaveløsning i det lokale NAV-kontoret og organisering av oppgaveløsning, kompetanse og ressurser som ikke trenger fysisk nærhet til bruker og som lokaliseres andre steder i den statlige etaten. Tanken er at en slik organisering samlet sett skal gi bedre brukerretting, en mer effektiv ressursbruk og bedre kvalitet i oppgaveløsningen. For det andre vil det innen den statlige etaten utvikles en mer eller mindre hierarkisk autoritetskjede fra øverste politiske myndighet og ned til det lokale nivå i hver kommune. Også kommunene har en styringslinje inn mot det lokale NAV-kontoret. Og for det tredje, det lokale NAV-kontoret skal organiseres som et (kontraktsbasert) partnerskap mellom stat og kommunens sosialtjeneste, og med høy grad av lokal autonomi, bla. annet med hensyn til hvor mye av den øvrige kommunale tjenesteytingen som skal innlemmes i NAV-kontoret..

Det er viktig at forskningsinnsatsen fanger opp alle disse tre dimensjonene, samtidig som den må være åpen slik at den muliggjør et økt fokus på de dimensjoner som kan vise seg å bli viktigere for hele reformens utvikling etter hvert som tiden går, - og på måter det ikke er mulig å forutsi i dag. Særlig er det viktig at forholdet mellom statsetaten og kommunenes sosialtjeneste vies et særskilt fokus. De lokale partnerskapet mellom disse to reguleres bl.a. gjennom krav til samlokalisering, krav om minimumsløsning mht hvilke kommunale oppgaver som skal legges til kontoret og krav om inngåelse av lokale forpliktende avtaler. Innenfor denne rammen legger reformen opp til at man lokalt kan velge egne løsninger bl.a. når det gjelder ledelsesmodell og oppgaver utover minimumsløsningen som kan legges til kontoret. Dette kan dermed skape styringsmessige utfordringer, bl.a. at staten må velge andre styringsteknikker overfor de tjenester i NAV-kontoret som tilligger kommunene enn det som er mulig i forhold til de tjenester som den selv har fullt ansvar for. Ut fra den lokale friheten som reformen åpner for må det forventes en stor variasjon i hvordan like oppgaver løses ulikt. Denne (antatte) variasjonen bør overvåkes, systematiseres og analyseres med sikte på å frambringe et grunnlag for erfaringsspredning og læring. Noen eksempler på problemstillinger/temaer som kan belyses her:

* Hva karakteriserer styringssystemet som er (blir) etablert mellom Arbeids- og Inkluderingsdepartementet, et nytt direktorat og de lokale NAV-kontorene? Sikrer det nye styringssystemet en mer helhetlig styring og samordning av innsatsen, både når det gjelder departementets styring av den nye etaten og direktoratets styring av etaten, herunder koordinering med ulike parter, bl.a. Sosial- og helsedirektoratet og fylkesmannens rolle overfor den kommunale sosialtjenesten?

* Reformen introduserer bruk av avtaler på ulike nivåer – en rammeavtale mellom AID og KS og lokale avtaler mellom den statlige etaten og den enkelte kommune. Hvor egnet er denne type avtaler som virkemiddel i gjennomføringen av reformen og til å nå reformens mål?

* Av hvem og på hvilke måter utnyttes friheten til å utforme lokale løsninger for NAV-kontoret, og i hvilken grad kan det observeres at noen løsninger er bedre enn andre for eksempel i forhold til samhandlingen med den statlige etatens mer sentraliserte produksjons-

/spesialenheter og i forhold til brukerne? Er interkommunale løsninger et vellykket grep? Hvordan påvirkes NAV-kontorets funksjonalitet og resultater av hvorvidt stat og kommune løser oppgaver for hverandre? Hva er effekten av ulike ledelsesvalg lokalt?

* Hvilke betydning i forhold til måloppnåelsen har variasjoner i tilstedeværelse av statlig tjenestemeny (bl.a. oppfølgingstjenester) på det enkelte NAV-kontor, og at statlige tjenester for en del småkommuner må utløses fra andre og større enheter i etaten? Hvordan påvirker kommunestørrelse og kommuneportefølje som inngår i arbeids- og velferdsforvaltningen NAV-kontorets styrbarhet, samarbeidsrelasjoner og resultater?

* I hvilken grad oppstår det spenninger mellom den statlige etaten og den kommunale sosialtjenesten i de lokale NAV-kontorene, hva bunner de i, hvordan håndteres de og i hvilken grad har de konsekvenser for arbeidsmåter, brukerbehandling og måloppnåelse?

* I hvilken grad utvikles det nye gråsoner av uavklarte ansvarsområder mellom NAV-kontoret og de deler av kommunal (og velferdsstatlig) tjenesteyting som ikke innlemmes i NAV-kontoret (for eksempel bolig- og innvandringsspørsmål, helse- og omsorgstjenester)?

* I hvilken grad karakteriseres den gradvise implementeringen av reformen av en administrativ, mer eller mindre instruert koordinering ovenfra, i forhold til innslaget av erfaringsutveksling, kopiering av og læring fra, gode og dårlige erfaringer?

NAV-reformens mål om å etablere et likeverdig forhold mellom stat og kommune på det lokale nivået, er en nyskapning i vår velferdspolitik. Den store variasjonen kommunene i mellom i befolkningssammensetning og problemomfang, den høye grad av autonomi som gis i å utforme det lokale NAV-kontoret, og det forhold at reformen innføres gradvis gjennom pilotkommuner, skaper gode betingelser både for å utvikle spennende forskningsdesign så vel som at det vil framkomme et unikt empirisk materiale for læring og korreksjon. Det er naturlig at den forskningssatsing som settes i gang gjennom denne evalueringen også tar et ansvar for å samle inn og systematisere en databank over lokale løsninger, og at en slik databank danner et empirisk grunnlag både for en løpende (årlig) rapportering og som grunnlagsmateriale for den mer endelige, langsiktige evalueringen.

III Demografi, kultur, kompetanse og profesjon

NAV-reformen bygger på en antakelse om at organisatorisk samordning som dette er en viktig betingelse for at virkemidlene som arbeids- og velferdsforvaltningen rår over kan ses, og virke, i sammenheng. Sentrale mål som helhetlighet, brukerretting og effektivitet hviler på at dette også skjer i praksis, - at de ansatte i NAV-kontoret også opplever det slik og faktisk utøver sin virksomhet ovenfor brukerne på nye måter. Samtidig er NAV-reformen mye mer enn en reorganisering av formelle strukturer og ansvarsforhold. Den er også en fusjon av ulike organisasjonskulturer, ulike kompetanser, fag, profesjonelle identiteter og tilvante arbeidsmåter. Hver av de gamle etatene og sosialtjenesten kommer med sine ansattes demografi, kunnskap, ferdigheter og holdninger, og det kan forventes at menneskelige og organisasjonssosiologiske forhold som normalt virker inn på reorganiseringer både i offentlig og privat sektor vil ha betydning for måloppnåelsen. Kjennetegn ved det personell man starter med innebærer både begrensninger og, gjennom rokeringer og nye samhandlingsrelasjoner, muligheter for økt effektivitet og måloppnåelse. Spenninger kan lamme, men ved klok

håndtering, også utløse entusiasme og ny energi. Evalueringen må legge opp slik at denne dimensjon ved reformen blir systematisk belyst.

Det er naturlig at forskningsinnsatsen innrettes slik at den kan fange opp hvorledes den enkelte medarbeider så vel som de ulike interesseorganisasjonene opplever effekter av reformen, både for egen arbeidssituasjon så vel som i forhold til brukerne. Reformens ambisjon om å virkeliggjøre et helhetlig, rasjonelt og effektivt samarbeid mellom ulike funksjoner og fagkompetanser må, for å kunne evalueres på dette punktet, i betydelig grad basere seg på data fra medarbeiderne selv. Hvordan dette skal gjøres, og hvilke aspekter ved de ansattes oppfatninger av arbeidsinnhold, kompetansebehov, oppgaver og måloppnåelse som skal belyses, kan med fordel overlates til forskningen selv. Noen eksempler på problemstillinger/temaer som kan belyses her:

* I hvilken grad utvikles felles verdier og holdninger i den nye organisasjonen? Hvilke kulturelle spenninger ligger i den nye arbeids- og velferdsforvaltningen, hvordan utvikles de over tid, og på hvilke måter legger dette føringer på graden av måloppnåelse? I hvilken grad vil slike spenninger forsterkes av at et kommunalt og et statlig arbeidsgiveransvar og avtaleverk vil være i funksjon samtidig innen en felles organisasjon?

* I hvilken grad utvikler NAV-reformen nye relasjoner mellom tradisjonelle faggrupper og profesjonelle identiteter, hvilke strategier nyttes for fostring av nye organisasjonsverdier og normer, - og hvilke virkninger har dette eventuelt på medarbeidernes oppfatninger og deres profesjonelle praksis?

* I hvilken grad utvikles det en bevisst rekrutterings- og kompetansepolitikk, og hvilke oppfatninger har nøkkelaktører om forholdet mellom ulike typer av personell, fagkompetanse, arbeidsformer og måloppnåelse?

Den nye arbeids- og velferdsforvaltningen vil ha ansvar for en omfattende personlig tjenesteyting. Den mellommenneskelige relasjon som karakteriserer møtet mellom bruker og forvaltning/tjenesteyter er et vesentlig kvalitetsaspekt ved selve tjenesteytingen. Det er derfor viktig at vitenskapelig opplegg som utformes for å belyse de ansattes opplevelser av reformen knyttes opp til den fjerde hovedproblemstillingen, nemlig brukernes perspektiv.

IV Brukernes perspektiv

Bedre tjenester til den enkelte borger er en sentral begrunnelse for offentlig ressursbruk. Brukernes behov skal settes i sentrum når den nye arbeids- og velferdsforvaltningen utformes. For å styrke arbeidslinja er det viktig at alle ytelser og tjenester som er innrettet mot personer som står i fare for å falle ut av arbeidslivet, eller som står utenfor arbeidslivet, ses i sammenheng. Rask avklaring av behov og riktig virkemiddelbruk er viktig for å sikre inkludering i arbeidslivet. Å sette brukernes behov i sentrum betyr et offentlig apparat hvor organisering, arbeidsmåter, rutiner, ledelse, personalpolitikk og administrativ kultur preges av respekt for brukeren, og gir brukeren mulighet til medvirkning og innflytelse på sin egen sak. Å tilpasse tjenestene til brukerens behov dreier seg derfor både om hvordan man møter brukeren og hva man kan gi brukeren. Det handler om service, respekt, informasjon, tilgjengelighet, kompetanse, diskresjon og et godt tilpasset tjenestetilbud. Det er derfor viktig at ulike brukergruppers erfaringer med den nye etaten følges løpende. Det foreligger allerede i dag data om brukerholdninger fra etatene. Det er naturlig at dette arbeidet systematiseres og

følges opp, bl.a. gjennom å kontrastere brukerholdninger/-erfaringer mellom gammel og ny forvaltningsform og at det etableres en mer rutinisert overvåkning av brukeoppfatninger. Dermed kan det etter hvert som reformen implementeres etableres et instrument som kan utnyttes til å studere sammenhenger mellom ulike lokale organiseringer av NAV-kontoret og brukernes opplevelser av å møte en helhetlig, brukerrettet tjeneste som skal gi rask avklaring av, og hjelp til å løse, deres behov. I en periode framover vil det i noen kommuner være etablert lokale NAV-kontor, mens andre kommuner vil ha den gamle organiseringen i funksjon. Dette skaper gode forutsetninger for sammenlikninger, og for tilbakeføring av erfaringer fra pilotkommunene.

Det er ikke alltid slik at brukernes subjektive oppfatninger er en god indikator på faglig kvalitet. Data om holdninger må suppleres med informasjon om faktiske resultater i form av endrede livssituasjoner for brukeren, - ut fra hva som erfaringsmessig er realistiske mål for godt faglig arbeid for ulike brukergrupper og problemkategorier. Likeså er det en viktig tidsdimensjon ved brukerholdninger. Brukere kan bli både mer negative eller mer positive i sin vurdering etter hvert som de også kan erfare virkningene av den tjenesteyting de mottok. Utviklingen av et evalueringsdesign for å belyse NAV-reformen fra et brukerperspektiv er en krevende oppgave, samtidig som det er viktig at evalueringen etablerer innsikt om hvordan brukerholdninger er (eller ikke er) knyttet systematisk sammen med trekk ved den nye arbeids- og velferdsforvaltningen. Det er et sentralt mål for evalueringen å identifisere hvilke institusjonelle trekk ved denne forvaltningen som er årsaksmessig forbundet med god brukerretting. Noen eksempler på problemstillinger/temaer som kan belyses her:

- * Hvilke trekk ved arbeids- og velferdsforvaltningen (og særlig NAV-kontorets) genererer hhv positive eller negative brukererfaringer, kontrollert for karakteristika ved ulike brukere?
- * Er det systematiske forskjeller i brukertilfredshet/-erfaringer mellom brukere som har enkle behov i forhold til dem som trenger mer omfattende bistand, og hvordan utvikler dette seg over tid?
- * I hvilken grad har brukermedvirkning på systemnivå innvirkning på utformingen av tjenestetilbudene? I hvilken grad bidrar brukermedvirkning på individnivå til myndiggjøring?
- * Kan det finnes motstridende brukerinteresser som i praksis begrenser muligheten for hva forvaltningen/tjenesteyterene oppfatter som godt, brukerrettet arbeid, - for eksempel mellom personvern hensyn og muligheten til koordinering av mange samvirkende tiltak?
- * Den nye arbeids- og velferdsforvaltningen vil satse offensivt på utvikling av de elektroniske kanalene, og spesielt de som legger til rette for selvbetjening over internett. Hvordan påvirkes ulike brukergrupper av dette? Hvilke forutsetninger har ulike brukergrupper for å nyttiggjøre seg de nye kanalene?
- * Hvilke spenninger kan tenkes å eksistere mellom reformens overordnede mål dersom et konsekvent brukerperspektiv legges til grunn? Hvordan er for eksempel forholdet mellom effektivitet og helhetlighet som mål for forvaltningsreformen og hensynet til individualisert brukerretting og rask avklaring av hjelpebehov som mål i tjenesteytingen?

V Den praktiske dimensjon

Å gjennomføre en omfattende reform innebærer å overkomme en lang rekke praktiske spørsmål. Betydningen av hvordan slike problemer løses kan legge viktige føringer på hvordan en reform til slutt faktisk blir gjennomført og om, og i hvilken grad, den når sine mål. En samordnet førstelinjetjeneste, og utvikling av en felles IT-plattform, oppfattes som viktig både for organisatorisk effektivitet og for økt brukervennlighet. Det er en sentral problemstilling å undersøke om slike gevinster faktisk oppnås. Samtidig omfatter NAV – reformen mer enn etablering av NAV-kontor. Hvorledes oppgaveløsning og ressurser organiseres bak NAV-kontoret - for eksempel i produksjons- og spesialenheter - kan være viktig for oppnåelsen både av reformens hovedmål og for effektivisering av forvaltningen. Det er derfor viktig at evalueringen fanger opp den geografiske og den teknologiske dimensjonen ved reformen. Noen eksempler på problemstillinger/temaer som kan belyses her:

* Hvilke betydning har geografisk lokalisering og innplassering av ulike funksjoner innen den statlige arbeids- og velferdsetaten på ulike nivåer, for å oppfylle hovedmålene for reformen?

* Hvordan løses problemene med samordning av informasjon fra de gamle til den nye etaten, og hvilke betydning har etablering av en felles IKT-plattform for nedbygging av gamle sektorbarrierer og utvikling av en mer effektiv og helhetlig informasjonsflyt og arbeidsform.

* Hvordan tematiseres og løses praktiske problemer som oppstår underveis, og i hvilken grad, og gjennom hvilke kanaler formidles praktiske erfaringer tilbake som korreksjon av den videre implementering av hele reformen?

* Et overordnet mål med NAV-reformen er effektivisering. Overføring av ressurser fra administrasjon til tjenesteproduksjon er et sentralt mål for å sikre at de interne ressurser utnyttes mer effektivt, blant annet gjennom bedre utnyttelse og utvikling av de ansattes kompetanse, endringer i rutiner mv. I hvilken grad bidrar reformen til å nå målet om en mer effektiv ressursutnyttelse? I hvilken grad klarer den nye etaten å utnytte potensialet i de store ressursene som ligger i den nye organisasjonen?

* I hvilken grad har fysiske løsninger (herunder fysisk tilgjengelighet, universal utforming, logistikk og kundeflyt) betydning i forhold til å oppfylle målene i reformen?

Ved store reformer er det ikke alltid de problemene som forutses som i praksis viser seg å skape størst hindre for god måloppnåelse, - av den enkle grunn at dem har aktørene stor bevissthet omkring. Det er de uventede og overraskende hendelser som kan ha størst betydning. Det er derfor viktig at evalueringen legges opp slik at den raskt vil kunne fange opp betydningen av faktorer vi i dag ikke kjenner.

Nærmere om indikatorer på måloppnåelse

Et forskningsbasert evalueringsprosjekt vil ha et stort behov for konkrete operasjonaliseringer av gode indikatorer for graden av oppnåelse av reformens hovedmål. Det finnes allerede en stor mengde data både fra de gamle etatene, fra den ordinære statistikkproduksjon og fra mer spesiallagede koblinger av data i regi av Statistisk Sentralbyrå. Samtidig arbeides det med, som ledd i gjennomføringen av reformen, å utvikle statistikk og nye data som kan gi løpende tilbakeføring av informasjon om den nye forvaltningens virksomhet.

I utgangspunktet må en forskningsbasert evaluering selv konstruere og velge ut hvilke indikatorer den mener er best egnet for å belyse de spørsmål den stiller. Samtidig er det store forskningsøkonomiske gevinster å hente ved å basere de vitenskapelige analysene på data som allerede foreligger eller vil foreligge ved at de er samlet inn for andre formål.

Som forberedelse til evalueringen bør det snarest utarbeides en samlet oversikt over hvilke data av relevans som eksisterer/vil bli produsert. Oversikten bør være summarisk og på overskriftsnivå gi opplysninger om type statistikk og hvem som produserer dataene. Denne oversikten bør foreligge innen sommeren 2006, og gjøres tilgjengelig i forbindelse med den første utlysningen av prosjekter. Sentrale dataprodusenter (NAV-interim, RTV, Aetat Arbeidsdirektoratet, SSB) kan gå sammen og lage en slik oversikt.

4 Nærmere om innretningen av evalueringsprosjektet

Den evaluering som skal settes i gang skal dels følge, dels konkludere om, en forvaltningsreform som gjennomføres i perioden 2006-2009. Den nye statlige arbeids- og velferdsetaten skal etter planen etableres fra andre halvår 2006 og utover. Etableringen av lokale arbeids- og velferdskontor vil skje gradvis fra ultimo 2006 og fram mot utgangen av 2009. De nye kontorene skal fylles med delvis ny kompetanse og nye arbeidsmåter mv. som kan komme til å strekke seg utover perioden for nye kontoretableringer.

Evalueringen skal starte opp så snart som praktisk mulig og i løpet av andre halvår 2006. Det vil være vanskelig å få fram solid empirisk belagte forskningsinnsikter i den tidlige fasen av evalueringen. Den forskningsbaserte evalueringen vil følgelig ha en relativt lang tidshorisont og bør strekke seg over 6-8 år.

Stortinget har lagt vekt på at evalueringen skal gjennomføres på en slik måte at det kommer *'kontinuerlige impulser og kritiske blikk utenfra under reformgjennomføringen'* ut fra et ønske om *'å trekke lærdommer underveis'* (St.prp. nr. 46 (2004-2005) s.79).

Kravet om at det skal kunne trekkes lærdommer underveis innebærer at forskningsinnsatsen må struktureres i minst to ulike prosjektfaser; en fase hvor formålet er å få fram resultater knyttet til områdene *dokumentasjon* og *korreksjon* og en annen fase hvor det særlig vektlegges å vinne innsikt om effektene av NAV-reformen som velferdsstatlig forvaltningsreform ivaretas. Det er viktig at det gis tidlig tilbakemelding om resultater knyttet til målene om dokumentasjon og korreksjon.

Det gode prosjektet klarer å etablere en faglig fruktbar og forskningsøkonomisk effektiv arbeidsdeling mellom disse to aktivitetene. Det er naturlig å se for seg en prosjektfase der en samler og systematiserer data, som så kommenteres og regelmessig rapporteres til de relevante brukere av denne kunnskapen. I første fase bør etableringsprosessen særlig i forbindelse med NAV-kontorene følges nøye, likeså utviklingen av holdningsmønstre hos ansatte og sentrale målgrupper. Likeså bør det raskt etableres en databank over de lokale partnerskapsavtalene som løpende oppdateres gjennom hele evalueringens løpetid. Rask formidling til oppdragsgiver og andre sentrale aktører vil være et viktig mål for denne aktiviteten.

De data som samles inn gjennom den første prosjektfasen og de problemstillinger og nye kunnskapsspørsmål som utkrystalliserer seg, vil være viktige input til den andre prosjektfasen.

Her bør forventningene om rapportering være mindre, slik at forskningsaktiviteten i større grad kan innrette seg på en grundigere, endelig rapportering til vel etter at reformen er innført i sin helhet.

Evalueringen må være åpen, slik at prosjekter og problemstillinger kan endres underveis: Nye innsikter vil gi opphav til mer spennende og relevante problemstillinger, reformens innhold og tidsskjema kan endres og det kan komme endringer i virkemiddelapparatet som gjør det nødvendig å endre forskningsdesignet. Det er viktig å ta høyde for at reformens utvikling og resultatene som evalueringsprosjekter frambringer kan og bør påvirke innretning og fokus for videre evalueringsarbeid. Evalueringen skal være en dynamisk prosess. Evalueringen kan for eksempel vise at problemstillinger man startet opp med ikke lenger er relevante, at enkelte problemstillinger er vanskelig eller i praksis umulig å følge opp og at nye spørsmål dukker opp. Også annen forskning og resultater av forsøksvirksomhet bør påvirke den samlede evalueringens vitenskapelige progresjon. Gjennomførings- og etableringsfasen vil samtidig være en utviklingsfase der enkelte sider ved reformen kan endres og tilpasses underveis. Det må derfor, og i samråd med oppdragsgiver, legges opp til en kontinuerlig revisjon av nøkkelspørsmål og problemstillinger, for eksempel en gang årlig.

For at en samlet evalueringsinnsats best mulig skal kunne møte de kunnskapsbehov som skisseres i dette evalueringsprogrammet, er det nødvendig å integrere spisskompetanse både fra forvaltnings- og velferdsforskning. Det er ikke ønskelig å spre evalueringen ut på en rekke mindre enkeltprosjekter og forskningsinstitusjoner. Sett på bakgrunn av reformens omfang og kompleksitet og behovet for forskjellige typer av forskningskompetanse, vil det være mest hensiktsmessig at forskningsmiljøer går sammen, avtaler en arbeidsdeling seg i mellom, med én ansvarlig prosjektledelse, og som leverer ett tilbud på hele evalueringen.

Prosjektbeskrivelsen må, med utgangspunkt i de fem hovedproblemstillingene som er reist, og ut fra de kunnskaps- og formidlingsbehov som er signalisert, redegjøre for hvordan prosjektet/forskergruppen vil løse fire oppgaver:

1. En omfattende forskningsbasert evaluering. Dette innebærer en evaluering av resultater/effekter av reformen, basert på forskningsmessig metodikk og kompetanse, for rapportering etter at NAV-reformen er fullt ut innført.
2. En evaluering der den organisatoriske og praktiske gjennomføringen av reformen dokumenteres, herunder oppbygging av en databank over de ulike lokale løsninger og de erfaringer som løpende gjøres her. Det må presenteres en særskilt plan for hvordan den løpende rapporteringen ovenfor oppdragsgiver er tenkt organisert og en sannsynliggjøring av hvordan dette vil være et verdifullt kunnskapsgrunnlag for de aktører som har ansvaret for reformens implementering.
3. Redegjøre for hvordan man vil benytte data som legges til grunn vil bli frambrakt uavhengig av prosjektet, hvilke egne rutiner som vil bli etablert for innsamling av andre data, samt presentere opplegg for eventuelle større (survey-)undersøkelser som planlegges gjennomført.
4. Utarbeide et opplegg for
 - a) Formidling og kunnskapsspredning som ivaretar behovet for en bredt anlagt og mangfoldig erfaringsspredning, for eksempel gjennom publikasjoner, nettbaserte løsninger og workshops/seminarer rettet mot relevante grupper.
 - b) Systematisk tilbakeføring av resultater fra evalueringen til beslutningstakere og andre sentrale aktører, slik at det blir mulig å forbedre tiltak og prosesser underveis. Oppdragstaker må redegjøre for hvordan oppgaven under a) og b) tenkes løst.

5. Organisering av evalueringsarbeidet

Arbeids- og inkluderingsdepartementet vil, bl.a. i kraft av sitt sektoransvar, ha det overordnede ansvaret for evalueringen. Dette innebærer et finansieringsansvar og et hovedansvar for utforming av innhold og innretning i evalueringen. Norges Forskningsråd har ansvaret for den praktiske organiseringen av evalueringen, og skal ivareta den faglige og administrative kontakten med forskningsmiljøet som skal utføre evalueringen. Oppdraget er gitt til NFR for å gi evalueringssatsingen en faglig kvalitet, integritet og uavhengighet, og for å kunne trekke på forskningsrådets rutiner og kompetanse i håndtering av forskningsprosjekter, løpende kontakt med forskerne og organisering av brukerrettet formidling fra prosjekter.

Departementet betrakter evalueringen av NAV-reformen som et separat oppdrag som skal holdes utenfor den ordinære programstrukturen hvor AID i dag finansierer den langsiktige forskningen. Grunnen til dette er at evalueringen av NAV-reformen krever en tettere styringsdialog mellom departement og forskningsrådet enn det som kan pålegges et ordinært forskningsprogram.


Samarbeidet med Forskningsrådet som konkretiseres gjennom dette mål- og rammedokumentet skal ha følgende innhold:

- 1) Det opprettes en styringsgruppe for evalueringen som er i funksjon i hele prosjektperioden. Styringsgruppen består av én person fra hhv. AID, KS, Arbeids- og velferdsetaten og SH-dir, samt tre personer oppnevnt av Forskningsrådet. Ved oppnevning av disse tre skal Forskningsrådet vektlegge hensynet til faglig uavhengig kompetanse, kjennskap til forvaltnings- og velferdsforskning og kompetanse i forskningsformidling ovenfor forvaltning og andre brukere.
- 2) Styringsgruppen har, og med tilgang til Forskningsrådet's ressurser og ordinære prosedyrer, ansvar for at det utarbeides et dokument som konkretiserer departementets kunnskapsbehov til en utlysningstekst, hvor forskningsmiljøer inviteres til konkurranse om å levere tilbud på hele evalueringen.
- 3) Styringsgruppen vurderer innkomne forslag og tildeler oppdraget. Når oppdrag er tildelt, tiltrer forskningsledelse styringsgruppen med møteplikt og forslagsrett. Det utarbeides i plan for evalueringen, som konkretiserer:
 - a) En tidsplan for når ulike enkeltprosjekter skal levere hva, og hvor det skilles klart mellom forskningsprosjekter med lang løpetid og de mer prosessorienterte evalueringene med løpende rapportering.
 - b) En rutine for (årlige) gjennomganger av rapportering, kunnskapsstatus og nye forskningsutfordringer.
 - c) Et opplegg for formidling og kunnskapsspredning bl.a. gjennom nettbaserte løsninger.

Planen godkjennes av AID, mens Forskningsrådet gis ansvaret for gjennomføring.

4) Det opprettes et forum for relevante brukere og interessenter, som møtes årlig i forbindelse med oppgavene spesifisert i punkt 3b.

5) Styringsgruppen har ansvaret for at de målsettinger, delrapporteringer og tidsplan for evalueringen som avtales mellom AID og Forskningsrådet nedfelles i kontrakten mellom utførende forskningsinstitusjon og Norges forskningsråd.



Publikasjonen kan lastes ned fra
www.forskningsradet.no/publikasjoner

Norges forskningsråd

Drammensveien 288
Postboks 564
1327 Lysaker

Telefon: 22 03 70 00
Telefaks: 22 03 70 01
post@forskningsradet.no
www.forskningsradet.no

Omslagsdesign: Design et cetera AS
Trykk: 07 Gruppen AS
Opplag: 200

Oslo, august 2014

ISBN 978-82-12-03357-3 (trykksak)
ISBN 978-82-12-03358-0 (pdf)