



Evaluering af virksomhedsovervågningen

Evaluering af første runde af undersøgelsen af virksomhedernes forebyggende arbejdsmiljøaktiviteter som led i overvågning af fremdriften i handlingsplanen for et rent arbejdsmiljø i år 2005

af

Olaf Rieper
Leena Eskelinen

akf forlaget
januar 2005

Forord

Denne rapport indeholder resultaterne af evalueringen af virksomheds-
overvågningen (VOV1). Evalueringen er en del af den anden runde af
virksomhedsovervågningen (VOV2), som blev igangsat efteråret 2003.
Virksomhedsovervågningerne har til formål at beskrive og analysere ud-
viklingen i de forebyggende aktiviteter på virksomhederne på de syv visi-
onsområder, som er indeholdt i regeringens handlingsprogram for rent ar-
bejdsmiljø år 2005.

Formålet med evaluering af VOV1 er at sikre tilpasning af design og
metoder til VOV2 samt at vurdere, i hvilken grad VOV1 opfyldte sit for-
mål.

Evaluatoren (akf) er en del af konsortiet for VOV2 med ansvaret for
evalueringen. Konsortiet består i øvrigt af COWI og Teknologisk Institut,
og det ledes af COWI.

Vi vil gerne rette en tak alle dem, der har medvirket med oplysninger
og viden til evalueringen og stillet sig til rådighed for interview.

Evalueringen er blevet gennemført af Olaf Rieper (docent i akf og
projektansvarlig) og Leena Eskelinen (seniorforsker i akf). Hanne-
Pernille Stax (lektor ved Institut for Almen og Anvendt Sprogvidenskab,
KU) og Anders Holm (lektor ved Sociologisk Institut, KU) har bidraget
med særlig sagkundskab og udarbejdet bilag. Camilla Teil Nielsen (stu-
dentermedarbejder i akf) har refereret interview og møder.

Evalueringen er finansieret af Arbejdstilsynet.

Olaf Rieper
Januar 2005

Indhold

Resumé	7
1 Sammenfatning og konklusion	9
1.1 Evalueringens genstand og metode	9
1.2 Design og metode i VOV1	10
1.3 Spørgeskemaer og vurderingsskemaer i VOV1	12
1.4 Den praktiske dataindsamlingsmetode i VOV1	14
1.5 Vurdering af visionsrapporter	14
1.6 Konklusion	15
2 Evalueringens genstand og metoder	19
2.1 Første runde af overvågningen.....	19
2.2 Metoder og fremgangsmåder i evalueringen	22
3 Design og metode i VOV1	26
3.1 Udtræk af arbejdssteder til telefoninterview	26
3.2 Filterspørgsmål	33
3.3 Interview med såvel A- som B-side.....	35
3.4 Interview med sikkerhedsrepræsentant versus anden medarbejder....	38
3.5 Validering og supplering af telefoninterview ved hjælp af arbejdsstedsbesøg	40
4 Spørgeskemaer og vurderingsskemaer i VOV1	47
4.1 Temaernes relevans for udviklingen samt kvantificering/	

kvalificering af besøgsdata	47
4.2 Fordelingen af generelle og visionsspecifikke spørgsmål	54
4.3 Spørgsmålenes (i telefoninterviewene) relevans for de udvalgte temaer	
4.4 Telefoninterviewspørgsmålenes sproglige udformning	61
5 Den praktiske dataindsamlingsmetode i VOV1	64
5.1 Praktisk tilrettelæggelse af dataindsamling	64
5.2 Variation mellem besøgs konsulenter ved opgørelserne fra arbejdsstedsbesøgene	72
6 Vurdering af visionsrapporter	75
7 Forslag til ændringer i design og metoder med henblik på VOV2	83
Bilag	
1 Oversigt over skriftligt materiale anvendt i evalueringen	87
2 Præcisering af evalueringsspørgsmål	89
3 Kriterier for vurdering af visionsrapporter, af spørgeskemaer til telefoninterview og af vurderingsskemaer til arbejdsstedsbesøg	95
4 Oversigt over de gennemførte interview	99
5 Interviewguider anvendt i evalueringen	101
6 Vurdering af dataudtrækket til VOV1 og anbefalinger til det videre arbejde. Udarbejdet af Anders Holm, Sociologisk Institut, KU	117
7 Sammenfatning af sproglig evaluering af spørgeskemaer til VOV1...	129
English Summary	147
Noter	149

Resumé

Denne rapport indeholder resultaterne af evalueringen af den første virksomhedsovervågning (VOV1), som Arbejdstilsynet igangsatte i 2000 som led i regeringens handlingsprogram for rent arbejdsmiljø år 2005. Virksomhedsovervågningerne har til formål at beskrive og analysere udviklingen i de forebyggende aktiviteter på virksomhederne på syv visionsområder, som er indeholdt i handlingsprogrammet, herunder indeklime, kemi, støj, tunge løft og ensidigt, gentaget arbejde, arbejdsulykker, det psykiske arbejdsmiljø og børn og unges arbejde.

Den første virksomhedsundersøgelse blev gennemført i 2000-2002 i tre dele, som omhandlede henholdsvis udvikling og validering af en metode til overvågning af virksomhedernes forebyggende aktiviteter samt telefoninterview til ca. 4.600 virksomheder (både ledelses- og medarbejderside), og arbejdsstedsbesøg på ca. 10% af de virksomheder, der blev telefoninterviewet. Som den første virksomhedsundersøgelse skulle VOV1 endvidere tjene som basis (statusmåling) for virksomhedernes forebyggende aktiviteter.

Formålet med evaluering af VOV1 er at sikre tilpasning af design og metoder til den næste runde af undersøgelsen (VOV2) samt at vurdere, i hvilken grad VOV1 opfyldte sit formål. Evalueringen er en del af den anden runde af virksomhedsovervågningen (VOV2), som blev igangsat efteråret 2003. Evalueringen bygger på kollegial vurdering af visionsrapporter, inklusive dataindsamlingsmetoder samt interview med bestiller og udførere af undersøgelsen.

Evaluatoren vurderer, at det grundlæggende design med telefoninterview på et stort antal arbejdssteder suppleret med arbejdsstedsbesøg er

hensigtsmæssigt. Der skal imidlertid foretages en række justeringer og udvikling af udtræksprocedurer, spørgeskemaer og vurderingsskemaer. Dette vil i sagens natur begrænse muligheden for sammenligning mellem VOV1 og fremtidige virksomhedsovervågninger.

Med hensyn til, i hvilken grad VOV1 levede op til sit formål, er det evaluatorens vurdering, at dette kun skete delvis. Som begrundelse for denne vurdering henleder evaluator opmærksomheden på mangler ved dokumentation af datagrundlaget og det faktiske datamateriale, metodernes egnethed til at måle forebyggende aktiviteter, gennemførelse af dataindsamlingen og mangler ved analyserne af datamaterialet på de syv visningsområder.

VOV1 udgør en platform og et erfaringsgrundlag for de fremtidige virksomhedsovervågninger, men mulighederne for at bruge VOV1 som basis (reference) for fremtidige undersøgelser af virksomhedernes forebyggende aktiviteter vurderes at være begrænsede. Der er således behov for en afklaring og revision af en række design- og metodemæssige aspekter ved VOV1 med henblik på at udvikle pålidelige og gyldige metoder, som kan anvendes i fremtidige undersøgelser af fremdriften i arbejdsmiljøarbejdet i virksomhederne.

1 Sammenfatning og konklusion

1.1 Evalueringens genstand og metode

Evalueringen retter sig mod den første virksomhedsovervågning (VOV1), som Arbejdstilsynet igangsatte i 2000. Evalueringen er en del af den anden runde af virksomhedsovervågningen (VOV2), som blev iværksat efteråret 2003. Virksomhedsovervågningerne har til formål at beskrive og analysere udviklingen i de forebyggende aktiviteter på virksomhederne på de syv visionsområder, som er indeholdt i regeringens handlingsprogram for rent arbejdsmiljø år 2005.

VOV1's datagrundlag er telefoninterview med såvel A- (ledelse) som B-side (medarbejderside) på ca. 3.200 virksomheder og besøg ved BST-konsulenter på ca. 320 virksomheder. Der er, som rapportering fra VOV1, udarbejdet syv visionsrapporter og en generel rapport.

Formålet med evaluering af VOV1 er at sikre tilpasning af design og metoder til VOV2 samt at vurdere, i hvilken grad VOV1 opfyldte sit formål. Det ligger uden for evalueringens formål at foretage en procesevaluering herunder at afdække de vilkår, som VOV1 blev gennemført under.

Evalueringen bygger på følgende elementer:

- En faglig (kollegial) vurdering af syv visionsrapporter og rapporten om det generelle arbejdsmiljøarbejde, spørgeskemaer til telefoninterview og vurderingsskemaer anvendt ved arbejdsstedsbesøg, herunder er der anvendt dels en sproglig konsulent, dels en statistisk konsulent.
- Interview med bestiller i Arbejdstilsynet (Kontor for Effekt og Analyse), med Overvågningsudvalget under Arbejdsmiljørådet, med de visi-

onsansvarlige i Arbejdstilsynet og med kontrahenten for VOV1 ved Arbejds miljøinstituttet (AMI) og Center for Alternativ Samfundsanalyse (CASA), samt med kvalitetskontrollører fra kontrahenten.

- Interview med udvalgte telefoninterviewere og besøgs konsulenter, koordinatoren for telefoninterviewdelen og projektlederen for besøgsdelen.

Evalueringen fandt sted fra primo marts til medio juli 2003. Udgangspunktet for evalueringen, hvis resultater sammenfattes i dette kapitel, er en række evalueringsspørgsmål, som er gengivet i bilag 2.

1.2 Design og metode i VOV1

Udtræk

Datamaterialets dækning af de udvalgte branchegrupper kan ikke vurderes, idet der ikke forelå bortfaldsanalyser i de rapporter, der forelå ved evalueringens start. Den lave deltagelse i flere branchegrupper inden for de enkelte visionsområder og især for små arbejdssteder gør det tvivlsomt, om datamaterialet giver et dækkende billede af visionsområderne. Uden bortfaldsanalyse, der kan sandsynliggøre, om generalisering er mulig, er det usikkert at foretage sammenligninger for disse branchegrupper vedkommende med en ny stikprøve for VOV2.

Evaluators foreslår, at man undlader at stille højere styrkekrav end 50%, som følger signifikansniveauet på 5%. Der vil i så fald skulle opnås en samlet stikprøve på 180-400 arbejdssteder i alt for VOV1 og VOV2, dvs. der kan opereres med 4-5 strata i hver vision med uændret krav om mindst 50 arbejdssteder i hver branchegruppe.

Den lave svarprocent (under 60) i nogle branchegrupper er en trussel mod at kunne generalisere til visionen og dermed for at foretage sammenligning over tid. Det foreslås derfor at gennemføre bortfaldsanalyser for disse branchegrupper vedkommende for at sandsynliggøre, om det er muligt at generalisere eller ej.

Det foreslås, at man opstiller en statistisk model for ændringerne på de enkelte parametre, fx en logistisk regressionsmodel. Modellen kunne indeholde forklarende faktorer for sammenhængen mellem de studerede

variabler på den ene side og branchegruppe og arbejdsstedsstørrelse på den anden. På denne måde kunne man opnå et klarere overblik over eventuelle systematiske effekter fra branchegruppe og arbejdsstedsstørrelse. Den samlede statistiske model vil kunne give information om, hvorvidt der er selvstændige effekter fra arbejdsstedsstørrelse og branchegruppe, men samtidig også information om, hvorledes den specifikke effekt af arbejdsstedsstørrelse påvirker arbejdsmiljøarbejdet inden for hver vision.

Filterspørgsmål

Filterspørgsmål beregnet på at udelade arbejdssteder, er ikke blevet anvendt af telefoninterviewerne til at undlade interview. Hvilke konsekvenser, det har for generaliserbarhed, afhænger af, hvor mange arbejdssteder det drejer sig om, og om de afviger systematisk fra de øvrige arbejdssteder i branchegruppen. Rapporterne indeholder ingen oplysninger om disse forhold.

Der er spørgsmål, der ufrivilligt kommer til at tjene som et filter for, hvad svarpersonerne mener, fx begrebet »kemikalier« eller »børn og unge i fritidsjob«.

Interview med både A- og B-side

Både A- og B-siden bør indgå i undersøgelsen i lige omfang, dels af arbejdsmarkedspolitiske grunde, dels fordi de to parter kan forventes at have forskellig viden om og forskellig indfaldsvinkel til arbejdsmiljøet. Generelt mener evaluator, at uoverensstemmende svar fra A- og B-siden bør kunne udnyttes bedre. Analyserne af forskelle mellem svar fra henholdsvis A- og B-siden bør foretages mere systematisk og dybdegående, forskelle bør diskuteres og mulige forklaringer gives.

Interview med sikkerhedsrepræsentant versus anden medarbejder

Der har i udbudsfasen været forskellige opfattelser af, om man burde vælge ansatte uden tillidshverv eller en sikkerhedsrepræsentant til interview. Evaluator vurderer, at valget af en sikkerhedsrepræsentant, hvor en sådan findes, er forsvarligt. Det er uklart, hvordan og efter hvilke kriterier sikkerhedsrepræsentanten blev udpeget til interview.

Validering og supplering af telefoninterview med arbejdsstedsbesøg

Evaluators mener, at det metodisk set er problematisk at validere telefoninterview med arbejdsstedsbesøg på grundlag af eksisterende data fra VOV1 primært fordi de to dataindsamlinger ikke var designet til dette formål. Dels indeholder telefoninterviewene på nogle visionsområder få spørgsmål, som er egnet til at blive valideret, dels er spørgsmålene ikke stillet på samme måde ved interview og besøg. Derimod er det i højere grad muligt at bruge besøgene fra VOV1 til at supplere og uddybe de problemstillinger, der er belyst gennem telefoninterviewene. Dette er imidlertid sket i meget begrænset omfang.

1.3 Spørgeskemaer og vurderingsskemaer i VOV1

Temaernes relevans

Temavalget på visionsområderne følger Arbejdstilsynets forslag til temaer i udbudsmaterialet, men der kan sættes spørgsmålstegn ved, om temaerne er operationaliseret på en sådan måde, at de afspejler visionspecifikke forebyggende aktiviteter. Vægten synes i højere grad dels at ligge på det generelle arbejdsmiljøarbejde, dels på kortlægning af arbejdsmiljøforhold, herunder eksponering/risikofaktorer, selv om det sidstnævnte kun undtagelsesvis var ønsket i AT's udbudsmateriale.

Kvantificering/kvalificering af besøgsdata

Med hensyn til forholdet mellem henholdsvis kvantitative og kvalitative vurderinger, der foretages ved arbejdsstedsbesøgene, mangler visionsrapporterne en argumentation for at anvende både en samlet kvantitativ og kvalitativ opgørelse samt en diskussion af, hvordan disse to skalaer forholder sig til hinanden. Den samlede kvantitative skala anvendes ikke i rapporterne; i enkelte rapporter anvendes dog enkeltspørgsmål fra vurderingsskemaet, som for de fleste konsulents vedkommende danner grundlag for den kvantitative vurdering. For den kvalitative vurdering findes ikke en tilsvarende måleparameter i telefoninterviewene. Derfor er

den ikke egnet til at blive brugt til validering; derimod kan den bruges uddybende.

Generelle og visionsspecifikke spørgsmål

De generelle (fælles) spørgsmål udgør over halvdelen af telefoninterviewene, hvilket ikke er i overensstemmelse med, at undersøgelsens formål var at sætte fokus på visionsspecifikke forhold på de syv udvalgte områder (jf. Arbejdstilsynets udbudsmateriale). Problemet forstærkes yderligere ved, at de fælles og de visionsspecifikke spørgsmål næsten ikke er blevet sammenholdt med hinanden i analysefasen. Evaluators vurdering er, at de generelle spørgsmål med meget fokus på den formelle organisering af arbejdsmiljøarbejdet og lidt fokus på forebyggende aktiviteter har en høj prioritet.

Spørgmålenes relevans

Spørgmålenes (i telefoninterviewene) relevans for de udvalgte temaer vurderes forskelligt af forskellige aktører, hvilket bl.a. kan skyldes, at de har forskelligt vurderingsgrundlag afhængigt af, om de lægger vægt på de forebyggende aktiviteter eller arbejdsmiljøets tilstand på de enkelte visionsområder. Måling af forebyggende aktiviteter har hele tiden været den overordnede målsætning, men der er sket en afklaring af prioriteringer i undersøgelsen undervejs i projektet. Således har opdeling i viden, holdning og handling ikke været specificeret i udbudsmaterialet (viden og holdning er dog nævnt på nogle visionsområder).

Evaluator vurderer, at der ikke i tilstrækkelig grad er målt forebyggende aktiviteter. Set i forhold til opdeling af spørgsmålene i viden, holdning og handling er områderne viden og holdning som forventeligt underbelyst. Desuden er det kun for en delmængde af spørgsmålene muligt at måle udvikling over tid.

Spørgmålenes sproglige udformning

Der er en kritik af flere spørgsmål og deres mulighed for at give gyldige svar i forhold til, hvad spørgsmålene skulle belyse, jf. bilag 7. Det gælder især for spørgeskemaerne til visionsområderne børn og unge, kemi, ulyk-

ker og støj; støjskemaet vurderes af sprogkonsulenten til at være materialets svageste.

1.4 **Den praktiske dataindsamlingsmetode i VOV1**

Tilrettelæggelse af dataindsamlingen

Der konstateres problemer med undersøgelsens gennemførelse, herunder mangler ved planlægning, koordinering og styring af hele projektet samt de to delprojekter. Der er tale om forhold, som kan forventes at have betydning for datakvaliteten. Angående telefoninterviewene handler det bl.a. om udtrækket, og om de rigtige arbejdssteder og arbejdsstedsrepræsentanter blev interviewet (mangler ved virksomhedsoplysninger) og om B-sidens deltagelse (systematisk lavere end A-sidens). Angående arbejdsstedsbesøg handler det bl.a. om manglende systematik og ensartethed med hensyn til, hvem der var respondenterne, og hvordan metoderne blev brugt.

Variation mellem besøgs konsulenter

Der er især variation mellem forskellige konsulenter kvalitativ vurdering ved arbejdsstedsbesøgene. Konsulenterne har foretaget vurderingerne forskelligt, og de har brugt skalaerne forskelligt. Variation er ikke blevet analyseret af kontrahenten. Sikring af større ensartethed/minimering af observatørvariation har ikke været prioriteret i projektet, fx har kollegial sparring mellem konsulenter og projektet været meget sparsom.

1.5 **Vurdering af visionsrapporter**

Vurderingen af de syv rapporter samt den generelle rapport peger på styrker og svagheder ved rapporterne. Styrker er, at de enkelte rapporter konkluderer retvisende i forhold til de præsenterede analyseresultater, men der mangler gennemgående forbehold med hensyn til, om udtrækket og datamaterialet giver et dækkende billede.

Rapporterne giver viden om nogle forhold, der ikke tidligere har været belyst så omfattende, og det er positivt, at Arbejdstilsynet har kunnet udnytte dele af datagrundlaget til overvågningsrapport 2001.

Evaluator har hæftet sig ved følgende svagheder. For det første er der mangel på ensartethed i analyserne på tværs af rapporterne, selv om deres formelle opbygning er tilnærmelsesvis parallel. Det vanskeliggør en læsning på tværs, og dermed er det svært at få et samlet overblik over resultaterne. For det andet mener vi, at givet at det ene formål med VOV1 var at udvikle og afprøve design og metoder, er der begrænset vurdering af de anvendte metoder og ej heller megen refleksion og diskussion af metodernes egnethed til at gentages senere. For det tredje er rapporterne generelt ikke lette at læse, heller ikke for en målgruppe som professionelle i arbejdsmiljøsystemet. De ujævne analyser og sine steder manglende omtale af og dermed manglende forbehold over for metodiske svagheder i rapporterne gør, at evaluator må sætte spørgsmålstegn ved, om rapporterne gennemgående lever op til den forskningsmæssige kvalitet, man må stille til rapporteringen af et så stort projekt. Der gennemførtes en vis overordnet kvalitetskontrol af udkast til rapporter, men evaluator vurderer, at det var en brist i kvalitetskontrollen, at den normale vurdering af rapporterne ved eksterne lektører ikke blev foretaget.

1.6 Konklusion

Det ene formål med evalueringen er at vurdere, om design og metoder anvendt i første runde (VOV1) bør revideres i anden runde (VOV2) med mulige konsekvenser for sammenligning mellem VOV1 og VOV2.

Evaluator vurderer, at det grundlæggende design med telefoninterview på et stort antal arbejdssteder suppleret med arbejdsstedsbesøg er hensigtsmæssigt. Der skal imidlertid foretages en række justeringer og udvikling af udtræksprocedurer, spørgeskemaer og vurderingsskemaer. Dette vil i sagens natur begrænse muligheden for sammenligning mellem VOV1 og VOV2.

For en detaljeret listning af evaluators forslag til VOV2 henvises til kapitel 7 om forslag til ændringer i design og metoder med henblik på VOV2. Her vil vi pege på de forslag, vi vurderer som de væsentligste.

Design og metoder:

1. Såfremt der ønskes sammenligning over tid af forholdene i enkelte branchegrupper, bør udtrækket af arbejdssteder forøges væsentligt, eller man bør renoncere på de statistiske krav til, på hvilket niveau man vil have afdækket forskelle. Alternativt kan man sammenligne en given vision (med branchegrupperne slået sammen) over tid. Det vil give et større datagrundlag for at gennemføre en regressionsanalyse, der kan afdække fx branchegruppeeffekt, tidseffekt eller effekt af arbejdsstedsstørrelse.
2. Overensstemmelse mellem A- og B-siden bør analyseres på en gennemskuelig måde, fx ved korrelationsanalyser. Lav overensstemmelse i svarene bør fortolkes og ikke ubetinget tages som udtryk for usikre besvarelser (lav gyldighed).
3. Besøgsdata bør primært anvendes til validering af telefoninterview for faktuelle spørgsmål. Den kvalitative del af besøgsdata bør i højere grad anvendes til at supplere telefoninterview og integreres med resultaterne fra telefoninterview i rapporterne.

Spørgeskemaer og vurderingsskemaer:

1. Udgangspunktet for revision af spørgsmål til telefoninterview bør være evaluators vurdering af, hvilke spørgsmål der er egnet til sammenligning over tid, jf. bilag 8.
2. Nye spørgsmål bør udvikles under hensyntagen til, a) at der mangler spørgsmål til at indikere arbejdsmiljøets tilstand (så aktivitetsniveauet på arbejdsstederne kan sættes i forhold til omfanget af deres problemer), b) at der mangler spørgsmål om forebyggende aktiviteter og om holdninger og viden. Endelig c) bør de formelle spørgsmål blandt de fælles spørgsmål reduceres.
3. Flere af de spørgsmål, der genbruges fra VOV1, bør justeres efter forslag fra sprogkonsulenten, jf. bilag 7.

Tilrettelæggelse af dataindsamling:

1. Der bør gennemføres en stram og konsekvent projektledelse og klar koordinering på tværs af visioner og datakilder således, at der opnås en fælles forståelse af metoderne, at tidsplanen overholdes, og at resultaterne analyseres og rapporteres ensartet.

2. Kriterier for valg af arbejdssteder til undersøgelsen og eventuel eksklusion bør afklares.
3. Virksomhedsoplysninger bør tjekkes, herunder om der er overensstemmelse mellem branchegruppe og adresseoplysninger ved telefoninterviewene. Det bør ligeledes afklares, hvem der er de rette lokale virksomhedsrepræsentanter, og sikres, at disse bliver interviewet.
4. Der bør ske en afklaring af, hvordan besøgene gennemføres/bruges i store virksomheder, da besøgene kun omfatter en del af virksomheden (vedr. spørgsmålet om validering og supplering i forhold til telefoninterviewene).

Visionsrapporter:

1. Der bør være en klar analysestrategi og detaljeret disposition for samtlige rapporter, inden analyse og skrivning påbegyndes.
2. De enkelte visionsrapporter bør være ensartet opbygget med klar påpegning af metodemæssige forbehold. Rapporterne bør indeholde en metodebeskrivelse og dokumentation af datamaterialet.
3. Det bør sikres, at der finder en ekstern kvalitetssikring af rapportudkast sted.

Det andet formål med evalueringen er at vurdere, i hvilken grad VOV1 levede op til sit formål. VOV1's formål var dels at udvikle og afprøve en metode til at overvåge fremdriften i arbejdsmiljøarbejdet i virksomhederne, dels at give et overblik over udviklingen i forebyggende aktiviteter på virksomhederne for hver vision og dermed tjene som basis (reference) for fremtidige undersøgelser af virksomhedernes forebyggende aktiviteter.

Evaluators vurderer, at VOV1 kun delvis har levet op til sit formål.

For det første er datamaterialet og datagrundlaget ikke dokumenteret i rapporterne, og derfor er det ikke muligt at vurdere disses kvalitet fyldestgørende. Den lave deltagelse (andel i udtrækket og svarprocent) i flere branchegrupper inden for enkelte visionsområder og for små virksomheder bidrager yderligere til at gøre det tvivlsomt, om datamaterialet giver et dækkende billede af branchegrupperne og visionsområderne. Dertil kommer, at det faktiske datamateriale i undersøgelsen er karakteriseret ved, at det er individbaseret og afspejler et ønske om generaliserbarhed af

resultaterne i forhold til beskæftigelse i de udvalgte branchegrupper. Det står i modsætning til, at bestilleren ønskede et udtræk, som er virksomhedsbaseret, og hvor såvel store, mellemstore som små virksomheder er ligeligt repræsenteret.

For det andet foreligger der kun sparsom beskrivelse og vurdering af de anvendte metoder og ej heller megen refleksion og diskussion af metodernes egnethed til at måle forebyggende aktiviteter (i henhold til de temaer, der var nævnt i Arbejdstilsynets udbudsmateriale) og af metodernes egnethed til at gentages senere. Endvidere er der mangler i en del af spørgsmålenes formulering, som er en trussel mod gyldigheden og pålideligheden i dataindsamlingen.

For det tredje er der konstateret manglende systematik ved planlægningen og gennemførelsen af dataindsamlingen både ved telefoninterviewene og arbejdsstedsbesøgene, hvilket kan forventes at have betydning for datakvaliteten.

For det fjerde er der en mangel på analyser, som er centrale for, at undersøgelsen kan bruges som reference for senere undersøgelser af virksomhedernes forebyggende aktiviteter på visionsområderne.

Af ovennævnte grunde må evaluator sætte spørgsmålstejn ved, om undersøgelsen gennemgående lever op til den forskningsmæssige kvalitet, man må stille til gennemførelsen af et så stort projekt.

Evaluator har noteret sig, at VOV1 udgør en platform og et erfaringsgrundlag, som, trods de kritiske vurderinger, vil muliggøre en bedre VOV2, end det ellers ville have været tilfældet.

2 **Evalueringens genstand og metoder**

2.1 **Første runde af overvågningen**

Evalueringens genstand er den første runde af undersøgelsen af virksomhedernes forebyggende arbejdsmiljøaktiviteter som led i overvågningen af femdriften i Handlingsprogram for rent arbejdsmiljø år 2005, som blev fremlagt af regeringen i 1996. I handlingsprogrammet opstilles de overordnede mål for den forebyggende arbejdsmiljøindsats frem til år 2005 indeholdende syv visioner om et rent arbejdsmiljø. De syv visioner omfatter formindskelse af følgende arbejdsmiljøproblemer: ulykker, udsættelse for kemiske stoffer, skader hos børn og unge, tunge løft eller ensidigt, gentaget arbejde, psykisk arbejdsmiljø, skader på grund af dårligt indeklima, og støj. Programmet udgør således en overordnet ramme for arbejdsmiljøets aktører for en 10-årig periode.

Kort efter fremlæggelsen af handlingsprogrammet nedsatte Arbejdsmiljørådet en arbejdsgruppe, der fik til opgave at overvåge fremdriften i de syv visioner. Arbejdsgruppen skulle udvikle metoder til overvågning af programmet. Arbejdstilsynet har fungeret som sekretariat til Arbejdsmiljørådets såkaldte overvågningsudvalg.

Overvågningsprogrammet omfatter tre søjler: 1) en overvågning af risikofaktorer og helbredsforhold/-skader på grundlag af eksisterende datakilder, 2) en overvågning af fremdriften i de forebyggende aktiviteter på virksomhederne, samt 3) en overvågning af de øvrige arbejdsmiljøaktørers forebyggende aktiviteter (jf. figur 2.1). Det overordnede formål med årligt at gennemføre analyser af en række indikatorer på udviklingen in-

den for de nævnte felter er at danne grundlag for eventuelle justeringer i de forebyggende aktiviteter på arbejdsmiljøområdet.

Overvågningen i den første søjle har været etableret længe før Handlingsprogrammets vedtagelse. Den bygger først og fremmest på Arbejdstilsynets registre over anmeldte arbejdsulykker og anmeldte arbejdsbetingede lidelser samt Arbejdsmiljøinstituttets nationale arbejdsmiljøkohorte (NAK), tidligere kaldt Lønmodtagerundersøgelsen.

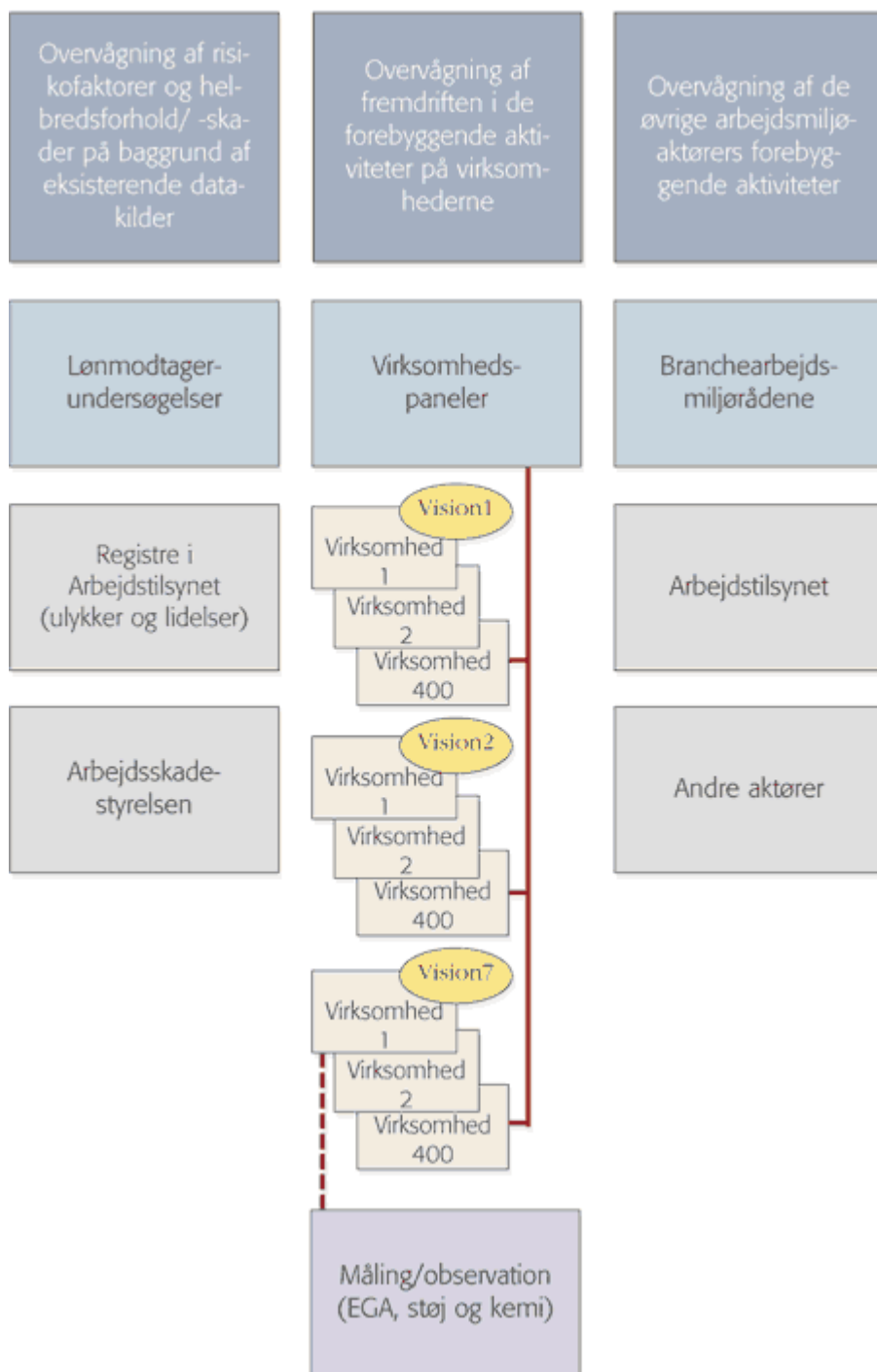
Overvågningen i den anden søjle begyndte i 2001 med den såkaldte virksomhedsovervågning (VOV1). Den anden runde af overvågning (VOV2) skal igangsættes i 2003, og det er hensigten at gennemføre endnu en i 2006 (VOV3).

Overvågningen i den tredje søjle omfatter for tiden Arbejdstilsynets aktiviteter samt branchegruppearbejdsmiljørådenes (BAR) aktiviteter.

Formålet med VOV1 er at udvikle og afprøve en metode til at overvåge fremdriften i arbejdsmiljøarbejdet på virksomhederne for de syv visioner, at analysere de indsamlede data samt at foretage rapportering af resultaterne. Sigtet er, at overvågningen skal tjene som basis (reference) for fremtidige undersøgelser af virksomhedernes forebyggende arbejdsmiljøindsatser samt give et grundlag for, at arbejdsmiljøets aktører kan justere deres indsats, hvis resultaterne peger i en forkert retning (Udbud af overvågning af fremdriften i 2005-handlingsplanen, 2000, s. 4).

VOV1 blev udbudt af Arbejdstilsynet i tre dele: en del som omhandler udvikling og validering af en metode til overvågning, analyse og rapportering, en anden del som vedrører dataindsamling ved telefoninterview til ca. 4.600 virksomheder (både A- og B-side), og en tredje del som vedrører dataindsamling ved arbejdsstedsbesøg på ca. 10% virksomheder, der blev telefoninterviewet.

Figur 2.1 De tre søjler i overvågningen af fremdriften i de syv visioner i »Handlingsprogrammet for et rent arbejdsmiljø år 2005« ifølge Arbejdstilsynet. Figuren afspejler, at der i forbindelse med første undersøgelsesrunde var tanker fremme om, at der også skulle foretages målinger i arbejdsmiljøet under arbejdsstedsbesøgene. Dette blev dog ikke realiseret. (Fra: <http://www.at.dk/Overblik/information/overvaag/model.gif> den 28. okt. 2002.)



Arbejdsmiljøinstituttet og CASA varetog udvikling og validering af en metode til overvågning samt analyse og rapportering. SFI Survey og Teknologisk Institut gennemførte dataindsamlingen ved henholdsvis telefoninterview og arbejdsstedsbesøg. Der blev indgået separate kontrakter mellem Kontor for Effekt og Analyse i Arbejdstilsynet og de fire kontrahenter. AMI og CASA udgjorde styregruppen for undersøgelsen, som det fremgår af forord til visionsrapporterne, hvor det også nævnes, at AMI og CASA, suppleret med underentreprenører, har gennemført undersøgelsen.

Opgaven startede sommeren 2000 og var planlagt til afslutning med en sammenfattende rapport marts 2002 (Kontrakt mellem Arbejdstilsynet og Arbejdsmiljøinstituttet, 2000). I realiteten var rapporteringen ikke endeligt afsluttet ved denne evalueringens start marts 2003. Der var således tale om en betydelig forsinkelse i forhold til den oprindelige tidsplan. Det ligger uden for evalueringens formål at afdække grundene hertil, men der vil punktvis i denne rapport blive angivet nogle formodede grunde.

2.2 **Metoder og fremgangsmåder i evalueringen**

Udgangspunktet for evalueringen af VOV1

akf, som er en del af konsortiet for at gennemføre VOV2, har ansvaret for at gennemføre evalueringen af VOV1. Evalueringen startede primo marts 2003 og afsluttes med denne rapport i september 2003.

Formålet med evalueringen er at vurdere det undersøgelsesdesign og de metoder, der blev anvendt i VOV1, således at man på den baggrund kan tilrettelægge anden undersøgelsesrunde på den mest hensigtsmæssige måde. Et andet, ligestillet formål er at vurdere, i hvilken grad første undersøgelsesrunde lever op til dens formål.

Resultaterne af evalueringen skal først og fremmest anvendes af de personer, der har ansvaret for anden undersøgelsesrunde, dvs. Arbejdstilsynet og Arbejdsmiljørådets overvågningsudvalg samt ikke mindst projektledelsen af VOV2.

Evalueringen vurderer, om det tidligere anvendte design og metoder er hensigtsmæssige, og vil pege på, hvilke forbedringer der bør foretages i det tidligere anvendte design og metoder, herunder de anvendte temaer, under hensyntagen til, at Arbejdstilsynets formål med overvågningen er at

registrere udvikling over tid, og at et eventuelt evalueringsresultat, der peger på gennemgribende ændringer i design og metoder, må tage dette hensyn i betragtning. Evalueringen er foretaget på grundlag af det i bilag 1 nævnte skriftlige materiale samt interview (jf. nedenfor), dvs. der er hverken foretaget reanalyser af datamaterialet eller vurdering af den konkrete fremgangsmåde ved dataudtrækket i forhold til stratificering og vægtning. Evalueringen vil pege på og begrunde, hvilke ændringer i design og metoder der er mest nødvendige samt redegøre for mulighederne for at måle udviklingen fra første til anden runde. Det endelige valg af ændringer foretages af Arbejdstilsynet og Arbejds miljørådets overvågningsudvalg.

De skriftlige materialer, som evalueringen bygger på, fremgår af bilag 1. Rapporter er medtaget i de versioner, hvori de forelå primo marts 2003.

Evaluator har udarbejdet tre notater til intern brug i Kontor for Effekt og Analyse (til september 2002 Kontor for Forskning og Udvikling) i Arbejdstilsynet og til projektledelsen af VOV2.

Fastlæggelse af evalueringsspørgsmål

Grundlaget for de spørgsmål, som er udgangspunktet for evalueringen, er udbudsmaterialet for VOV2, to møder med Kontor for Effekt og Analyse i Arbejdstilsynet, samt en høring hos Kontor for Effekt og Analyse i Arbejdstilsynet, kontrahenten ved AMI og CASA, samt leverandørerne i VOV1, dvs. Teknologisk Institut og SFI Survey. De således fastlagte evalueringsspørgsmål fremgår af bilag 2.

Faglig vurdering ved akf og særligt sagkyndige

Evaluators i akf har foretaget en faglig vurdering (såkaldt kollegial vurdering eller peer review) af de syv visionsrapporter og den generelle rapport, der forelå ved evalueringens start primo marts 2003. Endvidere har evaluators vurderet de syv spørgeskemaer anvendt til telefoninterview og de syv vurderingsskemaer anvendt ved arbejdsstedsbesøgene.

Som særligt sagkyndige har evaluators haft en statistisk konsulent til at vurdere procedurer for udtræk af virksomheder, samt en sprogkonsulent til at granske de syv spørgeskemaer anvendt til telefoninterview. Begge sagkyndige er kommet med forslag til VOV2, og disse forslag fremgår af bilagene 6 og 7.

Der er gennemført interview dels med personer, der var involveret i udbud og design af VOV1, samt med kontrahenten ved AMI og CASA, dels med personer, der indsamlede data.

Interview med personer, der var involveret i udbud og design af VOV1, samt med ledelsen hos kontrahenten af VOV1

Der blev gennemført gruppeinterview med bestiller, dvs. henholdsvis partsmedlemmer af overvågningsudvalget og medarbejdere i Kontor for Effekt og Analyse i Arbejdstilsynet. I Arbejdstilsynet blev yderligere de visionsansvarlige interviewet enkeltvis. Endvidere blev de to ledere hos kontrahenten af VOV1 interviewet, den ene ansigt til ansigt, den anden via telefon. Endelig blev de ansvarlige for den overordnede kvalitetskontrol kontaktet. Den ene svarede pr. mail, den anden blev interviewet personligt.

Oversigt over de gennemførte interview fremgår af bilag 4. De anvendte interviewguider fremgår af bilag 5.

Gruppeinterviewene blev optaget på lydbånd, og under enkeltpersoninterviewene blev der taget notater. Der foreligger referater fra alle interview til intern brug hos evaluator i akf.

Interview af besøgs konsulenter og telefoninterviewere

Der blev gennemført enkeltpersoninterview med henholdsvis telefoninterviewere hos SFI Survey, besøgs konsulenter fra BST'erne, COWI og TI, som gennemførte arbejdsstedsbesøg samt interview med projektlederen af besøgsdelen på Teknologisk Institut og med koordinatoren af telefoninterviewdelen hos SFI Survey. I bilag 4 gives en oversigt over de gennemførte interview.

Interviewpersoner blandt besøgs konsulenterne blev udvalgt således, at de repræsenterer de syv visionsområder i undersøgelsen. Der blev interviewet to besøgs konsulenter og en telefoninterviewer pr. visionsområde. Besøgs konsulenterne blev så vidt muligt udvalgt blandt dem, der er med i VOV2. Telefoninterviewere blev udvalgt tilfældigt blandt dem med de fleste interview på et visionsområde. Interviewene blev dels gennemført personligt, dels telefonisk.

Evalueringsspørgsmål i kursiv

Rapporten er struktureret efter de evalueringsspørgsmål, der behandles, jf. bilag 2. Evalueringsspørgsmålene er i de følgende kapitler gengivet med kursiv og deres nummering følger nummeringen i bilag 2.

3 Design og metode i VOV1

3.1 Udtræk af arbejdssteder til telefoninterview

Formålet med udtræk af en stikprøve af arbejdssteder var at etablere en første måling af forebyggende arbejdsmiljøaktiviteter på ca. 400 arbejdssteder pr. vision inden for branchegrupper, der er i fokus for arbejdsmiljøapparatets indsatser. Det skulle endvidere tilstræbes at dække hver branchegruppe med 50 arbejdssteder. Kravet var endvidere, at stikprøven var stratificeret efter arbejdsstedsstørrelse, således at man ved efterfølgende målinger kunne udtale sig sikkert om ændringer på små, mellemstore og store arbejdssteder.

- a) Det skal vurderes, om kravet om mindst 50 repræsentanter i en branchegruppe giver et brugbart materiale til senere dataopgørelser.*
- b) Det skal ligeledes vurderes, om fordelingen på virksomhedsstørrelse i små, mellemstore og store virksomheder bør være balanceret, forstået på den måde, at de er repræsenteret med hver en tredjedel.*

Faglig vurdering ved statistisk konsulent, jf. bilag 6

Dette afsnit bygger i hovedsagen på den indgåede kontrakt mellem Arbejdstilsynet og Arbejdsmiljøinstituttet, men også øvrige dokumenter indgår, herunder udbudsmaterialet til 2001-undersøgelsen.

Ifølge denne kontrakt ønskes et udtræk uden systematiske fejl. Der er aftalt et udtræk fra NAK-registret, hvorved der opnås repræsentation efter beskæftigelsesandele (dvs. virksomheder efter antal ansatte). I det oprindelige udbudsmateriale var der ønsket et udtræk fra CVR-registret (se næste afsnit), men dette krav er efterfølgende blevet fraveget til fordel for et

udtræk fra NAK. Der ønskes 400 virksomheder inden for hvert af de syv visionsområder og en stratificering efter virksomhedsstørrelse (små, mellemstore og store arbejdssteder). Det sidste sker i forbindelse med udførelsen af projektet.

Data kan illustreres af nedenstående tabel:

Branche-gruppe	Visionsområde 1			Visionsområde 7
Branche-gruppe1	0-4 ansatte	5-19 ansatte	Mere end 19 ansatte		
.....					
Branche-gruppe K					
I alt	400			400

Når der i tabellen vises strata (celler) på arbejdsstedsstørrelser, er det fordi der i processen omkring udtrækket er blevet diskuteret en yderligere stratificering på arbejdsstedsstørrelser. Stikprøven skal altså stratificeres på virksomhedsstørrelse (og ikke nødvendigvis på branchegrupper) inden for hvert visionsområde.

Der ønskes at kunne påvise ændringer på 10-15 procentpoint (absolutte ændringer) på et 5%’s signifikansniveau med en statistisk styrke på 80% (dvs. at der er stillet et højere styrkekrav end 50%, som følger signifikansniveauet på 5%) mellem visioner over tid, altså mellem den allerede foretagne undersøgelse (VOV1) og kommende undersøgelser. Kravet om at kunne påvise en forskel på 10-15% på et 5%’s signifikansniveau med en statistisk styrke på 50% medfører en samlet stikprøve i de to strata, der sammenlignes på mellem 180 og 400 observationer (virksomheder). Kravet om at kunne påvise en forskel på 10-15% med en statistisk styrke på 80% medfører en samlet stikprøve i de to strata, der sammenlignes, på mellem 400 og 820 observationer. Hensynet til statistisk styrke bestemmer med andre ord stikprøvestørrelsen.

Hvis der i hvert visionsområde udtages 400 arbejdssteder, kan der altså ikke stratificeres, hverken på branchegrupper eller på arbejdsstedsstørrelser, når hvert stratum (celle i tabellen ovenfor) skal indeholde mellem 200 og 400 observationer (arbejdssteder). Hvis kun signifikanskravet om at kunne påvise ændringer på minimum 15% lægges til grund for be-

stemmelsen af stikprøven, kan der opereres med op til 4-5 strata inden for hvert visionsområde. Så opnås et antal observationer, der er nødvendigt for at opnå et samlet antal observationer på ca. 180, når der sammenlignes to strata, hvis den ønskede styrke skal opnås.

Det anbefales derfor at ændre på kravet til styrken, til fx 70%. I så fald kræves 300 observationer i alt for at opnå den ønskede statistiske styrke eller 150 observationer i hvert stratum. Der kan i så fald sammenlignes 2-3 strata (fx små, mellemstore og store arbejdssteder) inden for hvert visionsområde med tilsvarende strata i andre visionsområder eller i senere undersøgelser.

c) Er udtrækket af arbejdssteder foregået på en måde, der svarer til det ønskede formål?

d) Hvordan er udtrækkene af arbejdssteder faktisk foretaget?

Kontor for Effekt og Analyse i Arbejdstilsynet

Kontoret havde ønsket et udtræk inden for de syv visionsområder, som bygger på såkaldt visionsrepræsentativitet, dvs. at hver vision er repræsenteret ved relativt få forudvalgte »problematiske« branchegrupper. Inden for disse brancher var det hensigten at udvælge et antal virksomheder af forskellig størrelse, så såvel store, mellemstore som små virksomheder var repræsenteret. Derudover skulle udtrækket være ligeligt fordelt mellem de udvalgte branchegrupper. Udtrækket opfylder ikke de af Arbejdstilsynet fremsatte krav til udtrækket, idet der er en skævhed i forhold til branchegrupper og arbejdsstedsstørrelse. I bilagstabel C-4 i Arbejdstilsynets rapport »Overvågning 2001 (Bilag A-F-delen)«, ses de deltagende arbejdssteder fordelt efter antal ansatte sammenholdt med fordelingen af samtlige virksomheder i de pågældende branchegrupper fra Danmarks Statistik. De deltagende virksomheder fordeler sig forskelligt fra fordelingen af samtlige virksomheder i de pågældende branchegrupper på antal ansatte.

Overvågningsudvalget

Der har været forhandlinger omkring udvælgelse af branchegrupper i udvalget. Konsekvensen har været, at nogle branchegrupper er repræsenteret

med få arbejdssteder, hvilket gør det problematisk at generalisere resultater på branchegruppeniveau.

Kontrahenten ved AMI og CASA

Udtrækket er stratificeret således, at en virksomheds chance for at blive udtrukket er i relation til dens størrelse, dvs. antallet af ansatte. Dette princip har ifølge kontrahenten ved AMI og CASA sikret, at størrelsesfordelingen mellem virksomhederne i udtrækket er den samme som i det danske samfund inden for de udvalgte branchegrupper.

- e) I hvilken udstrækning sikrer den faktisk anvendte fremgangsmåde for stikprøveudvælgelse den ønskede repræsentativitet på branchegrupper?*
- f) Hvilke konsekvenser har evt. manglende repræsentativitet i stikprøven af arbejdssteder for sammenligning mellem data fra runde 1 og runde 2?*

Kontrahenten ved AMI og CASA

Der er en tendens til, at bortfaldet er størst blandt de små virksomheder, og det gør, at branchegrupper med mange små virksomheder har en forholdsvis lav svarprocent (under 60). Årsagen til ikke-deltagelse kan enten være arbejdsmiljøproblemer eller problemer med at afsætte tid til undersøgelsen – eller en kombination af disse to forhold. Resultaterne er rapporteret også for branchegrupper med en lav deltagelse, men det er gjort med det forbehold, at arbejdsmiljøarbejdet er som resultaterne viser eller ringere.

Faglig vurdering ved statistisk konsulent

Evaluator er blevet opmærksom på, at hverken udbudsmaterialet eller kontrakten mellem AMI og AT indeholder krav om repræsentativitet på branchegruppeniveau. Evaluator har imidlertid i samarbejde med AT opstillet evalueringsspørgsmålene, herunder spørgsmål 1e) om repræsentativitet på brancheniveau, hvorfor der tages stilling til det.

Til denne undersøgelse er i alt udtrukket 4.641 virksomheder og heraf indgår 3.287 (70,8%) i undersøgelsen. Af de virksomheder, der indgik i undersøgelsen var 43% udtrukket via NAK, 42% udtrukket via CVR og

de sidste 15% via spørgeskemaundersøgelsen blandt skolesøgende børn under 18 år. Svarprocenten (A-siden) svinger fra 44% inden for fjerkræslagterier til 97% inden for pasning af børn. Generelt er bortfaldet størst inden for privat erhverv, især fremstilling og mindst blandt offentlige virksomheder.

Der burde have været foretaget en bortfaldsanalyse med henblik på at undersøge, hvorvidt der er systematisk bortfald inden for de branchegrupper, hvor bortfaldet er størst. Der burde have været gennemført en bortfaldsanalyse for hele materialet.

De 4.641 udvalgte virksomheder blev udtrukket fordelt på branchegrupper, ca. 40 i alt, inden for de syv visionsområder. Det fremgår, som tidligere nævnt, ikke, hvorvidt der yderligere også er stratificeret efter virksomhedsstørrelse, men der nævnes i visionsrapporterne arbejdssteder med under 4 ansatte, med mellem 5 og 19 ansatte og med mere end 19 ansatte¹.

Det er sværere at sammenligne over aggregerede strata. Hvis der fx ønskes at beregne en samlet andel, der svarer ja til et bestemt spørgsmål inden for et visionsområde, skal de forskellige strata vejes sammen efter et eller andet kriterium. Det er imidlertid ikke klart, hvad dette kriterium skal være. Et kriterium kunne være de enkelte stratas beskæftigelsesandele. Altså at arbejdsstederne vægtes efter den andel af beskæftigelsen, de har inden for de branchegrupper, der er udvalgt inden det enkelte visionsområde.

Der er blevet foretaget et udtræk baseret på i hovedsagen to registre, et der er virksomhedsbaseret (CVR), og et der er medarbejderbaseret (NAK). Efterfølgende er data fra det virksomhedsbaserede udtræk vægtet efter antal beskæftigede og er dermed sammenligneligt med det medarbejderbaserede udtræk. De enkelte strata er dermed repræsentative med hensyn til beskæftigelse. Der er tale om meget varierende bortfaldsprocenter fra branchegruppe til branchegruppe. Men der er ikke i de rapporter fra VOV1, der forelå ved evalueringens start, foretaget en bortfaldsanalyse, og det er derfor ikke muligt at udtale sig om repræsentativiteten i det foreliggende udtræk. Evaluator har fået oplyst, at der efterfølgende foreligger en nyere version af rapporten om det generelle arbejdsmiljøar-

bejde, som det imidlertid ikke har været praktisk muligt at inddrage i evalueringen.

Ved vurderingen af de enkelte visionsrapporter fremgår det, at der inden for flere visioner er lav svarprocent (under 60) i flere branchegrupper, samt at der er en særlig lav svarprocent for små arbejdssteder.

g) Hvilken betydning har det, at unge over 18 år blev behandlet som en del af ulykkesvisionen?

Kontrahenten ved AMI og CASA

Med hensyn til at have unge mellem 18 og 24 år under ulykkesvisionen er det forskernes anbefaling, at det er hensigtsmæssigt at behandle denne gruppe under ulykkesvisionen bl.a. fordi der ikke findes særlige regler for ansatte i 18-24-års-alderen.

Faglig vurdering ved statistisk konsulent

Der er indsamlet oplysninger om unge. Unge under 18 år er indsamlet i et udtræk baseret på uddannelsesinstitutioner og er dermed repræsentativt, mens unge på 18 år og derover er indsamlet via de stratificerede udtræk. Da strata ikke er repræsentativ, kan de to udtræk i princippet ikke sammenlignes, hvilket heller ikke har været meningen. Ydermere er oplysningerne om de unge på 18 år og derover blevet indsamlet i forbindelse med ulykkesvisionen. I det omfang at fremdriften for arbejdsmiljøet for de unge på 18 år og derover ikke er repræsenteret ved de samme branchegrupper som ulykkesvisionen, er denne del af undersøgelsen naturligvis problematisk.

Samlet vurdering

Udtrækket opfylder ikke Arbejdstilsynets krav, idet udtrækkets fordeling på arbejdsstedsstørrelse ikke svarer til størrelsesfordelingen totalt i de udvalgte branchegrupper.

Datamaterialets dækning af de udvalgte branchegrupper kan ikke vurderes, idet det ikke af det foreliggende materiale fremgår, at der er foretaget bortfaldsanalyser. (Evaluatoren har fået oplyst, at en nyere version af den generelle rapport indeholder bortfaldsanalyser, men den var ikke til-

gængelig betids for evalueringen). Mange branchegrupper har imidlertid en svarprocent på under 60, hvilket gør det problematisk at generalisere uden bortfaldsanalyse.

Den lave deltagelse i flere branchegrupper og især for små arbejdssteder gør generaliserbarhed tvivlsom til henholdsvis branchegruppe og små arbejdssteder. Uden bortfaldsanalyse, der kan sandsynliggøre, om de lave svarprocenter er problematiske for generaliserbarhed eller ej, er det usikkert at foretage sammenligninger for disse branchegrupper vedkommende med en ny stikprøve for VOV2.

Forslag til VOV2

Evaluators vil foreslå, at man undlader at stille særlige krav om statistisk styrke (som med det valgte signifikansniveau på 5% er 50%). Der vil i så fald skulle opnås en samlet stikprøve på 180-400 observationer, dvs. der kan opereres med 4-5 strata i hver vision, med uændret krav om mindst 50 arbejdssteder i hvert stratum. Alternativt kan man sikre en større samlet stikprøve i de strata, hvor der indgår mange branchegrupper, mens visionsområder med et mindre antal branchegrupper kunne belyses med en mindre stikprøve. På den måde kan samme statistiske styrke opnås i alle stikprøvens strata.

Den lave svarprocent (under 60) i nogle branchegrupper er en trussel mod generaliserbarhed til visionen og dermed for sammenligning over tid. Det foreslås derfor at gennemføre bortfaldsanalyser for disse branchegrupper vedkommende for at sandsynliggøre, om det er muligt at generalisere eller ej.

Den statistiske konsulent foreslår, at man opstiller en statistisk model for ændringerne på de enkelte parametre, fx en logistisk regressionsmodel. Modellen kunne indeholde forklarende faktorer for sammenhængen mellem de studerede variabler på den ene side og branchegruppe og arbejdsstedsstørrelse på den anden. Hvis der indgår andre arbejdssteds karakteristika i undersøgelsen, kan disse også indgå i analysen. På denne måde kunne man opnå et klarere overblik over eventuelle systematiske effekter fra branchegruppe og arbejdsstedsstørrelse. Selv om nogle af disse effekter muligvis går på tværs af hinanden, vil den samlede statistiske model stadigvæk give information om, hvorvidt der er selvstændige ef-

fekter fra arbejdsstedsstørrelse og branchegruppe, men samtidig også information om, hvorledes den specifikke effekt af arbejdsstedsstørrelse påvirker arbejdsmiljøarbejdet inden for hver branchegruppe. Man ville desuden opnå, at de estimerede sammenhænge bliver ufølsomme over for stratificerings- og vægtningsproblemer. Man opnår konsistente estimater af effekter af branchegruppe og arbejdsstedsstørrelse på arbejdsmiljøet, uanset om der vægtes eller ej, og uanset om stikprøven er stratificeret eller ej.

3.2 Filterspørgsmål

- a) *Har filterspørgsmål tjent deres formål, nemlig at udelade arbejdssteder, som mod forventning ikke er relevante for den pågældende vision?*
- b) *Der blev ikke anvendt filterspørgsmål inden for alle visionsområder, og det skal derfor vurderes, hvilke konsekvenser forskellig anvendelse af filterspørgsmål har for generaliserbarhed og for analyserne.*
- c) *En anden type filterspørgsmål er dem, der bruges i spørgeskemaerne til at styre overspring af spørgsmål. Disse spørgsmål vil også blive vurderet for deres logik og hensigtsmæssighed, jf. den sproglige gennemgang af skemaer.*

Kontrahenten ved AMI og CASA

Ifølge *kontrahenten ved AMI og CASA* blev filterspørgsmålene anvendt for ikke at stille irrelevante spørgsmål. Der blev anvendt to typer filtre. Den ene drejede sig om arbejdsstedsstørrelse (i den generelle del af telefoninterviewskemaet), da arbejdsmiljøreglerne er forskellige afhængigt af antallet af ansatte. Den anden type filtre blev anvendt til at sikre, at alene de arbejdssteder hvor arbejdsmiljøproblemet er relevant, svarer på bestemte spørgsmål. Konkrete eksempler på sådanne filterspørgsmål findes inden for visionerne »kemi« og »ergonomi«. Inden for begge visionsområder viste det sig, at nogle spørgsmål ikke var relevante for nogle af arbejdsstederne i udtrækket (fx bestemte kemikalier eller EGA), og disse spørgsmål blev så ikke stillet. Dette forhold har ifølge forskerne ikke konsekvenser for generaliserbarheden eller gyldigheden af resultaterne på visionsområdet. Det vil sige, at telefoninterviewene blev gennemført i alle

de virksomheder, der indgik i udtrækket, også selv om der var mange spørgsmål, der ikke var relevante for den pågældende virksomhed.

Den faglige (kollegiale) vurdering

For børn og unge-visionen gælder, at de første tre spørgsmål i spørgeskemaet bruges til at kontrollere, om arbejdspladsen har børn og unge ansat. Er det ikke tilfældet, indgår interviewet ikke i materialet. Repræsentativitetsspørgsmålet og frafaldet bliver diskuteret i rapporten, og det konstateres, at resultaterne af både telefoninterviewene og besøgene skal fortolkes med forsigtighed. Der sker bortfald i to omgange, først er der 35% af de udtrukne 747 arbejdspladser, der ikke opnås kontakt med. Af de 482, der opnås kontakt med, er der 288 arbejdspladser, der aktuelt har børn og unge ansat, og som indgår i undersøgelsen, dvs. endnu et bortfald på 40%. Dertil kommer, at det er uklart, om undersøgelsen alene skal omfatte børn og unge under 18 år i fritidsjob, og hvis dette er ønsket, om filterspørgsmålet har fungeret efter hensigten. Alt i alt er det uklart, hvad repræsentativiteten skal vurderes i forhold til.

Med hensyn til visionen ergonomi: Selv om visionsområdet faktisk består af to undersøgelser, EGA og tunge løft, er der kun foretaget ét udtræk, hvilket er uhensigtsmæssigt, da det er forskellige branchegrupper, der er problematiske i forhold til disse to arbejdsmiljøproblemer. Det viser sig, at udtrækket især er utilfredsstillende i forhold til EGA, da relevante EGA-branchegrupper (såkaldte traditionelle EGA-branchegrupper) kun udgør 37% af materialet i undersøgelsen (172 ud af 459 arbejdsgiver-svar). Derfor er branchegrupper af relevans for EGA-problematik så små, at deres repræsentativitet er tvivlsom.

For visionen kemi er nøgleordet i det indledende visionsspecifikke filterspørgsmål »kemikalier« et uheldigt begreb, hvis der ønskes oplysninger om fx maling, rengøringsmidler o.l. Det anbefales generelt at bruge formuleringen »kemiske produkter«. Det er muligt, at formuleringen »kemikalier« i spørgsmål 54 kan have filteret relevante respondenter fra i VOV1. Hvis spørgsmål 54 ændres til en formulering i VOV2, der medtager flere typer af kemiske produkter, vil sammenligning mellem VOV1 og VOV2 kunne blive problematisk.

Samlet vurdering

Der forekom to typer af filterspørgsmål. En type er spørgsmål, der anvendes til at styre overspring af irrelevante spørgsmål, fx spørgsmål om antal ansatte på arbejdsstedet, idet nogle spørgsmål kun skulle besvares af arbejdssteder over en bestemt størrelse.

Den anden type filterspørgsmål er dem, der anvendes til at udelade arbejdssteder, som mod forventning ikke er relevante for den pågældende vision.

Endelig er der også spørgsmål, der ufrivilligt kommer til at tjene som et filter for, hvad svarpersonerne mener, er omfattet af spørgsmålet, fx hvad begrebet »kemikalier« omfatter, eller betegnelsen »børn og unge under 18 år i fritidsjob«.

Filterspørgsmål for at udelade arbejdssteder er ikke ifølge telefoninterviewerne blevet anvendt af telefoninterviewerne til at undlade interview. Hvilke konsekvenser det har for generaliserbarhed afhænger af, hvor mange arbejdssteder, det drejer sig om, og om de afviger systematisk fra de øvrige arbejdssteder i branchegruppen. Rapporterne indeholder ingen oplysninger om bortfaldsanalyse.

3.3 Interview med såvel A- som B-side

I første undersøgelsesrunde skulle gennemføres interview med såvel A- som B-siden med det formål at give en mere nuanceret beskrivelse af handlingsplanens gennemslagskraft samt for at vurdere graden af overensstemmelse.

- a) *Hvilken informationsværdi har det givet at telefoninterviewe såvel A- som B-side?*
- b) *Hvordan har man håndteret uoverensstemmelser i besvarelsene?*

I første undersøgelsesrunde blev der i en række tilfælde kun interviewet den ene part.

- a) *Hvad er fordele og ulemper ved, at der på nogle arbejdssteder er interview med kun en ledelsesrepræsentant eller med kun en sikkerhedsrepræsentant eller en anden medarbejderrepræsentant?*

b) *Spørgsmålenes svarevne for henholdsvis ledelsesrepræsentant og sikkerhedsrepræsentant (her tænkes på eksponeringsspørgsmål)?*

Kontor for Effekt og Analyse i Arbejdstilsynet

Ifølge *Kontor for Effekt og Analyse* var den oprindelige idé at belyse, om A- og B-siden på arbejdsstederne har den samme vurdering. Hele sikkerhedssystemet på arbejdsstederne bygger på det princip, at begge parter er repræsenteret, og derfor er det af interesse at have de to svar. I praksis er der en hel del sammenfald i besvarelserne på baggrund af de foreliggende rapporter – bortset fra det psykiske arbejdsmiljø. Samtidig antyder de interviewede fra *Kontor for Effekt og Analyse*, at analyserne af overensstemmelse mellem A- og B-siden ikke er tilstrækkelige i rapporterne.

Overvågningsudvalget

I *Overvågningsudvalget* er der bred tilslutning til at interviewe både A- og B-sidens repræsentanter, bl.a. fordi hele arbejdsmiljøarbejdet i virksomhederne bygger på de to siders repræsentation. Der har været to hensigter med denne tilgang. På den ene side kan man tale om kvalitetsmæssig validering af resultaterne, når begge parter bliver hørt. På den anden side kan parternes (eventuelt divergerende) synspunkter bruges i den politiske debat. Udvalget savner imidlertid mere direkte analyser (end kappaværdier) og fortolkninger af de to siders svar.

Kontrahenten ved AMI og CASA

Ledelsen hos kontrahenten vurderer, at modellen for at interviewe både A- og B-sidens repræsentant har været god og vigtig. Dog er der forskelle mellem visionsområderne med hensyn til, hvor megen information de to sæt oplysninger har givet. Den største diskrepans er på det psykosociale område. Forskernes vurdering er, at de eksisterende data kan analyseres mere tilbunds gående; de foreliggende analyser er alene beskrivende. Fremgangsmåden med at interviewe både A- og B-side fordobler antallet af telefoninterviewene. Hvis der ønskes færre interview er det værd at overveje systematisk at interviewe A-siden (på grund af, at de er i besiddelse af flere faktuelle oplysninger) og interviewe en stikprøve af B-sidens repræsentanter.

Telefoninterviewere

Tre telefoninterviewere har erfaring med, at der er forskelle i svarevne eller svarvillighed mellem ledelsesrepræsentant og medarbejderrepræsentant. Medarbejderrepræsentanterne kan i nogle tilfælde være tilbageholdende eller forsigtige med deres svar, idet de eventuelt oplever intervieweren som kontrollant, eller ifølge en telefoninterviewer frygter reaktioner fra ledelsen. Der er også tilfælde, hvor medarbejderrepræsentanten ikke har den viden om et sagsforhold, som ledelsesrepræsentant har. Ledelsesrepræsentanter kan formulere sig, men de er i nogle tilfælde meget korte i deres svar og har tendens til at give mere positive vurderinger end medarbejderrepræsentanter.

Den faglige (kollegiale) vurdering

I de fleste visionsrapporter bemærkes manglende overensstemmelser mellem svar fra A- og B-siden, og mellem telefoninterview og besøg.

Generelt nøjes man med en beskrivelse af forskelle i A- og B-sidens besvarelser, men der gives sjældent specifikke forklaringer. Generelt tolkes overensstemmende besvarelser som tegn på pålidelighed. Der anvendes i flere rapporter et mål for grad af overensstemmelse (kappaværdier). Disse rapporteres i nogle af tabellerne, men der drages sjældent konsekvenser, fx at tal undlades, fordi overensstemmelsen er for lav. I stedet for kappaværdier kunne man anvende korrelationer for at vise grad af overensstemmelse.

Evaluator savner en nærmere diskussion af, hvornår uoverensstemmende svar mellem A- og B-siden har rod i, at A og B ser forskelligt på ting, fordi de har forskellige positioner, og hvornår der er tale om pålidelighedsbrist.

Samlet vurdering

Der er overvejende enighed om, at både A- og B-siden bør interviewes i lige fuldt omfang. Der er ligeledes enighed om, at analyserne af forskelle mellem svar fra henholdsvis A- og B-siden bør foretages mere systematisk og dybdegående, og forskelle bør diskuteres og mulige forklaringer gives.

Evaluators deler disse vurderinger. Både A- og B-siden bør indgå i lige omfang, dels af arbejdsmarkedspolitiske grunde, dels fordi de to parter kan forventes at have forskellig viden om og forskellig indfaldsvinkel til arbejdsmiljøet. Generelt mener evaluatoren, at uoverensstemmende svar fra A- og B-siden bør kunne udnyttes bedre, og at forskelle ikke behøver at have noget at gøre med pålidelighed.

3.4 **Interview med sikkerhedsrepræsentant versus anden medarbejder**

I første undersøgelsesrunde var hensigten på B-siden at interviewe sikkerhedsrepræsentanter, fordi man antog, at de var bedre i stand til at besvare spørgsmål om arbejdsstedets forebyggende aktiviteter. Oprindeligt var det Arbejdstilsynets tanke at gennemføre interview med flere arbejdstagere på det enkelte arbejdssted frem for en repræsentant, men den model blev ikke valgt i forbindelse med udviklingen af VOV I.

a) Hvad er fordele og ulemper ved at interviewe henholdsvis en sikkerhedsrepræsentant og en almindelig medarbejder?

Kontor for Effekt og Analyse i Arbejdstilsynet

Den oprindelige idé (jf. Udbud af overvågning af fremdriften i 2005 handlingsplanen, januar 2000, s. 6) var at vælge flere menige medarbejdere som medarbejderrepræsentant ved telefoninterviewene bl.a. fordi sikkerhedsrepræsentanten ikke altid kan forventes at have viden på en bestemt problematik, eller sikkerhedsrepræsentantens svar kan være »farvede« på grund af dennes særlige position på virksomheden. Det nævnes i udbuddet, at de ansatte skal have jobfunktioner, der er relevante for problemstillingen, og at virksomhederne aldrig selv må udpege, hvem af de ansatte der skal interviewes.

Dette blev fraveget i Udbud af dataindsamlingsdelen i projektet fra september 2000, hvor det nævnes (s. 2), at der skal interviewes 2 nøglepersoner på hver arbejdssted, typisk en sikkerhedsleder og en sikkerhedsrepræsentant, som er medlem af sikkerhedsudvalget.

Begrundelsen er, at en sikkerhedsrepræsentant har et kendskab til arbejdsmiljøarbejdet (jf. de generelle spørgsmål i telefoninterviewskemaerne), hvilket ikke kan forudsættes fra en menig medarbejder. Valget af en sikkerhedsrepræsentant var på sin måde et kompromis mellem disse forskellige hensyn og et praktisk valg, da det var svært ifølge forskerne at finde kriterier for, hvordan en menig medarbejderrepræsentant kunne udvælges (selv om dette var beskrevet i Arbejdstilsynets udbudsmateriale af januar 2000, s. 6). Det kan imidlertid diskuteres, hvordan sikkerhedsrepræsentanten er blevet udvalgt i undersøgelsen, herunder om repræsentanten er tilfældigt udvalgt eller med kendskab til visionsområdet, fx EGA-problematikken. Det er ikke oplyst i rapporterne, hvordan sikkerhedsrepræsentanterne er udvalgt.

Overvågningsudvalget

Overvågningsudvalgets repræsentanter har delvis forskellige holdninger til, hvorvidt det var en rigtig beslutning at udvælge en sikkerhedsrepræsentant som medarbejderrepræsentant i undersøgelsen. Nogle medlemmer pointerer, at det var det rigtige valg, fordi en sikkerhedsrepræsentant har mest viden. Andre synes, at en menig medarbejder ville være en god informant med henblik på at få et billede af niveauet af arbejdsmiljøarbejdet i virksomheden.

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA mener, at der ved at udvælge en sikkerhedsrepræsentant som repræsentant for medarbejdersiden blev sikret ensartethed i besvarelserne fra virksomhed til virksomhed. Det ville have været umuligt at vælge en menig medarbejder, som var den »rigtige« repræsentant for medarbejderne. Bl.a. ville der have været en fare for, at man udvalgte en medarbejder, som arbejdsgiversiden anbefaler i og med, at A-siden blev interviewet først og kontakten til B-siden gik gennem A-siden. Dertil kommer, at flere af spørgsmålene (de generelle spørgsmål) ikke ville have været muligt at stille til de menige medarbejdere, dvs. at de ikke har forudsætninger for at besvare spørgsmål, som fx handler om deltagelse i arbejdsmiljøuddannelse.

Telefoninterviewere

Der er forskel på vurderingerne hos telefoninterviewerne. Tre af dem mener, at sikkerhedsrepræsentanten er den mest relevante at interviewe. Begrundelserne, der nævnes, er, at en sikkerhedsrepræsentant normalt kan overskue et større område og har kendskab til flere medarbejderes synspunkter og eventuelle kritik af arbejdsmiljøet. Sikkerhedsrepræsentanten er ofte mere engageret i arbejdsmiljøproblematikker end en menig medarbejder. I modsætning hertil nævner en anden interviewperson, at det er en fordel at interviewe en menig medarbejder, fordi en sikkerhedsrepræsentant i højere grad varetager virksomhedens interesser og derfor ikke vil svare lige så ærligt som en menig medarbejder. Hos en interviewperson er det blandet, om sikkerhedsrepræsentant eller menig medarbejder er mest informativ at interviewe. Der er store forskelle indbyrdes mellem sikkerhedsrepræsentanterne i en grad, så menige medarbejdere i enkelte tilfælde vidste mere end nogle sikkerhedsrepræsentanter.

Samlet vurdering

Der har i udbudsfasen været forskellige opfattelser af, om man burde vælge ansatte uden tillidshverv eller en sikkerhedsrepræsentant til interview. Der er også lidt forskellige synspunkter hos de personer, evaluator har interviewet. Indtrykket er, at hvis interviewpersonerne lægger vægt på et overblik over de forebyggende aktiviteter og organiseringen af arbejdsmiljøarbejdet, vil man foretrække, at det var en sikkerhedsrepræsentant. Hvis man derimod lægger vægt på at få niveauet af de forebyggende aktiviteter og eksponering belyst, vil man foretrække menige medarbejdere.

Det er evaluators vurdering, at valget af en sikkerhedsrepræsentant i sikkerhedsudvalget, hvor et sådant findes, er forsvarligt.

3.5 Validering og supplering af telefoninterview ved hjælp af arbejdsstedsbesøg

Sigtet med at kombinere telefoninterview og arbejdspladsbesøg var, at arbejdsstedsbesøgene skulle validere såvel som supplere telefoninter-

viewene. Validering skulle forstås således, at faktuelle forhold, konstateret ved besøgene, skulle sammenlignes med informationer fra telefoninterviewet.

Kontor for Effekt og Analyse i Arbejdstilsynet

Repræsentanterne for Kontor for Effekt og Analyse understreger, at validering og supplering af telefoninterview ved hjælp af arbejdsstedsbesøg var to sideordnede, lige vigtige formål med undersøgelsen og en væsentlig del af undersøgelsens metodiske design. Vurderingen er, at formålet med at validere ikke er blevet opfyldt i den gennemførte undersøgelse. Der nævnes to eventuelle årsager dertil. For det første var besøgene måske ikke indrettet således, at de kan anvendes til at validere interviewdata. For det andet er analyserne mellem interview- og besøgsdata ikke gennemført på en måde, som sigter mod at validere telefoninterviewene. Det nævnes, at der råder forskellige opfattelser (henholdsvis mellem kontoret og kontrahenten) om, hvad der er mest validt, telefoninterviewene eller vurderingerne ved arbejdsstedsbesøgene; for kontoret er telefoninterviewene hoveddatakilden.

Med hensyn til spørgsmålet, hvorvidt det lykkedes at *supplere* telefoninterviewene ved hjælp af besøgsdata er det vurderingen af repræsentanter for Kontor for Effekt og Analyse, at besøgsdata hverken er brugt i tilstrækkelig grad eller analyseret på en tilfredsstillende måde i visionsrapporterne. Disse data er ikke en integreret del af rapporterne, og det har ikke været meningen, at besøgsdata skulle kvantificeres. I forhold til resourceforbrug er nytteværdien af besøgsdelen af undersøgelsen beskedent.

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA mener, at arbejdsstedsbesøgene er vigtige, idet de fungerer som målestok, så man kan vurdere »rækkevidden af telefoninterviewene«. Ved telefoninterviewene tenderer respondenterne at være mere positive, end »de burde være«, og derfor vil besøgene i bedste fald fungere som graduering af telefoninterviewene. Kontrahenten vurderer, at metoden med henblik på at validere kun delvis er udviklet, og de indikerer, at koblingen mellem telefoninterviewene og arbejdsstedsbesø-

gene ikke var helt gennemført, inden dataindsamlingen gik i gang bl.a. på grund af tidsmangel.

Med hensyn til at supplere telefoninterview ved arbejdsstedsbesøg er det kontrahentens vurdering, at besøgene resulterer i supplerende information, da virksomhedsrepræsentanter ved telefoninterview ikke kan vurdere forholdene på samme måde som arbejdsmiljøprofessionelle ved besøg. Telefoninterviewene alene har en begrænset værdi, og derfor er det vigtigt at have besøgsdelen med også i undersøgelsens næste runde. Det er kontrahentens vurdering, at data fra arbejdsstedsbesøgene kunne udnyttes i højere grad, end det er gjort i visionsrapporterne – både med henblik på koblingen med telefoninterviewene og med henblik på en generel metodeudvikling. Dette er ikke gjort dels på grund af tidsmangel, dels på grund af, at besøgsdelen ifølge Kontor for Effekt og Analyses udmelding primært skal bruges til evaluering af telefoninterviewene.

Besøgskonsulenter

Ifølge projektlederen for besøgsdelen var telefoninterviewene og arbejdsstedsbesøgene to selvstændige, adskilte dataindsamlinger. Det er både projektlederens og konsulenterne egen vurdering, at konsulenterne derfor ikke har forudsætninger for at svare på spørgsmålet om validering og supplering. Besøgskonsulenterne har imidlertid gjort overvejelser om, hvad arbejdsstedsbesøgene kan bidrage med, og disse synspunkter refereres kort nedenfor.

Nogle konsulenter konstaterer, at formålet med at validere – og supplere – telefoninterview ved hjælp af besøg ikke var et klart formuleret formål fra starten af undersøgelsen, mens andre mener, at de var bekendt med dette overordnede formål med undersøgelsen.

Principielt mener flere konsulenter, at det er »fornuftigt« at forsøge at validere og supplere telefoninterview gennem arbejdsstedbesøg. Ved interviewet nævner konsulenterne imidlertid forhold, som efter deres vurdering gør sammenligninger (validering) mellem telefoninterviewene og besøgene problematisk. For det første var virksomhedsrepræsentanterne delvis andre ved interview og besøg. For det andet dækkede besøget på større arbejdssteder typisk én afdeling, mens interviewet omhandlede hele virksomheden. Endvidere stiller en af konsulenterne følgende metodiske

spørgsmål: Hvad ønskede man at validere, når interviewene består af to vurderinger (A- og B-side) og besøgene af én vurdering?

Konsulenterne oplever, at besøgene giver en anden type information end telefoninterviewene i og med, at konsulenternes opgave var også at afdække arbejdsstedernes reelle arbejdsmiljøarbejde i forhold til det formelle (mens interviewene primært handlede om det formelle). Det er muligt at uddybe problemstillinger ved besøg, bl.a. indeholder besøgsdelen forklaringer på, hvorfor der er forskelle mellem oplysninger fra telefoninterview og besøg, og de kunne med fordel bruges i rapporter. Et andet eksempel på, at besøgene resulterede i uddybende informationer var oplysninger om, hvordan APV var blevet gennemført, og hvilke barrierer der havde været i den forbindelse. Nogle af de informationer, som konsulenterne indsamlede, findes i form af tekstdata, og konsulenterne er i tvivl, om disse oplysninger er blevet brugt overhovedet. Et konkret punkt, hvor der blev indsamlet supplerende informationer ved besøgene, var en liste over kemiske produkter på kemiområdet. De konsulenter, som har læst udkastet til visionsrapporten på »deres« visionsområde, oplever, at brugen af oplysninger fra besøgsdelen er meget begrænset.

De fleste konsulenter mener imidlertid, at det ville have været en fordel at kende telefoninterviewene ved besøgene, så man ikke skulle stille de samme spørgsmål på ny, og man havde mulighed for at nå mere ved besøget. Nogle konsulenter havde oplevet, at virksomhedsrepræsentanterne på nogle virksomheder var »lidt irriteret« over, at de skulle svare på samme spørgsmål som ved telefoninterview. Nogle mener, at gentagelse af samme spørgsmål kan opleves som kontrol. Samtidig nævner nogle konsulenter, at kendskab til telefonsvarene ville have givet en skævhed i konsulenternes vurderinger ved besøgene og dermed gjort valideringen umuligt. Flere konsulenter beskriver imidlertid, at kontrahentens begrundelse for, at konsulenterne ikke måtte se den enkelte virksomheds svar ved telefoninterview var, at virksomhederne var blevet lovet fortrolighed i forbindelse med telefoninterview – og ikke det metodiske argument om to uafhængige dataindsamlinger.

Den faglige (kollegiale) vurdering

Vi har for hvert visionsområde udvalgt de spørgsmål, i udgangspunktet

faktuelle spørgsmål, som efter vores vurdering er bedst egnet til at søge valideret². Vi har herefter sammenholdt disse udvalgte telefoninterviewspørgsmål med spørgsmålene i vurderingsskemaet, som besøgs konsulenterne har udfyldt i forbindelse med den kvantitative vurdering.

Som det fremgår af tabellen nedenfor, varierer antallet af spørgsmål, som ud fra den betragtning, at de handler om faktuelle forhold, er egnet til validering i telefoninterviewskemaet, i betydelig grad fra visionsområde til visionsområde – fra ingen spørgsmål på visionsområdet ulykker til 25 spørgsmål på børn og unge-området. Derudover er der stor variation fra vision til vision med hensyn til, hvor mange spørgsmål der matcher mellem telefoninterview og besøg. Der er flest identiske spørgsmål på visionsområderne børn og unge (17) og dernæst på visionsområderne indeklima og støj (henholdsvis 9 og 8 spørgsmål). På visionsområderne ergo-nomi er der 6 og kemi 5 spørgsmål, som er de samme ved telefoninterview og vurderingsskema, mens der på visionsområderne psykisk arbejdsmiljø og ulykker slet ikke findes spørgsmål, der helt matcher (jf. tabellen nedenfor). Kun to af de 23 spørgsmål, som handler om faktuelle forhold omkring generelt arbejdsmiljøarbejde, er identiske i telefoninterviewskema og vurderingsskema.

Antal spørgsmål som er identiske i telefoninterview og ved besøg for hver vision

<i>Visionsområde</i>	<i>Match</i>	<i>Omtrent match</i>	<i>Intet match</i>	<i>Faktuelle spørgsmål i alt</i>
Det generelle arbejdsmiljøarbejde	2	7	14	23
Børn & unge	17	2	6	25
Ergonomi (tunge løft og EGA)	6	4	5	15
Indeklima	9	0	0	9
Kemi	5	5	0	10
Psykisk arbejdsmiljø	0	6	3	9
Støj	8	4	1	13
Ulykker*	(1)	(2)	(-)	(3)

* Ingen faktuelle spørgsmål; de tre spørgsmål er skønsmæssige.

Samlet vurdering

Kontor for Effekt og Analyse mener ikke, at besøgsdata i sin nuværende form tjener de to sideordnede formål, som der var for denne del af undersøgelsen, nemlig at validere og supplere telefoninterviewdata. Dermed sætter bestilleren et spørgsmålstegn ved nytten af delundersøgelsen. Kontrahenten nævner mangler ved koblingen mellem telefoninterview og besøgsdata i metodeudviklingsfasen, men understreger vigtigheden af besøgene som en uafhængig datakilde og konstaterer samtidig, at informationerne fra besøgene ikke er udnyttet optimalt i de nuværende rapporter. Tidsnød i metodeudviklings- og rapporteringsfasen nævnes som årsager til mangler. Som udførere af arbejdsstedsbesøgene peger besøgs konsulenterne på forhold, som vanskeliggør sammenligning (validering) mellem telefoninterview og besøg, men de mener, at de eksisterende besøgsdata i højere grad kunne bruges uddybende, end det er tilfældet med visionsrapporterne. Kollegial vurdering af visionsrapporterne viste, at antallet af spørgsmål, som er egnet til at blive valideret ved besøg, varierer fra vision til vision, når der tages udgangspunkt i spørgsmål, som belyser faktuelle forhold (flesteprocent spørgsmål på visionsområderne børn og unge og støj, mindst på områderne psykisk arbejdsmiljø, kemi og ulykker). I de fleste tilfælde er der få spørgsmål, der er identiske mellem telefoninterview og besøg.

Evaluatoren deler ovennævnte vurderinger, dvs. det metodisk set er problematisk at validere telefoninterview med arbejdsstedbesøg på grundlag af eksisterende data fra VOV1, primært fordi de to dataindsamlinger ikke var designet til dette formål. Dels indeholder telefoninterviewene på nogle visionsområder få spørgsmål, som er egnet til at blive valideret, dels er spørgsmålene ikke stillet således, at det muliggør validering mellem telefoninterview og besøg. Dertil kommer, at respondenterne og/eller undersøgelsens genstand er delvis forskellig ved interview og besøg; det gælder især for store virksomheder. Endvidere er det ikke tydeliggjort, hvad der ønskes sammenlignet med telefoninterviewene ved besøg: enkeltspørgsmål i vurderingsskemaet (som indgår i den samlede kvantitative vurdering) eller den samlede kvalitative vurdering. Kontrahenten betragter den samlede kvalitative vurdering som hovedparameter ved arbejds-

stedsbesøgene, men den egner sig næppe til sammenligning, da telefoninterviewene ikke indeholder en tilsvarende parameter.

Derimod er der i højere grad mulighed for at bruge besøgene fra VOV1 til at supplere og uddybe de problemstillinger, der er belyst gennem telefoninterviewene. Dette er imidlertid sket i meget begrænset omfang.

4 Spørgeskemaer og vurderingsskemaer i VOV1

4.1 Temaernes relevans for udviklingen samt kvantificering/kvalificering af besøgsdata

Arbejdstilsynet havde ved første undersøgelsesrunde fastlagt en række temaer (problemstillinger), der skulle belyses, og som er beskrevet i Arbejdstilsynets udbudsmateriale. Det drejede sig om en serie spørgsmål om visionsspecifikke forebyggende aktiviteter. Derudover indeholdt nogle af spørgeskemaerne spørgsmål om risikofaktorer, idet man manglede viden om disse.

a) Er de fastlagte temaer (problemstillinger) i interviewene såvel som ved besøgene relevante for måling af udviklingen af de forebyggende aktiviteter på arbejdsstederne? Hvilke temaer er ikke relevante? Er der behov for nye temaer?

Kontor for Effekt og Analyse

Kontor for Effekt og Analyse understreger, at formålet med undersøgelsen var at beskrive »fremdriften inden for de syv visioner«, hvilket betyder, at hovedvægten skal være på visionsspecifikke forhold. Samtidig efterlyser Kontoret en systematik i undersøgelsen på tværs af visionsområderne, så resultaterne fra områderne kan sidestilles med hinanden.

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA har delvis en forskellig opfattelse af, hvordan temaerne blev til. Repræsentanten for Arbejdsmiljøinstituttet be-

skriver, at de endelige temaer vedrørende det generelle arbejdsmiljøarbejde er instituttets forslag efter drøftelser med bestilleren, Kontor for Effekt og Analyse. De visionsspecifikke spørgsmål bygger på de forslag, som de sagkyndige i Arbejdstilsynet foreslog med det forbehold, at nogle af temaerne ikke var egnede til fuldt ud at blive belyst ved telefoninterview. Repræsentanten for CASA mener i højere grad, at temaerne var givet i forvejen i den forstand, at de i stor udstrækning var formuleret af faglige eksperter. Der var i nogen grad dialog med de visionsansvarlige i Arbejdstilsynet, men konstrueringen af skemaerne var for kortfattet en proces på grund af den korte tidsplan. Ifølge CASA ville det være ønskværdigt med flere tværgående temaer/problemstillinger.

Kontrahenten ved AMI og CASA har ikke umiddelbart forslag til yderligere temaer – det vil forudsætte en betydelig indsats at komme med et sådant forslag med hensyn til samtlige visionsområder.

Endvidere er det projektledelsens vurdering, at det »rimeligt godt« lykkedes at måle de forebyggende aktiviteter i virksomhederne, især ved arbejdsstedsbesøgene.

På spørgsmålet om der er behov for at supplere med nye visionsspecifikke temaer eller temaer, som går på tværs af de syv visionsområder, er AMI og CASA's vurdering, at der i de foreliggende rapporter er meget fokus på de enkelte visioner. Det ville være oplagt at se på sammenhængen mellem det generelle arbejdsmiljøarbejde og indsatsen i forhold til visioner.

Besøgskonsulenter

Besøgskonsulenterne på de syv visionsområder oplyser at have fået at vide, at indholdet af besøgene var det samme som ved telefoninterviewene, men konsulenterne har ikke været med til at diskutere temaernes relevans for undersøgelsens formål. Størstedelen af konsulenterne har den umiddelbare vurdering, at temaerne var »fornuftige« og »relevante nok« for arbejdsmiljøarbejdet på visionsområdet. Kun få konsulenter tog egentlig stilling til, hvorvidt temaerne er relevante for at måle udviklingen af forebyggende aktiviteter. Disse konsulenter havde gennemgående kritik af, at der var for lidt fokus på at kortlægge de forebyggende aspekter og for megen vægt på det formelle og det lovgivningsmæssige.

Den faglige (kollegiale) vurdering

Det er vurderet, om de valgte temaer (også kaldet problemstillinger) i de enkelte rapporter er i overensstemmelse med Arbejdstilsynets udbud af januar 2000, og om de er relevante for undersøgelsens formål, som er at udvikle og afprøve metoder til overvågning af udviklingen i arbejdsmiljøarbejdet samt at danne grundlag for, at aktørerne eventuelt kan justere deres indsats.

Vi har fundet, at der er overensstemmelse mellem de temaer, der nævnes som eksempler i Arbejdstilsynets udbud, og de temaer, der tages op i visionsrapporterne, ligesom vi vurderer, at temaerne i hovedsagen er dækkende i forhold til Arbejdstilsynet udbud. Dog savner vi en langt mere grundig behandling og refleksion af de metodemæssige erfaringer og lærestykker, der måtte være gjort. Det var jo et af hovedformålene med undersøgelsen at udvikle og afprøve metoder til overvågning.

I flere visionsrapporter er der tale om direkte afskrift af temaerne og ingen eller liden udvikling af og begrundelse for problemstillingerne (ud over den operationalisering, der finder sted ved formulering af spørgsmål til skemaer). Det kunne have været ønskeligt med en udvikling af temaer om de forebyggende aktiviteter og begrundelser for temavalget, således at læseren fik en bedre forståelse for, hvorfor netop de problemstillinger er vigtige.

Rapporten *Det generelle arbejdsmiljøarbejde* behandler relevante temaer om de formelle forhold omkring organisering af arbejdsmiljøarbejdet. Vi vurderer imidlertid, at der bør skæres ned på de formelle ting og udvikles temaer, der kommer tættere på de forebyggende aktiviteter. Det undrer os, at denne rapport ikke indeholder en sammenfattende sammenligning på tværs af visionerne. En sammenfattende rapport er nævnt i kontrakten mellem Arbejdstilsynet og Arbejdsmiljøinstituttet som det sidste produkt af undersøgelsen.

b) Hvordan er data fra interviewguider anvendt ved arbejdsstedsbesøgene omsat til vurdering på de kvalitative vurderingsskalaer i vurderingsskemaerne?

Besøgskonsulenter

De fleste konsulenter gennemførte først et interview om de generelle spørgsmål, havde derefter en rundvisning og efterfølgende gennemførte de et interview om de visionsspecifikke spørgsmål. Det var den procedure, som konsulenterne skulle følge, men i praksis var der forskelle i gennemførelse af besøgene. De fleste konsulenter udfyldte kun interviewskemaet ved besøget, mens de udfyldte vurderingsskemaet og lavede den kvantitative og den kvalitative vurdering, i denne rækkefølge, efter besøget. En konsulent benyttede slet ikke interviewskemaet, men udfyldte vurderingsskemaet direkte på arbejdspladsen. En konsulent lavede den kvalitative vurdering forud for den kvantitative. Nogle konsulenter justerede deres kvantitative vurdering, hvis den samlede kvantitative score virkede »for høj« fx i forhold til den referenceramme, som konsulenten havde fra andre arbejdssteder.

Ifølge de fleste konsulenter var den *kvantitative* vurdering mere entydig end den kvalitative vurdering, idet der var en bestemt procedure, man kunne følge. Den kvantitative vurdering var dokumenteret i den forstand, at den tog udgangspunkt i interviewskemaet, som blev anvendt til at udfylde vurderingsskemaet. Flere konsulenter peger imidlertid på, at den kvantitative vurdering reelt består af flere vurderinger, idet konsulenterne skal omsætte verbale svar fra interviewskemaet til tal i vurderingsskemaet, og den er ikke mere præcis eller objektiv. Et andet problem var, at vejledningen ikke indeholdt alle svarkombinationer, så konsulenten måtte i flere tilfælde vælge mellem to svarkategorier på den samlede kvantitative skala. Det var en overvejelse, tilfældighed eller helhedsindtrykket, som var afgørende for den endelige score.

Den *kvalitative* vurdering var typisk »helhedsindtrykket«, eller »en slags opsummering af kvaliteten af arbejdsmiljøarbejdet«, som konsulenterne gjorde på basis af hele besøget. Flere konsulenter nævner, at rundvisningen samt interview om de visionsspecifikke spørgsmål var meget centrale i forhold til den kvalitative vurdering. Her spillede konsulentens ekspertise en stor rolle, fx at konsulenten er vant til at observere forhold på arbejdspladser og stille relevante supplerende spørgsmål. Den kvalitative vurdering er i højere grad subjektiv og er ikke på samme måde dokumenteret som den kvantitative. Konsulenternes begrundelser findes

som tekstfiler, men de er ikke i samme grad struktureret som de kvantitative oplysninger.

Den mest almindelige rækkefølge var, at konsulenterne først lavede den kvantitative vurdering (hvor de fleste konsulenter først overførte oplysninger fra interviewskemaet til vurderingsskemaet og derefter omsatte resultater til en samlet kvantitativ score ved hjælp af vejledningen i vurderingsskemaet) og til sidst den kvalitative vurdering. Men der var forskelle i praksis mellem konsulenterne.

De fleste konsulenter mener, at de lavede den kvantitative og den kvalitative vurdering uafhængig af hinanden, og at den kvantitative vurdering på ingen måde »overdømmer« den kvalitative – snarere omvendt. Men samtidig giver flere konsulenter udtryk for, at de to vurderinger i praksis ofte lå tæt på hinanden. Nogle konsulenter er af den opfattelse, at de to vurderinger »ideelt set burde være ens«, men at det ikke altid var tilfældet.

- c) *I vurderingsskemaerne skulle foretages både en kvantitativ og kvalitativ opgørelse. Den kvalitative opgørelse er beskrevet som den endelige vurdering af arbejdsstedet. Hvordan er de kvantitative og kvalitative opgørelser anvendt i praksis af besøgs konsulenterne i forhold til hinanden? Er det fx sådan, at den kvantitative opgørelse overdømmer den kvalitative (selv om det modsatte var hensigten fra kontrahentens side)?*
- d) *Tjener vurderingsskemaets kvantitative og kvalitative opgørelser i det hele taget undersøgelsens formål?*

Kontrahenten ved AMI og CASA

Angående det metodiske spørgsmål om, hvilken status den kvalitative og den kvantitative vurdering har i forhold til hinanden, er kontrahenten ved AMI og CASA's svar, at den kvantitative del er »et led i metodeudviklingen«, dvs. et rent hjælperedskab, som ikke var ment til at blive brugt som data. Derfor burde den måske ikke have været rapporteret.

Besøgs konsulenter

De fleste konsulenter vurderer, at den kvantitative og den kvalitative vur-

dering supplerer hinanden, og at det derfor er vigtigt at bevare begge. Hvis der skulle vælges mellem de to vurderinger, foretrækker de fleste helt afgjort den kvalitative (kun en konsulent vil eventuelt droppe den kvalitative vurdering og nøjes alene med den kvantitative), idet de mener, at den giver supplerende oplysninger i forhold til de faktuelle forhold, som den kvantitative vurdering måler. Den kvalitative vurdering indeholder vigtig information om kvaliteten og seriøsiteten af arbejdsmiljøarbejdet.

Nogle konsulenter mener imidlertid, at den kvalitative vurdering ikke kan stå alene, dvs. at det er vigtigt »at have nogle tal« i form af den kvantitative vurdering. En konsulent er meget kritisk over for brugen af den samlede kvantitative skala og anbefaler, at der i stedet bruges enkeltspørgsmål fra vurderingsskemaet. En konsulent synes, at skalaerne i de enkelte spørgsmål i vurderingsskemaet og den kvantitative og den kvalitative vurderingsskala er for unuancerede og ønsker en større detaljeringsgrad. En konsulent understreger, at det på trods af metodiske vanskeligheder er prisværdigt at bruge denne type metoder ved arbejdsstedsbesøg, bl.a. fordi det ellers vil være umuligt at systematisere vurderinger og sammenligne resultater over tid.

Den faglige (kollegiale) vurdering

Visionsrapporterne indeholder ingen argumentation for at anvende både en samlet kvantitativ og kvalitativ opgørelse, ej heller en diskussion af, hvordan disse to skalaer forholder sig til hinanden. Den samlede kvantitative skala anvendes ikke i rapporterne; i enkelte rapporter anvendes dog enkeltspørgsmål fra vurderingsskemaet, som for de fleste konsulents vedkommende danner grundlag for den kvantitative vurdering. Den kunne formodes at være et hjælpemiddel til at foretage den kvalitative vurdering. Det virker umiddelbart unødvendigt med den samlede kvantitative vurdering.

Samlet vurdering

Bestilleren efterlyser et tydeligere fokus på forebyggende aktiviteter i forbindelse med valg af temaer og udtrykker kritik af omfanget af spørgsmål om eksponeringer og generelle temaer i forhold til omfanget af

visionsspecifikke forhold. Kontrahenten udtrykker generel tilfredshed med temavalget, både hvad angår forebyggende aktiviteter og visionspecifikke temaer og har ikke umiddelbart forslag til yderligere temaer. Ifølge den kollegiale vurdering er temavalget i overensstemmelse med temaer nævnt i Arbejdstilsynets udbudsmateriale. Dog er der lagt mindre vægt på forebyggende aktiviteter.

Ifølge kontrahenten er den kvalitative vurdering central; den kvantitative del er et hjælperedskab og »et led i metodeudviklingen«. Kontrahenten gør ikke yderligere rede for de to instrumenters forhold til hinanden. Besøgskonsulenterne har brugt interviewguider primært til at svare på enkelte spørgsmål i vurderingsskemaet og dermed til den samlede kvantitative vurdering. Ved den kvantitative vurdering fulgte de fleste konsulenter denne faste procedure (som er dokumenteret), mens den kvalitative vurdering er »helhedsindtrykket«, som er en ekspertvurdering af kvaliteten af arbejdsmiljøarbejdet og bygger på en kombination af forskellige oplysninger under hele besøget. Ifølge konsulenterne er den kvalitative vurdering vigtigst, idet den indeholder ny information både i forhold til den kvantitative vurdering og telefoninterviewene. Konsulenterne har ikke kendskab til brugen af besøgsdata i analyserne hvad angår validering.

Temavalget på visionsområderne følger Arbejdstilsynets forslag til temaer i udbudsmaterialet, men der kan sættes spørgsmålstejn ved, om temaerne er operationaliseret på en sådan måde, at de afspejler visionspecifikke forbyggende aktiviteter. Vægten synes i højere grad dels at ligge på det generelle arbejdsmiljøarbejde (fylder over halvdelen af telefoninterviewspørgsmålene), dels på kortlægning af arbejdsmiljøforhold, herunder eksponering/risikofaktorer, selv om det sidstnævnte kun undtagelsesvis var bestillerens ønske.

Visionsrapporterne indeholder ingen argumentation for at anvende både en samlet kvantitativ og kvalitativ opgørelse, ej heller en diskussion af, hvordan disse to skalaer forholder sig til hinanden. Den samlede kvantitative skala anvendes ikke i rapporterne; i enkelte rapporter anvendes dog enkeltspørgsmål fra vurderingsskemaet, som for de fleste konsulents vedkommende danner grundlag for den kvantitative vurdering. Den kvalitative vurdering er problematisk, idet den i høj grad bygger på en subjektiv vurdering og derfor kan forventes at variere fra konsulent til

konsulent, og fordi der ikke findes en tilsvarende måleparameter i telefoninterviewene. Derfor er den ikke egnet til at blive brugt til validering; derimod kan den bruges uddybende som en indikator på kvaliteten af arbejdsmiljøarbejdet i virksomhederne.

4.2 **Fordelingen af generelle og visionsspecifikke spørgsmål**

- a) *Er fordelingen af generelle og visionsspecifikke spørgsmål i såvel telefonspørgeskemaerne som i vurderingsskemaer hensigtsmæssig?*
- b) *Kan de visionsspecifikke spørgsmål udvides og de generelle formindskes?*

Kontor for Effekt og Analyse i Arbejdstilsynet

Kontor for Effekt og Analyse angiver, at de ikke har ønsket en opdeling i et generelt tema for alle visioner og visionsspecifikke temaer; visionerne har hele tiden været vigtigst for Arbejdstilsynet. Kontoret savner, at resultaterne fra de generelle spørgsmål (som rettelig bør benævnes fælles spørgsmål for de syv visionsområder, da der ikke er tale om en generel undersøgelse af arbejdsmiljøarbejdet), som bl.a. omhandler organisering af arbejdsmiljøarbejdet, herunder arbejdspladsvurderinger i virksomhederne, analyseres i forhold til svarene på de visionsspecifikke spørgsmål, således at man fik en kobling mellem de to sæt af data. Arbejdstilsynet mener, at der ikke er arbejdet tilstrækkeligt med de visionsspecifikke spørgsmål i forbindelse med udformningen af skemaer, og at fælles spørgsmål kun har en værdi, hvis de kobles med de visionsspecifikke resultater.

Overvågningsudvalget

I Overvågningsudvalget er DA's synspunkt, at man ikke skal spørge om noget, man ikke bruger senere i analyserne. Hensigten var at anvende de generelle spørgsmål om arbejdsmiljøarbejde til at forklare resultater fra det visionsspecifikke område; de to sæt af spørgsmål skulle krydses med hinanden, men er ikke set brugt sådan. De øvrige medlemmer har ingen

klar holdning til spørgsmålet, men de synes, at det er vigtigt at have generelle spørgsmål med i skemaet.

Visionsansvarlige i Arbejdstilsynet

De visionsansvarlige i Arbejdstilsynet er gennemgående af den opfattelse, at der er for mange fælles spørgsmål og for få visionsspecifikke spørgsmål. For støjvisionen er vurderingen dog, at fordelingen er hensigtsmæssig. Flere understreger, at om de mange fælles spørgsmål er relevante afhænger af, hvordan man vil anvende dem i analyserne.

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA vurderer, at fordelingen af fælles spørgsmål og visionsspecifikke spørgsmål var hensigtsmæssig, men siger også, at koblingen mellem de generelle og de visionsspecifikke spørgsmål er mangelfuld i de eksisterende rapporter.

Den faglige (kollegiale) vurdering

De generelle spørgsmål udgør over halvdelen af alle spørgsmål i telefoninterviewet. Under hensyn til, at det generelle arbejdsmiljøarbejde ikke var nævnt i Arbejdstilsynets udbud og i betragtning af, at de generelle spørgsmål er lidt benyttet i de visionsspecifikke rapporter, mener vi, at antallet af generelle spørgsmål bør reduceres. Dette forslag skal også ses i lyset af, at vi har vurderet, at der er for mange spørgsmål om den formelle organisering af arbejdsmiljøarbejdet og for få om de gennemførte forebyggende aktiviteter.

Samlet vurdering

Alle adspurgte parter, bortset fra kontrahenten ved AMI og CASA, har det synspunkt, at de generelle spørgsmål fylder for meget i forhold til de visionsspecifikke. Bestilleren har ikke ønsket en opdeling i et generelt tema for alle visioner og visionsspecifikke temaer. Det nævnes, at de generelle spørgsmål har deres berettigelse alene, hvis de sammenholdes med de visionsspecifikke spørgsmål, hvilket i begrænset omfang er tilfældet i visionsrapporterne. De generelle spørgsmål udgør over halvdelen af alle spørgsmål i telefoninterviewene.

De generelle (fælles) spørgsmål udgør en betydelig del af telefoninterviewene, hvilket ikke er i overensstemmelse med, at undersøgelsens formål var at sætte fokus på visionsspecifikke forhold på de syv udvalgte områder (jf. Arbejdstilsynets udbudsmateriale). Problemet forstærkes yderligere ved, at de fælles og de visionsspecifikke spørgsmål ikke er blevet sammenholdt med hinanden i analysefasen. Evaluatorens vurdering er, at de generelle spørgsmål med stærkt fokus på den formelle organisering af arbejdsmiljøarbejdet og relativt lidt fokus på forebyggende aktiviteter har en for høj prioritering.

4.3 **Spørgsmålenes (i telefoninterviewene) relevans for de udvalgte temaer**

Arbejdstilsynet angiver, at det ikke altid har været muligt at belyse områderne: Viden, holdninger og handlinger på arbejdsstederne.

- a) I hvilken udstrækning kan spørgsmålene belyse disse områder?*
- b) Er de anvendte spørgsmål relevante for at belyse de udvalgte temaer?*
- c) Kan de enkelte spørgsmål anvendes til at belyse en fremdrift, eller kan de anvendes uddybende?*

Ad a) Kan spørgsmålene belyse viden, holdning og handling?

Kontor for Effekt og Analyse

Kontor for Effekt og Analyse udtrykker, at viden, holdning og handling kun i ringe omfang dækkes af spørgsmålene i spørgeskemaerne. Kontrahenten har ikke tænkt i tredelingen, hvilket bl.a. kan skyldes, at der har været mange aktører, der har fremsat ønsker om emner, der skulle belyses, herunder de visionsansvarlige i Arbejdstilsynet. Ansvarer er delvis Arbejdstilsynets i og med, at undersøgelsens forståelsesramme er blevet mere tydelig undervejs i forløbet. Viden var nævnt i udbudsmaterialet, men det er ifølge kontoret et af de områder, der er mindst tilgodeset i spørgeskemaerne. Ud over udbudsmaterialet foreligger der bl.a. et indikator-katalog, som kontrahenten har kendskab til.

Visionsansvarlige i Arbejdstilsynet

De visionsansvarlige i Arbejdstilsynet har i begrænset omfang taget stilling til, om viden, holdning og handling var dækket i skemaerne. For visionerne psykisk arbejdsmiljø og børn og unge nævnes, at det kun i mindre grad har været tilfældet.

Den faglige (kollegiale) vurdering

Vi har foretaget en kategorisering af samtlige spørgsmål på viden, holdning og handling. Hovedindtrykket er, at antallet af spørgsmål i de tre grupper varierer en del mellem visionsområderne. Der er dog kun få spørgsmål, der belyser holdning, og på de fleste visionsområder er der flere spørgsmål om handling end om viden. Der findes faktisk kun holdningsspørgsmål på områderne psykisk arbejdsmiljø og ulykker, og bortset fra området børn og unge er antallet af vidensspørgsmål begrænset. Der er flest handlingsspørgsmål på områderne kemi og støj. På ulykkesområdet er alle spørgsmålene holdningsspørgsmål.

Ad b) Er de anvendte spørgsmål relevante for at belyse de udvalgte temaer?

Kontor for Effekt og Analyse

Kontor for Effekt og Analyse siger »ja, delvis«, men har vanskeligt ved at svare generelt på spørgsmålet. Arbejdstilsynet lægger vægt på forebyggelse, og derfor indgår begrebet også i titlen på overvågningsundersøgelsen. Men kontrahenten synes delvis at have glemt overskriften: fokus er de påvirkninger, ansatte er udsat for, og dette perspektiv har fået en fremtrædende plads i spørgeskemaet i forhold til virksomhedernes forebyggende aktiviteter. Men Arbejdstilsynet må også tage noget af kritikken på egen kappe, bl.a. fordi der har været mange fagspecialister i Arbejdstilsynet (de visionsansvarlige), der skulle høres.

Overvågningsudvalget

Spørgeskemaerne har været til en høringsrunde i Overvågningsudvalget – dog har kun en af de interviewede (repræsentanten for DA) været medlem af udvalget på det tidspunkt. DA har fralagt sig ansvaret for kemispørge-

skemaet og støjspørgeskemaet. Ligeledes stiller repræsentanterne for LO og SID spørgsmålet, om spørgeskemaet på støjområdet (og evt. på kemiområdet) kan anvendes til at måle udvikling over tid.

Visionsansvarlige i Arbejdstilsynet

Relevansen er af nogle visionsansvarlige i Arbejdstilsynet vurderet ud fra, hvad de havde behov for af informationer, mens andre vurderer spørgsmålenes relevans i forhold til det sigte, undersøgelsen havde, og de temaer, der var nævnt i udbudsmaterialet. Visionsansvarlige for henholdsvis ergonomi-, kemi- og ulykkesvisionen savner centrale emner dækket ind i spørgeskemaet, herunder en måling af tilstanden af arbejdsmiljøet inden for visionen, så man kan holde spørgsmålene om de forebyggende aktiviteter op mod omfanget af det pågældende problem.

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA vurderer, at de forebyggende aktiviteter i virksomhederne er blevet målt på en rimelig måde, men vurderer samtidig, at det er specielt arbejdsstedsbesøgene, der har været hensigtsmæssige i den henseende, og at de valgte skalaer har været hensigtsmæssige.

Den faglige (kollegiale) vurdering

I tre visionsrapporter (indeklima, kemi og støj) vurderes, at spørgsmålene er relevante for at belyse rapportens temaer. I de resterende fire visionsrapporter (børn og unge, ergonomi, psykisk arbejdsmiljø og ulykker) samt den generelle rapport er der kun tale om delvis relevans.

På områderne børn og unge, ergonomi og psykisk arbejdsmiljø er spørgsmål om forebyggende aktiviteter og foranstaltninger vægtet lavt. I ulykkesrapporten er der nævnt temaer, som ikke belyses ved spørgsmål. Med hensyn til den generelle rapport vurderes det, at der er for megen vægt på de formelle forhold omkring organisering af arbejdsmiljøarbejdet, mens de forebyggende aktiviteter ikke bliver vægtet særlig højt. Der til kommer, at »viden og kvalifikationer« fortolkes for snævert som deltagelse i arbejdsmiljøuddannelse.

Ad c) Kan de enkelte spørgsmål anvendes til at belyse en fremdrift eller kan de anvendes uddybende?

Kontor for Effekt og Analyse

Kontor for Effekt og Analyse vurderer, at få spørgsmål kan anvendes til sammenligning over tid. En mindre del af spørgsmålene er bæredygtige i denne henseende, bl.a. fordi der er spurgt til flere ting i ét spørgsmål. Dette viste Kontorets egne analyser i forbindelse med udarbejdelsen af Overvågningsrapporten 2001.

Visionsansvarlige i Arbejdstilsynet

De visionsansvarlige i Arbejdstilsynet svarer gennemgående, at det er usikkert, om spørgsmålene kan (gen)anvendes til sammenligning over tid. Især den sagkyndige for henholdsvis kemi- og ulykkesvisionen ser store vanskeligheder forbundet hermed.

Den faglige (kollegiale) vurdering

Der er ganske stor forskel mellem visionsområderne med hensyn til, hvor mange spørgsmål, vi skønner, kan anvendes til at belyse en udvikling (eller mangel på samme) over tid; antallet af spørgsmål varierer fra 3 til 13. I det generelle skema vurderes 25 spørgsmål at kunne genanvendes. I oversigten nedenfor er nævnt antallet af hovedspørgsmål i de enkelte visioner.

Oversigt over spørgsmål, der kan genanvendes

Vision	Antal spørgsmål (hovedspørgsmål), der kan genanvendes
Det generelle arbejdsmiljøarbejde	25
Børn og unge	5
Ergonomi	13
Indeklima	8
Kemi	3
Psykosocial	9
Støj	Skemaet som helhed er ikke velfungerende
Ulykker	1 (3 underspørgsmål)

I skemaet om det generelle arbejdsmiljøarbejde er det således vurderet, at ca. halvdelen af spørgsmålene kan genanvendes. I ergonomiskemaet er

det ligeledes omtrent halvdelen af spørgsmålene. I det psykosociale skema er det knap 30% af spørgsmålene. I indeklimatekemaet er det knap 20% af spørgsmålene, der vurderes at kunne genanvendes. I skemaerne fra børn og unge, kemi, støj samt ulykker vurderes andelen af spørgsmål, der kan genanvendes, lavere.

Samlet vurdering

De forskellige adspurgte aktører har forskellige opfattelser af, hvad spørgsmålene skal være relevante i forhold til. Bestiller kritiserer spørgsmålenes manglende relevans for tredelingen: viden, holdning og handling, men indrømmer, at den ikke var klart fremsat ved undersøgelsens start. I Overvågningsudvalget er der en vis skepsis over for spørgeskemaerne på støjområdet og kemiområdet. De visionsansvarlige i Arbejdstilsynet på tre visionsområder (ergonomi, kemi, ulykker) er kritiske over for manglende spørgsmål om arbejdsmiljøets tilstand, så man kan holde spørgsmålene om de forebyggende aktiviteter op mod omfanget af det pågældende problem. Kontrahenten ved AMI og CASA mener, at de forebyggende aktiviteter er målt tilfredsstillende.

Spørgsmålenes anvendelighed til at vurdere udvikling over tid er de forskellige informanter gennemgående skeptiske overfor. Ifølge den kollegiale vurdering er spørgsmålene i visionsrapporterne indeklimatekemi og støj relevante for at belyse de udvalgte temaer, mens der i de øvrige rapporter, inklusive den generelle rapport, er for lidt vægt på forebyggende aktiviteter. Holdningsaspektet er i høj grad og vidensaspektet i nogen grad underbelyst, mens der er flest spørgsmål om handling. Bortset fra områderne psykisk arbejdsmiljø og støj er antallet af spørgsmål, der kan anvendes til at belyse udvikling over tid, under 10.

Spørgsmålenes relevans for de udvalgte temaer vurderes forskelligt, hvilket bl.a. kan skyldes, at forskellige aktører har forskelligt vurderingsgrundlag. Desuden er der sket afklaring af prioriteringer i undersøgelsen undervejs i projektet. Måling af forebyggende aktiviteter har dog hele tiden været den overordnede målsætning, men opdeling i viden, holdning og handling har ikke været specificeret i udbudsmaterialet. Der er ikke i tilstrækkelig grad målt forebyggende aktiviteter, områderne viden og holdning er underbelyst, og det er kun i et mindre omfang muligt at måle

udvikling over tid. Det indebærer, at kun en del af spørgsmålene kan genanvendes i VOV2. Der har manglet systematik i metoderne mellem visionsområderne, hvilket gør områderne svært sammenlignelige.

4.4 **Telefoninterviewspørgsmålenes sproglige udformning**

Den overordnede problematik er, om spørgsmålene er sprogligt formuleret på en sådan måde, at svarpersonerne vil opfatte dem på samme måde. Sprogkonsulenten har vurderet hvert spørgsmål på dimensionerne: grad af hensigtsmæssig formulering, samt mulighed for sammenligning mellem VOV1-spørgsmål og forslag til ændringer i spørgsmål. Sprogkonsulenten har ligeledes foreslået forbedrede spørgsmålsformuleringer, hvor det er relevant.

Kontor for Effekt og Analyse

Kontor for Effekt og Analyse vurderer, at kemi- og støjskemaerne ikke er valide. Der var også det problem, at flere spørgsmål kunne forstås på flere forskellige måder. Disse ting burde have været rettet med det samme, og kontrahenten burde have foretaget et mere omhyggeligt kvalitetstjek af spørgeskemaerne, inden de blev anvendt. Kontrahenten havde oversat udenlandske skemaer eller anvendt spørgsmål fra skemaer, som den alligevel anvendte i andre sammenhænge. Det har ikke været tilfredsstillende.

Visionsansvarlige i Arbejdstilsynet

De fem visionsansvarlige, der har besvaret spørgsmålet, udpeger spørgsmål, der er uhensigtsmæssige, og som bør ændres. For ulykkesvisionen er vurderingen, at de fleste spørgsmål har manglende gyldighed, fordi svarpersonerne vil svare politisk og ikke ud fra, hvad der faktisk foregår.

Overordnet kvalitetskontrollør hos kontrahenten for VOV1

Ifølge forskningsdirektøren var det i flere tilfælde ikke muligt at anvende validerede skemaer, og skemaerne skulle udarbejdes specielt til VOV1.

Der var ikke ressourcer til en egentlig validering af spørgeskemaer, men der skete en kvalitetssikring ved pilotafprøvning. Telefoninterviewskemaet blev kvalitetssikret ved afprøvning på antagelig 25-50 personer. Forskningsdirektøren vurderer det som en kvalitet i sig selv, at man anvendte SFI Survey, som har trænedede interviewere og er indstillet på at leve op til reproducerbarhed og indtastningskontrol.

Telefoninterviewere

Under interviewet til evalueringen gennemgik telefonintervieweren hele spørgeskemaet og fortalte, hvilke spørgsmål der havde virket uhensigtsmæssige, og hvorfor de havde det. Der var forskel fra den ene telefoninterviewer til den anden, hvor godt de huskede erfaringerne fra interviewene. På ergonomiområdet havde telefonintervieweren kommentarer til flest spørgsmål, mens der kun var kommentarer til få spørgsmål på det psykosociale område.

Ifølge telefoninterviewernes erfaringer indeholdt flere spørgsmål begreber, der ikke var umiddelbart forståelige for svarpersonen, eller som havde en uklar betydning, fx »analyse«, »statistik«, »kortlægning«. Flere spørgsmål blev vurderet som for lange og for svære. Det betød, at telefonintervieweren skulle læse det samme spørgsmål op flere gange eller blev nødt til at bruge en anden formulering, før svarperson forstod det. Det kunne også betyde, at telefonintervieweren måtte læse svarkategorier op i tilfælde, hvor det ikke var instruksen. Telefonintervieweren på kemi- og støjområdet nævner, at der er forskelle mellem besvarelserne fra en given virksomhed på samme spørgsmål. Og dette skyldes, ifølge telefonintervieweren, i flere tilfælde, at svarpersonerne har forskellig viden om området. På ulykkesområdet er to telefoninterviewere blevet interviewet. De har hver især kommentarer til en halv snes spørgsmål, hvoraf tre spørgsmål går igen. Telefoninterviewerne på samme visionsområde kan således have delvis forskellige erfaringer med de enkelte spørgsmål.

Den faglige (kollegiale) vurdering

Sprogkonsulenten har vurderet hvert spørgsmål med hensyn til grad af hensigtsmæssig formulering. Det er anført, om et spørgsmål bør ændres, og der er givet forslag hertil samt vurderet muligheden for at sammenlig-

ne VOV1-spørgsmål med de nye forslag. Ved vurderingen af de enkelte spørgsmål har sprogkonsulenten anvendt en række teoretiske kriterier for »god spørgeteknik« samt erfaringer fra empirisk forskning i effekter af forskellige udformninger af spørgeskemaspørgsmål.

Hvert enkelt spørgsmål vurderes i den sammenhæng, det indgår i. Det betyder, at der ikke altid er foreslået den samme løsningsmodel til at løse tilsyneladende parallelle problemtyper i forskellige spørgsmål.

Hvert enkelt spørgsmål er ved evalueringen forsynet med to koder. En kode, der angiver den sproglige kvalitet af spørgsmålet i VOV1, og en kode, der angiver, hvorvidt de foreslåede ændringer vil gøre sammenligning mellem VOV1 og VOV2 problematisk.

Sammenfatningen af den sproglige evaluering af spørgeskemaerne er vedlagt som bilag 7.

Generelt for flere visioner er spørgsmålene præget af:

- utilstrækkelige filtergange.
- uhensigtsmæssig udformning af spørgsmål til oplæsning i interview.
- unødvendigt komplekse sætningskonstruktioner.
- unødvendigt »teknisk« ordvalg.
- inkonsistent begrebsanvendelse omkring centrale størrelser.
- urealistisk respondentopgave.

I tre af visionerne (børn og unge, kemi og støj) er der hyppigt foreslåede ændringer, som er af en sådan karakter, at sammenligning mellem VOV1 og VOV2 næppe er mulig for flere spørgsmåls vedkommende, jf. bilag 7.

Samlet vurdering

Der er rejst kritik af en del af spørgsmålene og deres mulighed for at give gyldige svar i forhold til, hvad spørgsmålene skulle belyse. Det gælder især spørgeskemaerne til visionsområderne børn og unge, kemi, ulykker og støj; støjskemaet vurderes af sprogkonsulenten at være materialets svageste.

5 Den praktiske dataindsamlingsmetode i VOV1

5.1 Praktisk tilrettelæggelse af dataindsamling

Var den praktiske gennemførelse af telefoninterviewene og arbejdsstedsbesøgene samt kontakten med arbejdsstederne hensigtsmæssig i forhold til at opfylde undersøgelsens formål?

Var træningen af besøgs konsulenterne tilstrækkelig?

Hvordan håndteredes det forhold, at besøgs konsulenterne ved nogle arbejdsstedsbesøg stod over for at skulle udvælge arbejdsområder, der skulle være i fokus på et arbejdssted, hvor der kunne være en mangfoldighed af forskellige arbejdsområder?

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA udtaler, at tidsplanen skred i dataindsamlingsfasen. Det betød, at telefoninterview og arbejdsstedsbesøg kom til at foregå parallelt og ikke først telefoninterview og derefter besøgene. Der var endvidere dårlig kobling mellem SFI Survey, som havde ansvaret for telefoninterview og Teknologisk Institut, som stod for besøgene. SFI Survey og Teknologisk Institut havde ikke sammentænkt deres tilbud. Det betød, at dokumentationen af data blev mangelfuld, og at der var en lang fase med fejlretning på grundlag af håndskrevne noter. Det havde atter til følge, at analysefasen blev udskudt.

Træningen af besøgs konsulenterne havde ifølge kontrahenten et fornuftigt omfang med en generel dag om metoder og en visionsspecifik dag. Der var øvelser med to fælles arbejdsstedsbesøg, hvor metodeudviklerne og besøgs konsulenterne var sammen. Med hensyn til at besøgs konsulenterne

terne ved nogle arbejdsstedsbesøg stod over for at skulle udvælge arbejdsområder, der skulle være i fokus på et arbejdssted, hvor der kunne være en mangfoldighed af forskellige arbejdsområder, blev det håndteret ved, at besøgs konsulenterne blev anmodet om at vurdere, hvor de største problemer var, og hvor der var indsatser på virksomhederne. Det sidstnævnte, det forebyggende, var det vigtigste.

Telefoninterviewere

Den generelle procedure ved telefoninterviewene var, at SFI sendte adresser over virksomheder med et kontaktnavn, samt brev og brochurer til telefonintervieweren, som derpå medunderskrev brevet og sendte det sammen med det andet materiale i håndskrevet kuvert til virksomheden. Ved den efterfølgende telefonkontakt spurgte telefonintervieweren så efter den relevante ledelsesrepræsentant, og ledelsesrepræsentanten blev spurgt om navnet på en sikkerhedsrepræsentant (eller evt. en anden medarbejderrepræsentant). Telefoninterviewerne mente, at denne fremgangsmåde fungerede tilfredsstillende; dog fulgte alle interviewere ikke proceduren præcist – nogle af dem ringede først til virksomheden for at få oplyst navnet på den relevante ledelsesrepræsentant, som de så sendte materialet til.

Enkelte telefoninterviewere nævnte problemer i form af, at virksomhedsadresserne kom sent og i flere omgange, og at besøg i nogle tilfælde var foretaget før telefoninterviewet, selv om det skulle have været foregået i omvendt rækkefølge. Endvidere opdagede telefoninterviewerne for virksomheder med flere arbejdssteder, at der i nogle tilfælde var uoverensstemmelse mellem telefonnummer og adresseoplysninger, og hvor alene virksomhedens adresseoplysninger passede på den branchegruppe, virksomheden var udvalgt indenfor. Flere telefoninterviewere oplevede, at svarpersoner på B-siden ikke kendte til undersøgelsen i forvejen og var måske derfor tilbageholdende med deres svar. Det var meningen, at ledelsesrepræsentanten formidlede information om undersøgelsen videre til B-sidens repræsentant.

*Koordinator*en for telefoninterviewdelen havde ansvaret for, at det aftalte antal telefoninterview blev gennemført inden for den planlagte tidsramme, og at materialet som interviewerne skulle bruge, var på plads. Interviewere fra samtlige fem geografiske distrikter, som SFI Survey den-

gang var opdelt i, medvirkede i undersøgelsen. En bestemt distriktsleder var ansvarlig for bestemte visionsområder, og de udvalgte interviewere gennemførte interviewene på et givet visionsområde i hele landet. Distriktslederne har en koordinerende rolle i forhold til interviewerne, og det var distriktslederne, der valgte interviewerne til undersøgelsen. Det centrale kriterium ved udvælgelsen var, at intervieweren havde tid til at gennemføre opgaven inden for den aftalte tidsperiode. Derudover skulle de være »gode til opgaven«, hvor distriktslederen brugte sine tidligere erfaringer med interviewere. Det var ikke et princip, at antallet af interviewere skulle være færrest muligt – i praksis var der 20-40 interviewere pr. visionsområde, og nogle af dem gennemførte kun få interview. Nogle interviewede på flere visionsområder.

Ifølge koordinatoren blev gennemførelse af telefoninterviewene vanskeliggjort ved, at udtrækket ikke var på plads, da telefoninterviewene skulle gennemføres. Det største problem var kvaliteten af virksomhedsoplysninger – koordinatoren siger på eget initiativ, at kvaliteten var til et 2-tal på skala 0-10, hvor 10 er den bedste værdi. SFI Survey havde meget ekstra arbejde med at tjekke oplysninger (der var mangler ved virksomhedens navn, telefonnummer, adresse og antal ansatte; nogle virksomheder var udtrukket to gange til samme vision) og en betydelig del af virksomhederne i udtrækket faldt fra i forbindelse med denne kvalitetskontrol. SFI Survey har ikke dokumentation for, hvor mange virksomheder, der alt i alt faldt fra, men der var behov for flere nye, supplerende udtræk, så koordinatoren fik adresser 30-40 gange fra Arbejdsmiljøinstituttet, som stod for udtrækket. Grunden til, at materialet ikke var »godt nok«, er ifølge koordinatoren, at man havde valgt at bruge materiale fra andre, tidligere undersøgelser. En anden type frafald var, at virksomhedsoplysninger ikke viste sig korrekte, da virksomheden blev kontaktet af telefonintervieweren. Desuden var der virksomheder, der ikke ønskede at medvirke. Telefoninterviewerne har noteret årsager til disse typer frafald, og de er videresendt til Arbejdsmiljøinstituttet sammen med datamaterialet. Mangler ved virksomhedsoplysninger betød, at interviewene ikke kunne gennemføres inden for den ønskede tidsperiode, og dette voldte problemer med at planlægge og gennemføre arbejdsstedsbesøgene. SFI udtrak 10% af de

virksomheder, hvor et telefoninterview var gennemført, til arbejdsstedsbesøgene.

Ifølge koordinatoren er gennemførelse af en virksomhedsbaseret undersøgelse ikke helt uproblematisk, da der sker mange ændringer i virksomheder (de flytter/lukker/fusionerer/ændrer navn), og registre (også DS og CVR) vil i et vist omfang have uopdaterede oplysninger. Det er vigtigt at tage højde for denne problematik, når denne type undersøgelse planlægges, og ressourcerne vurderes.

Med hensyn til forskel i deltagelsen mellem A- og B-side vurderer koordinatoren, at den ikke kan undgås, når kontakten med B-siden var gennem ledelsesrepræsentanten. I nogle virksomheder tør medarbejderrepræsentanten ikke deltage; i andre virksomheder prioriteres det ikke af ledelsen. Ifølge koordinatoren vil en acceptabel forskel i deltagelsen være ca. 10%.

Besøgskonsulenter

Projektlederen for besøgsdelen havde lavet aftaler med virksomhederne for de fleste konsulents vedkommende; kun på visionsområdet psykisk arbejdsmiljø var det konsulenterne selv, der lavede en del af aftalerne. Generelt foregik de fleste besøg som aftalt: konsulenterne blev »pænt modtaget«, virksomhedsrepræsentanter var forberedt på besøget, og besøgene blev gennemført, så de omfattede et interview om det generelle arbejdsmiljøarbejde, en rundvisning på arbejdsstedet og et interview om de visionsspecifikke forhold. Besøgskonsulenterne har imidlertid også gjort erfaringer, som det er vigtigt at få drøftet og afklaret, så de praktiske forhold ved besøgene kan blive mere hensigtsmæssige og systematiske. Der er dels tale om generelle erfaringer på tværs af alle syv visionsområder, dels om erfaringer, som er specifikke for et bestemt område.

Det viste sig, at nogle virksomheder havde glemt besøget eller *ikke havde afsat den fornødne tid*. Nogle gange var begge virksomhedsrepræsentanter ikke til stede (især medarbejderrepræsentant), eller den ene part var til stede en del af besøgstiden (typisk ledelsesrepræsentant). Med hensyn til konsulenternes arbejdstid var besøgene ikke altid placeret hensigtsmæssigt i forhold til hinanden. Det kan bl.a. skyldes, at alle adresser ikke var tilgængelige, da besøgene skulle planlægges. På børn og unge-området

blev besøgene gennemført i kontortiden, hvor børn og unge ikke var på arbejde.

Flere konsulenter nævner, at nogle af de virksomheder, som de besøgte, ikke var relevante med hensyn til visionsområdet. De aftalte besøg blev imidlertid gennemført, men konsulenterne ved ikke, om de er med i det endelige datamateriale. Fx var der virksomheder, som ikke var relevante i forhold til EGA, selv om de var udvalgt til besøgsdelen for visionsområdet EGA/tunge løft. Tilsvarende var der virksomheder, der aldrig havde haft børn og unge ansat, eller virksomheder som ikke længere havde den del af produktion, som var relevant for branchegruppen. Og der var virksomheder, som havde fungeret i så kort tid (under et år), at det ikke var muligt at gennemføre vurderingen i sin helhed.

Endvidere rejste besøgene i nogle tilfælde spørgsmål om *virksomheds/arbejdsstedsbegrebet* i undersøgelsen, herunder hvordan virksomhederne var udvalgt til undersøgelsen, og hvorfor telefoninterviewene var blevet gennemført. I nogle tilfælde var besøgsadressen i administrationen, mens produktionen foregik et helt andet sted; og det var produktionen, som var relevant for visionen og burde have været i fokus både ved telefoninterviewene og besøgene (det gjaldt fx for kemi- og ulykkesvisioner). I andre tilfælde omfattede besøget en mindre del af virksomheden, mens telefoninterviewene omfattede hele virksomheden. Det sidstnævnte gjaldt for flere store virksomheder og betød, at respondenterne systematisk var forskellige ved telefoninterviewene og besøgene. På nogle visionsområder var størstedelen af de besøgte virksomheder store.

Nogle konsulenter oplevede, at virksomheden *ikke var velinformeret* om undersøgelsen. I enkelte tilfælde, typisk i små virksomheder, oplevede virksomheden besøget som kontrol.

Konsulenterne vurderer, at *træningen* i brugen af vurderingsmetoder var nødvendig, men ikke efter alles mening tilstrækkelig. Nogle ville have ønsket sig endnu mere systematisk træning fx i form af flere pilotbesøg, hvor metoderne blev afprøvet såvel i større som i mindre virksomheder. Et enkelt pilotbesøg på ulykkesområdet var ikke særlig relevant, og en enkelt af de interviewede konsulenter var slet ikke med på et pilotbesøg. Flere konsulenter understreger, at de ikke kunne have gennemført besøgene uden en solid arbejdsmiljøbaggrund med kendskab til lovene, APV o.l. Nogle

konsulenter oplevede, at virksomhedsrepræsentanter netop kritiserede telefoninterviewere, at disse ikke var professionelle, dvs. de manglede kendskab til arbejdsmiljøet.

Flere konsulenter beskriver, at *vurderingsmetoderne* ikke var færdigudarbejdet/ udviklet, da besøgene gik i gang. Fx blev metoderne på børn og unge-området sendt til konsulenterne i sidste øjeblik. På andre områder – fx ergonomi – var der ting, som ikke var blevet rettet til, selv om nogle af manglerne var blevet diskuteret på træningsmødet. Generelt havde konsulenterne mange kommentarer til metoder på træningsmøderne, men alle fejlene var ikke blevet rettet, bl.a. på grund af tidsnød. Det gjaldt især de visionsspecifikke spørgsmål. Nogle af problemerne blev først opdaget undervejs, og de kunne måske være undgået, hvis skemaerne var blevet afprøvet grundigere. Alene den generelle del af skemaerne var blevet afprøvet af kontrahenten. Nogle af spørgsmålene var ubehagelige at stille, idet de tog udgangspunkt i negative forhold. Nogle af spørgsmålene var ulogiske, idet de gik ud fra, at der har været et problem, og det var ikke muligt at vælge svarkategori »irrelevant«. Flere konsulenter har opdaget skalaproblemer især i forbindelse med den kvantitative vurdering, og de fandt databaseprogrammet lidet brugervenligt.

Konsulenternes rolle var praktisk gennemførelse af besøgene, hvilket ifølge flere konsulenter var »relativt begrænset« i forhold til, at de i starten havde en forestilling om, at deres rolle var en mere integreret del af projektet med et tættere samarbejde med forskerne og inddragelse af konsulenternes ekspertise på de respektive visionsområder. Flere konsulenter nævner, at de blot udførte besøgene, men at erfaringerne fra besøgsdelen aldrig blev samlet op. Efter træningen var overstået, var dialogen mellem konsulenterne og forskerne minimal. *Sparring mellem den fagligt ansvarlige og konsulenterne* undervejs i projektet via e-mail-grupperne fungerede ikke efter hensigten.

Projektlederen for besøgsdelen nævner flere forhold, som hun blev opmærksom på i forbindelse med den praktiske koordinering af arbejdsstedsbesøgene:

1. *Tidsproblemer*: Tidsrammen for telefoninterviewene blev ikke over-

holdt, og derfor opstod der problemer med at planlægge besøgene på en hensigtsmæssig måde. Problemerne var størst på to visionsområder, børn og unge og ulykker. Forsinkelserne var frustrerende og demotiverende for konsulenterne. Situationen blev vanskeliggjort ved, at SFI Survey ikke informerede om forsinkelser på en behørig måde; ifølge projektlederen var kommunikationen fra SFI Surveys side »totalt fraværende«.

2. *Samarbejde med virksomheder*: Den oprindeligt planlagte besøgstid på 5-6 timer fra forskernes side var »helt urealistisk«. Virksomhederne ønsker ikke at afsætte så lang tid til besøget, og i praksis aftalte projektlederen en kortere tidsramme. Det er evaluatorens indtryk, at projektlederen »overtalte« virksomhederne til at deltage ved at aftale en kortere besøgstid, men samtidig skete det i praksis, at konsulenterne brugte mere tid pr. besøg, end det var aftalt – og måske derfor kunne begge siders repræsentanter ikke altid være med hele besøgstiden. *Informationsmateriale*: Projektlederen betegner det brev, som blev sendt til virksomhederne sammen med en folder om projektet, »formelt og kortfattet«. Hun har ikke fået tilbagemeldinger om dette materiale.
3. *Samarbejde mellem besøgs konsulenterne og forskerne*: Set ud fra projektlederens ståsted udviklede samarbejdet mellem konsulenterne og forskerne sig ikke konstruktivt. Det blev »et mere splittet end samlet hold«. Der var en opsplitning mellem dem, der udviklede metoderne og dem, der brugte metoderne. Konsulenterne blev ikke inddraget i projektet – fx oplevede de ikke, at der var plads til deres kommentarer til metoder på de enkelte visionsområder, selv om de var specialister på de pågældende områder. Som et eksempel nævner projektlederen ulykkesområdet, hvor »akademikere« overhørte konsulenternes forslag.
4. *Udvælgelse af virksomhedsrepræsentanterne*. Da projektlederen skulle aftale arbejdsstedsbesøgene, blev hun opmærksom på, at det ikke var entydigt, hvem der var valgt som svarperson ved telefoninterviewene, og om man ved telefoninterviewene havde været opmærksom på forskellen mellem »virksomhed« og »arbejdssted«. Ifølge projektlederen havde man »flere gange« telefoninterviewet en forkert person – fx en sikkerhedsrepræsentant for kontorpersonalet i København, mens produktionen lå »et helt andet sted«. Dette problem var ifølge projektlede-

ren størst på medarbejdersiden og opstod som oftest i de store virksomheder. Nogle gange aftalte projektlederen andre kontaktpersoner end dem, der var blevet interviewet, men hun konstaterer ved interviewet, at hun måske ikke burde have gjort det, fordi hun ikke gjorde det systematisk.

5. *Metoderne* – interviewguider og vurderingsskemaer – blev kritiseret af konsulenterne, fx at scoringen/skalaerne ikke var på plads. Efterfølgende viste det sig, at konsulenterne ikke havde udfyldt vurderingsskemaerne systematisk, bl.a. manglede der svar. Der fulgte en arbejdsproces, hvor en studentermedhjælper gennemgik skemaerne og rekonstruerede de manglende svar sammen med konsulenterne.

Samlet vurdering

Kontrahenten, telefoninterviewere og besøgs konsulenter samt projektlederen for besøgsdelen og koordinatoren for telefoninterviewdelen udtaler sig om tidsproblemer og koordineringsproblemer, hvilket vanskeliggjorde gennemførelsen af telefoninterviewene, arbejdsstedsbesøgene og analysefasen. Endvidere peger koordinatoren for telefoninterviewdelen på alvorlige mangler ved virksomhedsoplysninger i udtrækket fra kontrahenten, hvilket bl.a. havde som konsekvens, at udtrækket skulle suppleres utallige gange. Ligeledes nævner telefoninterviewere mangler ved virksomhedsoplysninger samt information om undersøgelsen til B-siden. Træning af besøgs konsulenterne vurderes gennemgående positivt både af kontrahenten og konsulenterne. Derimod nævner konsulenterne flere forhold, som var utilfredsstillende ved besøgene, bl.a. tidsanvendelse pr. besøg, relevansen af det besøgte arbejdssted, information om undersøgelsen i små virksomheder. Mangler ved vurderingsmetoderne og mangelfuld sparring undervejs i projektet samt mangelfuldt samarbejde og koordinering med det øvrige projekt nævnes som barrierer for at kunne gennemføre dataindsamlingen tilfredsstillende både af konsulenterne og projektlederen for besøgsdelen.

Der konstateres problemer med undersøgelsens gennemførelse, herunder mangler ved planlægning, koordinering og styring af hele projektet samt de to delprojekter. Der er tale om forhold, som kan forventes at have betydning for datakvaliteten. Angående telefoninterviewene handler det

bl.a. om kvaliteten af datamaterialet på de syv visionsområder og om, hvorvidt de rigtige virksomheder og virksomhedsrepræsentanter blev interviewet samt om B-sidens deltagelse (systematisk lavere end A-sidens). Angående arbejdsstedsbesøg handler det bl.a. om manglende systematik og ensartethed med hensyn til, hvem der var respondenterne, og hvordan metoderne blev brugt.

5.2 **Variation mellem besøgs konsulenter ved opgørelserne fra arbejdsstedsbesøgene**

Var observatørvariationen mellem arbejdsstedsbesøgene truende for observationernes pålidelighed?

Hvad blev gjort for at sikre ensartet praksis inden for de enkelte visioner?

Kontrahenten ved AMI og CASA

Kontrahenten ved AMI og CASA udtaler, at øvelserne viste samstemmende resultater, dvs. der var ikke store afvigelser mellem besøgs konsulenternes vurderinger på de arbejdsstedsbesøg, der blev foretaget i forbindelse med træningen af besøgs konsulenterne.

Besøgs konsulenter

Ifølge projektlederen for besøgsdelen var kriteriet for udvælgelsen af besøgs konsulenter, at de havde en vis erfaring på visionsområdet (minimum tre år), fx var ergonomi-konsulenterne typisk ergoterapeuter eller fysioterapeuter. Der var to områder, hvor konsulenterne var anderledes end ellers i projektet. Konsulenterne på støjområdet var mere teknisk orienteret end konsulenterne ellers, og konsulenterne på børn og unge-området var personer, der har kendskab til reglerne på området, men ikke tre års erfaring. Besøgs konsulenterne blev udvalgt på basis af indsendte cv'er; i princippet to konsulenter fra Vestdanmark og to fra Østdanmark pr. visionsområde.

Der var fælles træning for konsulenterne i starten af projektet, og der var et pilotbesøg henholdsvis for østkonsulenterne og vestkonsulenterne

på de enkelte visionsområder, men konsulenterne blev ikke samlet senere i projektet. De havde mulighed for at stille spørgsmål til den fagligt ansvarlige på sit visionsområde via e-mail – det var en forsker hos kontrahenten, der var udnævnt som fagligt ansvarlig. Konsulenternes vurdering på flere visionsområder er, at disse e-mail-grupper *ikke* fungerede efter hensigten. Den mest almindelige form for sparring var, at konsulenterne selv tog kontakt med en anden konsulent i tvivlsspørgsmål.

Konsulenterne har delte meninger om, i hvilket omfang der er forskelle mellem vurderinger på det samme visionsområde. Nogle er af den opfattelse, at vurderingerne er rimelig ens, mens andre mener, at der er forskelle, og atter andre synes ikke, at de kan give et generelt svar. De fleste konsulenter mener ikke, at de er i stand til at vurdere, om eventuelle forskelle har betydning for datakvaliteten. Konsulenterne har ikke fået tilbagemeldinger om eventuelle *interobservatørvariationer* fra projektet, så deres svar bygger på deres egne indtryk dels på basis af kontakter med andre konsulenter, dels på basis af egne besøg.

De fleste konsulenter mener, at der kan være forskelle i enkeltspørgsmål i vurderingsskemaet, men at disse forskelle ikke har en større betydning for den samlede *kvantitative* vurdering. Derimod formoder de, at variationen mellem konsulenterne sandsynligt er større med hensyn til den *kvalitative* vurdering. Endvidere nævner konsulenterne fx på støj- og ulykkesområdet en anden type problem, som kan betyde variation i de kvantitative vurderinger, nemlig at den kvantitative skala er hårdere i forhold til små virksomheder, og de ligger derfor systematisk lavere end større virksomheder.

Konsulenternes forskellige praksis i brugen af den kvantitative og den kvalitative skala er blevet beskrevet tidligere. Nogle konsulenter beskriver, at de brugte hele skalaen, mens andre var mere tilbøjelige til at bruge de midterste skalapoint. En konsulent vurderer, at han systematisk var hårdere i sine vurderinger end fx den fagligt ansvarlige på visionsområdet.

Endvidere peger flere konsulenter på en anden type variation, som de er blevet opmærksomme på undervejs i projektet, nemlig *intraobservatørvariation*. Flere har oplevet, at deres egen referenceramme flyttede sig undervejs, dvs. de vurderede de første virksomheder anderledes end de

sidste. Således udviklede konsulenterne deres egen kalibrering af skalaerne, og den kan være forskellig fra konsulent til konsulent.

Flere konsulenter understreger betydningen af træning og systematisk gennemførte pilotbesøg inden selve besøgene bliver gennemført. Derudover vil det være vigtigt at have en velfungerende sparring mellem alle konsulenterne på et visionsområde undervejs i projektet. En måde at minimere intraobservatørvariation på vil være, at alle konsulenterne gennemgår deres eget besøgs materiale sidst i projektet – i VOV1 havde man ikke denne mulighed, fordi konsulenterne skulle indtaste oplysningerne fra arbejdsstedsbesøgene løbende.

Samlet vurdering

Kontrahenten hos ledelsen for VOV1 mener på baggrund af træningssessionerne ved projektets start, at der ikke kan forventes større interobservatørvariation mellem konsulenterne på de enkelte visionsområder. Besøgs-konsulenterne mener, at den kvalitative vurdering i højere grad er påvirket af variation mellem konsulenterne, men de kan ikke vurdere omfanget af denne variation eller dens eventuelle betydning for resultaternes pålidelighed. Konsulenterne har også bemærket, at deres egen referenceramme flyttede sig undervejs i projektet. Interobservatørvariation er ikke forsøgt analyseret i undersøgelsen.

Konsulenternes vurdering er, gennem konkrete erfaringer, at der er variation både mellem forskellige konsulents vurderinger og internt i én konsulents vurderinger og især i den kvalitative vurdering. Konsulenterne har foretaget vurderingerne forskelligt (de mener ikke, at der har været en fast procedure fra projektets side), og de har brugt skalaerne forskelligt. Variation er ikke analyseret af kontrahenten. Sikring af større ensartethed/minimering af observatørvariation har ikke været prioriteret i projektet, fx har kollegial sparring mellem konsulenter og projekter været meget sparsom.

6 Vurdering af visionsrapporter

I det følgende vurderes de foreliggende rapporters forskningsmæssige kvalitet, deres relevans og deres formidlingsværdi. Den faglige vurdering er foretaget for hver visionsrapport for sig. Her bringes de gennemgående konklusioner.

Kontor for Effekt og Analyse i Arbejdstilsynet

Kontor for Effekt og Analyse mener, at undersøgelsen giver nogle data, der er unikke, også i international sammenhæng. Kontorets repræsentanter vurderer, at man antagelig kan bruge de indsamlede data fra større virksomheder, men næppe fra mindre og mellemstore virksomheder. Metode- og materialebeskrivelser holder en for lav kvalitet for kontrahenten som er et forskningsinstitut. Derfor kan man være betænkelig ved at bruge resultaterne som helhed. Men debatten har fået et løft. Arbejdstilsynet kunne ønske sig en langt bedre udnyttelse af data end de foreliggende rapporter. Derfor har Arbejdstilsynet fået foretaget analyser af deldata fra undersøgelsen til den årlige overvågningsrapport, som kontoret udarbejder.

Kontor for Effekt og Analyse giver udtryk for, at entreprenøren burde vide, hvilken kvalitet Arbejdstilsynet efterspurgte. Arbejdstilsynet ønskede en konsulentopgave gennemført på internationalt niveau (dvs. på et niveau, der vil kunne opfylde kravene til publikation i et internationalt videnskabeligt tidsskrift), men det har kontrahenten ikke levet op til. Direkte forespurgt vurderer repræsentanter for Kontor for Effekt og Analyse ved interview, at rapporterne gennemgående er til et 5-tal på 13-skalaen.

Der er mange forklaringer på, hvorfor det ikke gik bedre. Arbejdstilsynet anfører følgende:

- Organisering af projektet var vanskeliggjort af, at mange forskergrupper var involveret, og den tværgående koordinering oversteg entreprenørens kræfter.
- Tidsrammerne var stramme – men kendt fra starten – og det har muligvis vanskeliggjort dialogen mellem kontrahenten og Arbejdstilsynet.

Målgruppen for visionsrapporterne er parterne på arbejdsmiljøområdet, centralt og ned til BAR'ene (branchearbejdsmiljørådene). Arbejdstilsynet har holdt oplæg om undersøgelsen, men kun på forespørgsel. Formidlingen hæmmes af langt overskreden tidsplan og for dårlig kvalitet i entreprenørens rapportering.

Overvågningsudvalget

Overvågningsudvalget har ikke set de endelige rapporter og kan derfor ikke besvare spørgsmålet om vurdering af den forskningsmæssige kvalitet af rapporterne, men det er DA's vurdering, at kontrahenten har undervurderet opgavens omfang, at projektledelsen har været mangelfuld, samt at rapporternes formidlingsværdi er ekstremt ringe.

Det er en enstemmig vurdering i udvalget, at det er svært at bruge resultaterne i deres nuværende form. Medlemmerne efterlyser mere dybdegående analyser og vurderinger samt overblik og konklusioner (LO), mere brugervenlighed i rapporteringen (HK), og at rapporterne var mere operative (FTF) og tættere på brugernes virkelighed (KL), uden at det går ud over faglighed og saglighed. Data fra arbejdsstedsbesøgene burde i højere grad inddrages i visionsrapporterne, idet dette materiale kan underbygge resultaterne fra telefoninterviewene (SID).

Resultaterne, som Kontor for Effekt og Analyse har fremlagt i Overvågningsrapporten 2001 (udgivet i december 2002), er gode og anvendelige. Resultaterne anvendes i Arbejdsmiljørådets arbejde i forskellige undergrupper, og man håber, at undersøgelsen kan bringe debatten et stort skridt videre. Resultaterne vil bl.a. blive anvendt på en konference til efteråret om psykisk arbejdsmiljø, som Arbejdsmiljørådet arrangerer.

Organisationerne og branchearbejdsmiljørådene er blandt mulige brugere af resultaterne. Fx har DA til hensigt at anvende resultaterne politisk,

når det er hensigtsmæssigt, idet de vil kunne give et objektivi grundlag for debatten. Men der kan hurtigt ske det, at rapporterne bliver for gamle, hvis de ikke snart kommer i endelig udgave.

Kontrahenten ved AMI og CASA

Ledelsen af kontrahenten udtaler, at rapporterne er beskrivende og meget lidt analytiske. Det deskriptive har en forskningsmæssig kvalitet, men der burde laves flere multivariate analyser og bedre validering af metoderne. Videnskabeligt er der meget at gå videre med, men rapporterne er gode nok som den første rapportering. Kontrahenten har søgt og fået midler til at gå videre med analyserne i et efterfølgende projekt.

Ledelsen af kontrahenten vurderer, at undersøgelsen har betydelig praktisk relevans. Der er målinger, der ikke er set før, fx om forskellige tiltags udbredelse, og der er nye, væsentlige resultater om det generelle arbejdsmiljøarbejde. Arbejdsstedsbesøgene har kvalificerede resultater fra telefoninterview, fx om APV.

Ifølge ledelsen levede rapporterne imidlertid ikke helt op til Arbejdstilsynets udbud, de er for forskelligartede. Der manglede en klar forestilling hos de visionsspecifikke skrivegrupper om, hvordan analysen skulle gennemføres. Dertil kom, at involvering af eksperter (de visionsansvarlige i Arbejdstilsynet) med hver deres dagsorden betød, at undersøgelsen blev trukket i forskellige retninger.

Kontrahenten havde søgt Arbejdstilsynet om midler til evaluering af metoderne, men fik afslag.

Den faglige (kollegiale) vurdering

VOV1' s formål var dels at udvikle og afprøve en metode til at overvåge fremdriften i arbejdsmiljøarbejdet i virksomhederne (dvs. evt. ændringer over tid), dels at foretage analyser af de indsamlede data og rapportere resultaterne for derved at give et overblik over udviklingen i forebyggende aktiviteter på virksomhederne for hver vision med henblik på, at indsatserne fra arbejdsmiljøsystemet evt. kan justeres.

Analyser

Arbejdstilsynets udbudsmateriale indeholder forbavsende lidt om, hvilke

analyser der ønskes foretaget. Med hensyn til, på hvilket kvalitetsniveau analysen skal gennemføres, fremgår det af Kontrakt mellem Arbejdstilsynet og Arbejdsmiljøinstituttet under afsnittet om betalingsbetingelser, at det er en forudsætning, at produktet kan leve op til nationale og internationale standarder. I Udbud af overvågningen af fremdriften i 2005-handlingsplanen, januar 2000, s 6, fremgår det, at der skal laves analyser på arbejdsgiverudtalelser for sig og på ansattes udtalelser for sig inden for hvert visionsområde. Implicit fremgår det, at Arbejdstilsynet tillige ønsker resultaterne analyseret i forhold til branchegruppe, overensstemmelse mellem telefoninterview- og besøgsdata, samt virksomhedsstørrelse (antal ansatte).

Da Arbejdstilsynets udbudsmateriale ikke differentierer metodemæssigt og formidlingsmæssigt mellem de enkelte visioner, havde evaluator forventet en ensartet opbygning og analysestrategi for de syv visionsrapporter. Rapporterne er imidlertid forskellige med hensyn til, hvilke analyser der foretages. I de fleste rapporter, men ikke alle, foretages analyser, hvor svarene fra telefoninterview fordeles på branchegrupper, A- og B-side-besvarelser og virksomhedsstørrelse. Tabelopstillingerne er ikke ensartede fra rapport til rapport, og omfanget af kommentarer til tabellerne varierer betydeligt, i nogle rapporter er kommentarerne alt for sparsomme. Analyserne er gennemgående holdt beskrivende med få forsøg på fortolkninger eller forklaringer på data. Præsentationen af besøgsdata er som regel holdt i et kapitel for sig og er heller ikke ensartet på tværs af rapporterne. Uensartetheden efterlader indtryk af en svagt koordineret analyse- og skrivefase, og ses ikke begrundet i indholdsmæssige forskelle mellem visionerne. Evaluator savner i de fleste rapporter mere udførlig tekst til tabelopstillingerne og mere forklaring og perspektivering. Det gælder især den generelle rapport, hvor vi tillige havde forventet en analyse af forholdet mellem branchegrupper og arbejdspladsstørrelse, og på tværs af visionerne. Den generelle rapport har ikke karakter af en rapport, som sammenfatter resultaterne fra alle visionerne. Det var ellers at forvente efter kontrakten mellem Arbejdstilsynet og Arbejdsmiljøinstituttet, hvor der som produkt bl.a. nævnes en sammenfattende rapport.

Den gennemgående fattige kommentering af tabeller og de få forsøg på fortolkninger og forklaringer af de beskrivende resultater giver et

ufærdigt indtryk af visionsrapporterne. Som produkter fra en stor overvågningsrunde er det evaluators vurdering, at de kun nødtørftigt lever op til nationale og internationale standarder.

Udvikle og afprøve metoder til at måle udvikling over tid

Kontrahenten har udviklet og afprøvet metoder. Kontrahentens vurdering af de afprøvede metoder, efter at have afprøvet dem i praksis i stor skala, er imidlertid meget sparsom. Der er gennemgående kun beskeden diskussion af metoderne i visionsrapporterne og meget få refleksioner over, om det overordnede design og om de anvendte spørgeskemaer, vurderingskemaer og interviewguider måler så gyldigt, pålideligt og præcist, at de kan anvendes igen i en senere undersøgelse og give mulighed for sammenligning af data. Evaluator havde forventet en mere omfattende og mere grundig vurdering og diskussion af de anvendte metoder ud fra, at et overordnet sigte med undersøgelsen var at udvikle metoder, der egnede sig til at foretage sammenligninger over tid.

Rapporternes resumé samt konklusions- og diskussionskapitler

Det er vurderet, om rapporternes resumé samt konklusions- og diskussionskapitler er begrundet i og har en god sammenhæng med de resultater, der præsenteres i rapporterne.

Generelt er resumé- og konklusionsafsnittet velbegrundet i de fremlagte resultater. Men der er gennemgående ikke taget tilstrækkeligt forbehold for udtrækkets størrelse, og de usikre generaliseringsmuligheder er ikke drøftet. Resuméafsnittene er generelt dækkende for rapporternes indhold, men det er forvirrende, at der ofte er et ganske stort overlap mellem resumé- og konklusionsafsnit. Evaluator vil gøre opmærksom på, at denne vurdering alene bygger på teksten i rapporterne. Det er et andet spørgsmål, om grundlaget for analyserne er tilstrækkeligt.

Rapporternes læsevenlighed

Evaluator forudsætter, at rapporternes målgruppe er arbejdsmiljøprofessionelle og særligt arbejdsmiljøinteresserede i offentlige myndigheder og hos arbejdsmarkedets parter.

Hver for sig er rapporterne gennemgående overskueligt disponeret. Ofte vanskeliggøres læsningen imidlertid af en rapportering med mange store tabeller og kun lidt tekst. Flere resuméafsnit er holdt i en opremsende telegramstil, der vanskeliggør læsning. Det virker unødvendigt, at flere siders tekst (otte sider) går igen i alle rapporterne. Det havde været mere elegant at holde de generelle metodebemærkninger i den generelle rapport og så henvise til denne i de enkelte visionsrapporter.

Den generelle rapport er ikke formidlingsvenlig. Der er tale om et ufærdigt, ikke gennemarbejdet manuskript med mange sproglige fejl. Resultaterne er svært overskuelige og mangler fokus. Evaluatoren er opmærksom på, at der er tale om den version af rapporten, der forelå ved evalueringens start, marts 2003, som ikke var den endelige version. Af ressourcemæssige grunde og af hensyn til tempoet i evalueringen har det ikke været muligt at medtage den sidste version af denne rapport i evalueringen.

Kvalitetskontrol af VOV1

Kvalitetssikringen foregik forskelligt på de to institutter, der udgjorde styregruppen for VOV1. På det ene foregik kvalitetssikringen ved den løbende faglige sparring blandt medarbejderne. Der var ingen systematisk overordnet kvalitetskontrol. På det andet institut var der en systematisk kvalitetskontrol, som omfattede følgende:

Der var ingen faglig kvalitetssikring af selve tilbuddet. I *planlægningsfasen* vurderede forskningsdirektøren de statistiske styrkeberegninger.

Under gennemførelsen af projektet blev telefoninterviewskemaet kvalitetssikret ved afprøvning på antagelig 25-50 personer.

I rapporteringsfasen læste forskningsdirektøren rapportudkastene i en tidligere version, end den evaluator har. Tilbage meldingen blev givet til den pågældende hovedforfatter i mundtlig form. Bemærkningerne til rapporterne havde karakter af at påpege uforståelige formuleringer og diskrepanser, fx mellem telefoninterview og arbejdsstedsbesøg.

Ifølge forskningsdirektøren var rapporterne meget forskellige, men det lykkedes til sidst at få et paradigme for, hvordan de skulle udarbejdes. Der var ingen formkrav til rapporterne fra starten af projektet. Fx var det

uklart, hvordan arbejdsstedsbesøgene skulle rapporteres, og hvordan de kunne supplere telefoninterviewene.

Der blev ikke foretaget vurdering af rapportudkast ved eksterne lektører, som ellers er proceduren på det pågældende institut, idet Arbejdstilsynets fagspecialister havde rapportudkastene til høring. Forskningsdirektøren vurderede, at det ikke ville være fornuftigt at have en ekstern vurdering af rapportudkast, idet det kunne betyde en kortslutning mellem de eksterne læsere og Arbejdstilsynets fagspecialister.

Forskningsdirektørens vurdering er, at de første udkast til rapporter, han havde set, var mangelfulde, men det lykkedes trods alt at lande på et udmærket koncept for rapporterne. Der var klart problemer knyttet til selve processen, som demotiverede forskerne. På den positive side blev det bemærket, at man havde fået et ganske godt datasæt med fornuftige svarprocenter. Og dette datasæt kan, sammen med NAK-registret, sige noget om, hvordan arbejdsmiljøsystemet rent faktisk virker.

Han betragter Arbejdstilsynet snævert som målgruppe for rapporterne, men påpeger også, at målgruppen for rapporterne ikke er blevet specificeret.

Samlet vurdering

Vurderingen af de syv rapporter samt den generelle rapport peger på en række svagheder ved rapporterne. Svagheder, der på sin vis deles af de interviewpersoner, der er blevet interviewet om rapporterne. Men der er betydelig forskel på, hvor hårdt man bedømmer dem. Den ubetingede hårdeste bedømmelse er fra bestiller, Kontor for Effekt og Analyse i Arbejdstilsynet, mens ledelsen af kontrahenten for VOV1 vurderer deres produkter mildere, men alligevel med en betydelig selvkritik.

Evaluator har hæftet sig ved følgende forhold. For det første er der en påfaldende mangel på ensartethed i analyserne på tværs af rapporterne, selv om deres formelle opbygning er tilnærmelsesvis parallel. Det vanskeliggør en læsning på tværs og dermed er det svært at få et samlet overblik over resultaterne. For det andet mener vi, at givet at det ene formål med VOV1 var at udvikle og afprøve design og metoder, er der for begrænset vurdering af de anvendte metoder og beskeden refleksion og diskussion af metodernes egnethed til at gentages senere. For det tredje er rapporterne

generelt ikke lette at læse, heller ikke for en målgruppe som professionelle i arbejdsmiljøsystemet. De ujævne analyser og de sine steder manglende omtale og dermed manglende forbehold over for metodiske svagheder i rapporterne gør, at evaluator må sætte spørgsmålstejn ved, om rapporterne gennemgående lever op til den forskningsmæssige kvalitet man må stille til rapporteringen af et så stort projekt. Der gennemførtes en vis overordnet kvalitetskontrol af udkast til rapporter, men evaluator vurderer, at det var en brist i kvalitetskontrollen, at den normale vurdering af rapporterne ved eksterne lektører ikke blev foretaget.

Det er ikke evaluators opgave at komme med forklaringer på, hvad der førte til det givne resultat, da der netop ikke har været tale om en procesevaluering. Der er dog givetvis mange forklaringer på, hvad der førte til det opnåede resultat. Undersøgelsens karakter som en udviklingsopgave, undersøgelsens fordeling på flere kontrahenter og samarbejdet mellem de involverede aktører er formodentlig blandt de medvirkende faktorer. Vi har hæftet os ved, at krav og ønsker til analysernes karakter og til indholdet af visionsrapporterne ikke var helt klart i udbudsmaterialet, samt at der øjensynligt har været for mange »kokke om suppen«. Det ville kræve en usædvanlig fast hånd fra kontrahentens side at komme helskindet gennem et sådant forløb.

Det positive er, at rapporterne giver et overblik over nogle forhold, der ikke tidligere har været belyst så omfattende, og at Arbejdstilsynet har kunnet udnytte dele af datagrundlaget til Overvågningsrapport 2001.

7 Forslag til ændringer i design og metoder med henblik på VOV2

Design og metoder:

1. Såfremt der ønskes sammenligning over tid af forholdene i enkelte branchegrupper, bør udtrækket af virksomheder forøges væsentligt, eller man bør renoncere på de statistiske krav til, på hvilket niveau man vil have afdækket forskelle. Alternativt kan man sammenligne en given vision (med branchegrupperne slået sammen) over tid. Det vil give et større datagrundlag for at gennemføre en regressionsanalyse, der kan afdække fx branchegruppeeffekt, tidseffekt eller effekt af virksomhedsstørrelse.
2. Ensartede sammenligninger af A- og B-sidens besvarelser bør gennemføres for alle visioner. Overensstemmelse mellem besvarelser fra A- og B-siden bør analyseres på en gennemskuelig måde, fx ved korrelationsanalyser. Lav overensstemmelse i svarene bør fortolkes og ikke ubetinget tages som udtryk for usikre besvarelser (lav gyldighed).
3. Det er vigtigt at afklare formålet med arbejdsstedsbesøgene i forhold til telefoninterviewene med henblik på, hvordan arbejdsstedsbesøgene anvendes til at a) validere og b) supplere telefoninterviewene. Det foreslås, at validering foretages primært i forhold til enkeltspørgsmål om faktuelle forhold, hvor telefoninterview suppleres ved at stille uddybende spørgsmål om bestemte faktuelle forhold af særlig interesse for visionsområdet og ved en kvalitativ vurdering af kvaliteten af arbejdsmiljøarbejdet. Endvidere er det væsentligt, at besøgsdata analyseres på en ensartet måde på tværs af visionsområderne, og at besøgs-

data i højere grad bliver integreret i rapporterne og ikke står som et adskilt kapitel.

4. Kriterierne for, hvordan medarbejderrepræsentanten udvælges til telefoninterview bør præciseres.
5. Den lave svarprocent (under 60) i nogle branchegrupper er en trussel mod at kunne generalisere til hele branchegruppen, og dermed for sammenligning af givne branchegrupper over tid. Det foreslås derfor at gennemføre bortfaldsanalyser for disse branchegrupper vedkommende for at sandsynliggøre, om det er muligt at generalisere eller ej.

Spørgeskemaer og vurderingsskemaer:

6. Det er vigtigt at afklare, hvilke parametre fra arbejdsstedsbesøgene der bruges henholdsvis til at a) validere og b) supplere telefoninterviewene. Evaluatoren anbefaler, at den samlede kvantitative vurdering undlades ved besøgene og kun enkeltspørgsmål primært om de faktuelle forhold medtages. Den kvalitative vurdering af arbejdsmiljøarbejdet bibeholdes, og grundlaget og kriterierne for denne vurdering diskuteres i forbindelse med træningen af konsulenterne.
7. Vægtningen mellem de generelle (fælles) og de visionsspecifikke spørgsmål ændres således, at antallet af de fælles spørgsmål reduceres. Designet for VOV2 præciseres med henblik på at analysere sammenhænge mellem de generelle og de visionsspecifikke spørgsmål.
8. Udgangspunktet for revision af spørgsmål til telefoninterview bør være evaluators vurdering af, hvilke spørgsmål der er egnet til sammenligning over tid.
9. Nye spørgsmål bør udvikles under hensyntagen til, a) at der mangler spørgsmål til at afdække virksomhedens opfattelse af arbejdsmiljøets tilstand (så aktivitetsniveauet på virksomhederne kan sættes i forhold til omfanget af deres problemer), b) at der mangler spørgsmål om fo-

rebyggende aktiviteter, og om holdninger og viden. Endelig c) bør de formelle spørgsmål blandt fælles spørgsmål reduceres.

10. Flere af de spørgsmål, der genbruges fra VOV1 bør justeres efter forslag fra sprogkonsulenten. Ændringer af spørgsmål til telefoninterviewskemaet er gengivet i separat bilag, hvor sprogkonsulenten har anført forslag til ændringer af de enkelte spørgsmål i alle syv skemaer.

Tilrettelæggelse af dataindsamling:

11. Stram og konsekvent projektledelse og klar koordinering på tværs af visioner og datakilder, således at der opnås en fælles forståelse af metoderne, at tidsplan overholdes og resultaterne analyseres og rapporteres ensartet.
12. Afklaring af kriterier for valg af virksomheder til undersøgelsen og evt. eksklusion.
13. Virksomhedsoplysninger bør tjekkes, herunder om der er overensstemmelse mellem branchegruppe og adresseoplysninger ved telefoninterviewene. Det bør ligeledes afklares, hvem der er de rette lokale virksomhedsrepræsentanter.
14. Der bør ske en afklaring af, hvordan besøgene gennemføres/bruges i store virksomheder, da besøgene her kun omfatter en del af virksomheden (vedr. spørgsmålet om validering og supplerings i forhold til telefoninterviewene).
15. Ved træning af konsulenter bør der lægges vægt på pilotbesøgene, hvor man kalibrerer konsulenternes vurderinger, så der sikres størst mulig ensartethed. Træningen bør foregå i mindre, mere homogene grupper.
16. Opmærksomhed på information om undersøgelsen til virksomheder (både til A- og B-side), hvor bl.a. fortroligheden af oplysningerne understreges.

17. Sparring i forhold til vurderingsmetoderne mellem konsulenter og forskere/fagligt ansvarlige. Der bør være mulighed for en løbende kontakt mellem den ansvarlige og konsulenterne. Der bør ligeledes være mulighed for, at konsulenterne mødes efter besøgene er foretaget og eventuelt justerer vurderingerne for at bevare ensartethed i brug af vurderingsinstrumenterne.

Visionsrapporter:

18. Der bør være en klar analysestrategi og detaljeret disposition for samtlige rapporter inden analyse og skrivning påbegyndes.

19. De enkelte visionsrapporter bør være ensartet opbygget med klar påpejning af metodemæssige forbehold. Rapporterne bør indeholde en metodebeskrivelse og dokumentation af datamaterialet.

20. Det bør sikres, at der finder en ekstern kvalitetssikring af rapportudkast sted.

Bilag 1

Oversigt over skriftligt materiale anvendt i evalueringen

Materialer, der forelå før evalueringens start:

- Visionsrapporter og en generel rapport fra VOV1 (i den udgave, som de forelå ved evalueringens start primo marts 2003), i alt otte rapporter
- Spørgeskemaer anvendt ved telefoninterview i VOV1, i alt syv skemaer
- Vurderingsskemaer anvendt ved virksomhedsbesøg i VOV1, i alt syv skemaer
- Interviewguider anvendt ved virksomhedsbesøg i VOV1, i alt 21 guider
- Referater fra møder i overvågningsudvalget, som omhandlede VOV1
- Kontrakt mellem Arbejdstilsynet og SFI-Survey, marts 2001
- Udbud af dataindsamlingsdelen i projektet »Overvågning af fremdriften i 2005-handlingsplanen«, september 2000
- Kontrakt mellem Arbejdstilsynet og Arbejdsmiljøinstituttet, maj 2000
- Udbud af overvågning af fremdriften i 2005-handlingsplanen, januar 2000
- Overvågning 2001. Udviklingstendenser i arbejdsmiljøet for de syv visioner i handlingsprogrammet om et rent arbejdsmiljø år 2005. Arbejdstilsynet, 2002
- Arbejdstilsynet, 2002. Bilagsrapport – Del 1 (inkl. bilag A-F). Arbejdstilsynet, 2002.

Materialer, der er udarbejdet under evalueringen:

- Tre notater til internt brug i Kontor for Effekt og Analyse i Arbejdstilsynet:
 - Notat 1: Resultat af den kollegiale vurdering, akf, maj 2003
 - Notat 2: Resultat af interview med bestiller og forskere, akf, maj 2003
 - Notat 3: Resultat af interview med telefoninterviewere, besøgs konsulenter og kvalitetskontrollører, akf, juni 2003
- Diverse mødereferater af møder om hvert af de tre notater mellem Kontor for Effekt og Analyse i Arbejdstilsynet, projektleder af VOV2 og evaluator.

Bilag 2

Præcisering af evalueringsspørgsmål

På nuværende tidspunkt (7.3.2003) bygger præciseringen på de i projektbeskrivelsen nævnte hovedspørgsmål suppleret med en rundspørge hos de organisationer, der gennemførte VOV1, samt læsning af AT's udbudsmateriale for VOV1, samt AMI/CASA's tilbud. Endvidere har et møde med sprogkonsulenten bidraget til præciseringer, som er indarbejdet i nedenstående.

AMI/CASA meddelte, at de synes, at spørgsmålene allerede er meget omfattende og i høj grad afspejler, at der stadig er et stort metodeudviklingsbehov for VOV2. De har derfor ikke supplerende spørgsmål. Det samme gælder for SFI Survey.

TI havde nogle enkelte kommentarer, der er indarbejdet i nedenstående.

Præcisering af evalueringsspørgsmål vedr. design og metode

1. Udtræk af arbejdssteder til telefoninterview

Formålet med udtræk af en stikprøve af arbejdssteder var at etablere en første måling af forebyggende arbejdsmiljøaktiviteter på ca. 400 arbejdssteder pr. vision inden for branchegrupper, der er i fokus for arbejdsmiljøapparatets indsats. Det skulle endvidere tilstræbes at dække hver branchegruppe med 50 arbejdssteder. Kravet var endvidere, at stikprøven var stratificeret efter arbejdsstedsstørrelse, således at man ved efterfølgende målinger kunne udtale sig sikkert om ændringer på små, mellemstore og store arbejdssteder.

- a) Det skal vurderes, om kravet om mindst 50 repræsentanter i en branchegruppe giver et brugbart materiale til senere dataopgørelser.
- b) Det skal ligeledes vurderes, om fordelingen på virksomhedsstørrelse i små, mellemstore og store virksomheder bør være balanceret, forstået på den måde, at de er repræsenteret med hver en tredjedel.
- c) Er udtrækket af arbejdssteder foregået på en måde, der svarer til det ønskede formål?
- d) Hvordan er udtrækkene af arbejdssteder faktisk foretaget?
- e) I hvilken udstrækning sikrer den faktisk anvendte fremgangsmåde for stikprøveudvælgelse den ønskede repræsentativitet på branchegrupper?
- f) Hvilke konsekvenser har evt. manglende repræsentativitet i stikprøven af arbejdssteder for sammenligning mellem data fra runde 1 og runde 2?
- g) Hvilken betydning har det, at unge over 18 år blev behandlet som en del af ulykkesvisionen?

2. Filterspørgsmål

- a) Har filterspørgsmål tjent deres formål (der var ikke på forhånd fra AMI/CASA's side tænkt over filterspørgsmål)? De blev stillet for nogle visioners vedkommende, men der er aldrig bevidst opstillet et formål med dem, nemlig at udelade virksomheder, som mod forventning ikke er relevante for den pågældende vision.
- b) Der blev ikke anvendt filterspørgsmål inden for alle visionsområder, og det skal derfor vurderes, hvilke konsekvenser forskellig anvendelse af filterspørgsmål har for generaliserbarhed og for analyserne.
- c) En anden type filterspørgsmål er dem, der er brugt i spørgeskemaerne til at styre overspring af spørgsmål. Disse spørgsmål vil også blive vurderet for deres logik og hensigtsmæssighed, jf. spørgsmål 10.

3. Interview med såvel A- som B-side

I første undersøgelsesrunde skulle gennemføres interview med såvel A- som B-siden med det formål at give en mere nuanceret beskrivelse af handlingsplanens gennemslagskraft samt for at vurdere graden af overensstemmelse.

- a) Hvilken informationsværdi har det givet at telefoninterviewe såvel A- som B-side?
- b) Hvordan har man håndteret uoverensstemmelser i besvarelserne?

4. Interview med sikkerhedsrepræsentant versus anden medarbejder

I første undersøgelsesrunde var hensigten på B-siden at interviewe sikkerhedsrepræsentanter, fordi man antog, at de var bedre i stand til at besvare spørgsmål om arbejdsstedets forebyggende aktiviteter. Oprindeligt var det Arbejdstilsynets tanke at gennemføre interview med flere arbejdstagere på det enkelte arbejdssted frem for en repræsentant, men den model blev ikke valgt i forbindelse med udviklingen af VOV1.

- a) Hvad er fordele og ulemper ved at interviewe henholdsvis en sikkerhedsrepræsentant og en almindelig medarbejder?

5. Interview med ledelsesrepræsentant versus medarbejderrepræsentant

I første undersøgelsesrunde blev der i en række tilfælde kun interviewet den ene part.

- a) Hvad er fordele og ulemper ved, at der på nogle arbejdssteder er interview med kun en ledelsesrepræsentant eller med kun en sikkerhedsrepræsentant eller en anden medarbejderrepræsentant?
- b) Spørgsmålenes svarevne for henholdsvis ledelsesrepræsentant og sikkerhedsrepræsentant (her tænkes på eksponeringsspørgsmål)?

6. Validering og supplerings af telefoninterviewene ved hjælp af arbejdsstedsbesøgene

Sigtet med at kombinere telefoninterview og arbejdspladsbesøg var, at arbejdsstedsbesøgene skulle validere såvel som supplere telefoninterviewene. Validering skulle forstås således, at faktuelle forhold, konstateret ved besøgene, skulle sammenlignes med informationer fra telefoninterviewene. Vi ved imidlertid fra Teknologisk Institut (TI), at besøgs konsulenterne ikke havde kendskab til data fra telefoninterview, når arbejdsstedsbesøgene blev foretaget.

- a) Spørgeskemaerne anvendt ved telefoninterview vil blive sammenholdt med besøgs skemaet for at vurdere, i hvilken udstrækning validering af interview ved besøg er mulig. Det vil blive vurderet, hvilke spørgsmål

fra telefoninterviewene der er relevante at validere ved besøg, og om disse spørgsmål er dækket ind ved besøgsvurderingsskemaerne.

- b) Hvordan blev sammenligning mellem data fra telefonsurvey og fra arbejdsstedsbesøgene foretaget?
- c) Hvordan håndteredes manglende overensstemmelse?
- d) Hvordan blev data fra arbejdsstedsbesøgene anvendt til at supplere data fra telefonsurvey?
- e) Hvad er i det hele taget fordele og ulemper ved at anvende arbejdsstedsbesøgene til at validere og supplere telefoninterviewene?

Præcisering af evalueringsspørgsmål vedr. spørgeskemaer, interviewguider og vurderingsskemaer

7. Temaernes relevans for udviklingen samt kvantificering/kvalificering af besøgsdata

Arbejdstilsynet havde ved første undersøgelsesrunde fastlagt en række temaer (problemstillinger), der skulle belyses, og som er beskrevet i AT's udbudsmateriale. Det drejede sig om en serie spørgsmål om visionsspecifikke forebyggende aktiviteter. Derudover indeholdt nogle af spørgeskemaerne spørgsmål om risikofaktorer, idet man manglede viden om disse risikofaktorer.

- a) Er de fastlagte temaer (problemstillinger) i interviewene såvel som ved besøgene relevante for måling af udviklingen af de forebyggende aktiviteter på arbejdsstederne? Hvilke temaer er ikke relevante? Er der behov for nye temaer?
- b) Hvordan er data fra interviewguider anvendt ved arbejdsstedsbesøgene omsat til vurdering på de kvalitative vurderingsskalaer i vurderingsskemaerne?
- c) I vurderingsskemaerne skulle foretages både en kvantitativ og kvalitativ opgørelse. Den kvalitative opgørelse er beskrevet som den endelige vurdering af arbejdsstedet. Hvordan er de kvantitative og kvalitative opgørelser anvendt i praksis af besøgs konsulenterne i forhold til hinanden? Er det fx sådan, at den kvantitative opgørelse »overdømmer« den kvalitative (selv om det modsatte var hensigten fra AMI/ CASA's side)?

- d) Tjener vurderingsskemaets kvantitative og kvalitative opgørelser i det hele taget undersøgelsens formål?
- e) Er data fra arbejdsstedsbesøgene udnyttet i tilstrækkelig grad i visionsrapporterne?

8. Fordelingen af generelle og visionsspecifikke spørgsmål

- a) Er fordelingen af generelle og visionsspecifikke spørgsmål i såvel telefonspørgeskemaerne som i vurderingsskemaerne hensigtsmæssig?
- b) Kan de visionsspecifikke spørgsmål udvides og de generelle formindskes?

9. Spørgsmålenes (i telefoninterviewene) relevans for de udvalgte temaer (problemstillinger)

Arbejdstilsynet angiver, at det ikke altid har været muligt at belyse områderne: viden, holdninger og handlinger på arbejdsstederne.

- a) I hvilken udstrækning kan spørgsmålene belyse disse områder?
- b) Er de anvendte spørgsmål relevante for at belyse de udvalgte temaer?
- c) Kan de enkelte spørgsmål anvendes til at belyse en fremdrift, eller kan de anvendes uddybende?

10. Telefoninterviewspørgsmålenes sproglige udformning

Den overordnede problematik er, om spørgsmålene er sprogligt formuleret på en sådan måde, at svarpersonerne vil opfatte dem på samme måde. Sprogkonsulenten vil vurdere hvert spørgsmål på dimensionerne: grad af hensigtsmæssig formulering samt mulighed for sammenligning mellem VOV1-spørgsmål og forslag til ændringer i spørgsmål. Sprogkonsulenten vil således også foreslå forbedrede spørgsmålsformuleringer, hvor det er relevant.

- a) Hvor optræder uklare formuleringer og flertydighed?
- b) Hvor optræder vanskelige sætningskonstruktioner?
- c) Hvor optræder svært forståelige ord – i forhold til målgruppen?
- d) Hvor optræder ledende formuleringer?
- e) Forekommer spørgsmålsdesign, der giver øget risiko for utidig afbrydelse under oplæsning?
- f) Forekommer inkonsistens mellem spørgsmål og svarformuleringer?

- g) Forekommer uklarhed i filterspørgsmål og filterinstruktioner?
- h) Er der risiko for respons-bias grundet konteksteffekter?
- i) Forekommer uhensigtsmæssige spørgeforløb, fx overlap eller pludselige skift i emne eller fokus, der ikke er tydeligt markeret?
- j) Er der en logisk sammenhæng mellem spørgsmål og svarkategorier?
- k) Er svarmulighederne udtrykt med et passende almindelig brugt ord (fx bruges udtrykket fremragende som svar på nogle af spørgsmålene).

Præcisering af spørgsmål vedrørende den praktiske dataindsamlingsmetode

11. Praktisk tilrettelæggelse af dataindsamling

Var den praktiske gennemførelse af telefoninterviewene og arbejdsstedsbesøgene samt kontakten med arbejdsstederne hensigtsmæssig i forhold til at opfylde undersøgelsens formål?

Var træningen af besøgs konsulenterne tilstrækkelig?

Hvordan håndteredes det forhold, at besøgs konsulenterne ved nogle arbejdsstedsbesøg stod over for at skulle udvælge arbejdsområder, der skulle være i fokus på et arbejdssted, hvor der kunne være en mangfoldighed af forskellige arbejdsområder?

12. Interpersonel variation ved opgørelserne fra arbejdsstedsbesøgene

Var observatørvariationen mellem arbejdsstedsbesøgerne truende for observationernes pålidelighed?

Hvad blev gjort for at sikre ensartet praksis inden for de enkelte visioner, jf. fx ovennævnte om udvælgelse af arbejdsområder?

Sammenhængen i visionsrapporter

13. Vurdering af sammenhængen i de enkelte visionsrapporter

Der vil blive foretaget en vurdering af sammenhængen mellem formål, problemstilling, design, metoder, analyse og resultater i de enkelte visionsrapporter, dvs. at akf ved gennemlæsning af rapporterne vurderer, om sammenhængen mellem formål, problemstillinger, anvendte metoder, datagrundlag og konklusioner er klar og logisk.

Bilag 3

Kriterier for vurdering af visionsrapporter, af spørgeskemaer til telefoninterview og af vurderingsskemaer til arbejdsstedsbesøg

Kriterier for vurdering af visionsrapporter

Vurdering af sammenhæng i rapporterne

1. Hvor dækkende er de temaer/problemstillinger behandlet, der fremgår af AT's udbud af 7.1.2000?
 - a. Er de fastlagte temaer (problemstillinger) i interviewene såvel som ved besøgene relevante for måling af udviklingen af de forebyggende aktiviteter på arbejdsstederne?
 - b. Hvilke temaer er ikke relevante?
 - c. Er der behov for nye temaer?
2. Hvor hensigtsmæssigt er det valgte surveydesign og de valgte metoder (telefoninterview og besøg på arbejdssteder) for at besvare de nævnte problemstillinger og for at leve op til undersøgelsens formål at måle fremdrift over tid?
3. I hvilket omfang behandles begrænsninger og brister, hvad angår generaliserbarhed til de udvalgte branchegrupper af de indsamlede data i rapporten? I hvilket omfang skønner akf, at der er begrænsninger og brister?
 - a. Hvordan er udtræk af arbejdssteder faktisk foretaget?

- b. Hvilken betydning har det, at unge over 18 år blev behandlet som en del af ulykkesvisionen?
 - c. Der blev ikke anvendt filterspørgsmål inden for alle visionsområder, og det skal derfor vurderes, hvilke konsekvenser forskellig anvendelse af filterspørgsmål har for generaliserbarhed og for analyserne (spørgsmålet behandles også ved interview med AMI/CASA).
4. I hvilket omfang behandles begrænsninger og brister, hvad angår gyldighed og pålidelighed af de indsamlede data i rapporten? I hvilket omfang skønner akf, at der er begrænsninger og brister?
- a. Hvilken informationsværdi har det givet at telefoninterviewe såvel A- som B-side?
 - b. Hvordan har man håndteret uoverensstemmelser i besvarelserne?
 - c. Tjener vurderingsskemaets kvantitative og kvalitative opgørelser i det hele taget undersøgelsens formål?
 - d. Er data fra arbejdsstedsbesøgene udnyttet i tilstrækkelig grad i visionsrapporterne?
5. Er analysen foretaget på en måde, der er gennemskueligt beskrevet i rapporten? Og i hvilket omfang giver analysen svar på de problemstillinger, rapporten tager op?
6. Er konklusionerne velbegrunder ud fra de gennemførte analyser?
7. Er sammenfatningen dækkende for rapportens indhold?
8. Er rapporten klar og overskuelig struktureret, og er dens sprog og form velegnet for dens målgruppe?

Kriterier for vurdering af spørgeskemaer til telefoninterview

Sprogkonsulenten vurderer spørgeskemaerne ud fra de spørgsmål, der er nævnt i evalueringsspørgsmål 10 (spørgeskemaernes sproglige udformning). De nedenfor nævnte kriterier er dem, som akf anvender.

Spørgmålenes relevans

- a. Er de anvendte spørgsmål relevante for at belyse de udvalgte temaer?
- b. I hvilken udstrækning kan spørgsmålene belyse områderne viden, holdning og handling på arbejdsstederne?
- c. Kan de enkelte spørgsmål anvendes til at belyse udvikling over tid?

Ad c. Spørgmålenes genstand:

Hvilke spørgsmål er:

- i. Faktuelle (indeholder kun lidt skøn fra svarpersonens side, genstanden kan observeres)
- ii. Skønsmæssige (genstanden kan i princippet observeres, men udstrækning i rum eller tid gør, at skønnet ved besvarelsen bliver stort)
- iii. Vurderingsmæssige (spørgsmålet drejer sig om årsags-virkningsforhold, der er svært tilgængelig for observation, og spørgsmål som indeholder et stort element af fortolkning/vurdering).
- iv. Normative/holdningsmæssige (genstanden er svarpersonens indstilling, ønsker, prioriteringer o.l.) som ikke kan gøres til genstand for observation (andet end indirekte)

Det er antagelsen, at sammenligning over tid bedst kan foretages ud fra faktuelle spørgsmål og i mindre grad ud fra de skønsmæssige spørgsmål. Mens sammenligning over tid ikke er mulig for de vurderingsmæssige og normative spørgsmål.

Fordeling af generelle og visionsspecifikke spørgsmål

- a. Er fordelingen af generelle og visionsspecifikke spørgsmål i såvel telefonspørgeskemaerne som i vurderingsskemaerne hensigtsmæssig?
- b. Kan de visionsspecifikke spørgsmål udvides og de generelle formindskes?

Kriterier for vurdering af vurderingsskemaer til arbejdsstedsbesøg

Spørgsmålenes relevans

- a. Er de anvendte spørgsmål relevante for at belyse de udvalgte temaer?
- b. I hvilken udstrækning kan spørgsmålene belyse områderne viden, holdning og handling på arbejdsstederne?
- c. Kan de enkelte spørgsmål anvendes til at belyse en udvikling eller kan de anvendes uddybende og/eller supplerende?

Kvantitative og kvalitative opgørelser

- a. Tjener vurderingsskemaets kvantitative og kvalitative opgørelser undersøgelsens formål?
- b. Hvad er fordele henholdsvis ulemper ved, at vurderingsskemaerne indeholder både en kvantitativ og kvalitativ opgørelse?

Fordeling af generelle og visionsspecifikke spørgsmål

- a. Er fordelingen af generelle og visionsspecifikke spørgsmål i vurderingsskemaerne hensigtsmæssig?
- b. Kan de visionsspecifikke spørgsmål udvides og de generelle formindskes?

Sammenhold af spørgeskema med vurderingsskema

Spørgeskemaerne anvendt ved telefoninterview vil blive sammenholdt med besøgsvurderingsskemaet for at vurdere, i hvilken udstrækning validering af interview ved besøg er mulig. Det vil blive vurderet:

- a. Hvilke spørgsmål fra telefoninterviewene er relevante at validere ved besøg?
- b. Er disse spørgsmål dækket ind i vurderingsskemaerne?

Bilag 4

Oversigt over de gennemførte interview

Interview	Deltagere	Evalueringsspørgsmål ifølge akf's projektbeskrivelse af marts 2003
AT, Kontor for Effekt og Analyse (gruppeinterview)	Fem personer	1-5, 7-9, 13
Arbejdsmiljørådets overvågningsudvalg (gruppeinterview)	Én repræsentant for følgende organisationer: ARF, DA (telefoninterview), FTF, HK, KL, LO, SID, AT (Kontor for Effekt og Analyse)	1-5, 7-9, 13
AT, visionsansvarlige (individuelt interview)	Mindst en repræsentant for hvert af de syv visionsområder i undersøgelsen: børn og unge, ergonomi (EGA/tunge løft), indeklima, kemi, psykisk arbejdsmiljø, støj, ulykker	8-10
Ledelsen hos kontrahenten for VOV1 (individuelt interview)	AMI (pr. mail og telefon) og CASA	Alle 13 spørgsmål
Koordinator for telefoninterviewere	SFI-Survey	4, 5, 10, 11
Telefoninterviewere (individuelt interview)	Mindst en repræsentant for hvert af de syv visionsområder bortset fra børn og unge	4, 5, 10, 11
Projektleder for besøgs-konsulenter	TI (nu Københavns Kommune)	6-12
Besøgs-konsulenter (individuelt interview enten personligt eller telefonisk)	To besøgs-konsulenter pr. visionsområde (én fra Vestdanmark, én fra Østdanmark)	6-12
Interview med kvalitetskontrollører	AMI (individuelt interview); CASA (pr. e-mail)	13

Bem.: I forhold til akf's projektbeskrivelse blev følgende evalueringsspørgsmål ikke belyst ved interviewene: Kontor for Effekt og Analyse: 2, 5 og 7; de visionsansvarlige i AT: 2, 5 og 7; ledelsen hos kontrahenten for VOV1: 5, 9 og 10.

Ca. halvdelen af de interviewede visionsansvarlige i AT har været med til at foreslå spørgsmål til VOV1.

Der er ikke interviewet telefoninterviewere fra børn og unge-området, da nogle af de interviewere, der ifølge vores oplysninger skulle have interviewet børn og unge, ikke havde gjort det.

Oversigt over evalueringsspørgsmål:

- 1 Udtrækket af arbejdssteder
- 2 Filterspørgsmål
- 3 Interview såvel A- som B-side
- 4 Interview sikkerhedsrepræsentant vs. anden medarbejderrepræsentant
- 5 Interview ledelsesrepræsentant vs. medarbejderrepræsentant
- 6 Validering af telefoninterviewene ved hjælp af arbejdsstedsbesøgene
- 7 Temaernes relevans og kvantificering af besøgsdata
- 8 Fordelingen af generelle og visionsspecifikke spørgsmål
- 9 Spørgsmålenes relevans for temaerne
- 10 Spørgsmålenes entydighed
- 11 Praktisk gennemførelse
- 12 Observatørvariationen
- 13 Sammenhængen i rapportering.

Bilag 5

Interviewguider anvendt i evalueringen

1. Spørgsmål til AT, Kontor for Effekt og Analyse
2. Spørgsmål til Overvågningsudvalget
3. Spørgsmål til de visionsansvarlige i AT på de enkelte visionsområder
4. Spørgsmål til kontrahenten for VOV1 ved AMI og CASA
5. Spørgsmål til telefoninterviewere (og koordinatoren for telefoninterviewdelen)
6. Spørgsmål til besøgs konsulenter (og projektlederen for besøgsdelen)
7. Spørgsmål til kvalitetskontrollører

1. Spørgsmål til AT, Kontor for Effekt og Analyse i forbindelse med evalueringen af første overvågningsrunde af fremdriften i forebyggende aktiviteter på virksomhederne.

Formål med overvågningsrunden 2001

Der er to overordnede formål med VOV1: dels at udvikle og afprøve metode, dels at analysere og rapportere resultater som grundlag for, at aktørerne evt. kan justere deres indsats.

- a. Hvordan var vægtningen mellem disse to formål?
- b. Blev vægtningen ændret under gennemførelsen af undersøgelsen?

Udtræk af arbejdssteder til telefoninterview

Formålet med udtræk af en stikprøve af arbejdssteder var at etablere en første måling af forebyggende arbejdsmiljøaktiviteter på ca. 400 arbejdssteder pr. vision inden for branchegrupper, der er i fokus for arbejdsmiljøapparatets indsatser. Det skulle endvidere tilstræbes at dække hver branchegruppe med 50 arbejdssteder. Kravet var endvidere, at stikprøven var stratificeret efter arbejdsstedsstørrelse, således at man ved efterfølgende målinger kunne udtale sig sikkert om ændringer på små, mellemstore og store arbejdssteder.

- a. I hvilket omfang levede de faktiske udtræk af arbejdspladser til telefoninterview og arbejdsstedsbesøg efter jeres vurdering op til kravene?
- b. På hvilke punkter var der væsentlige mangler med konsekvenser for resultaternes generaliserbarhed til de udvalgte brancher?

Interview med såvel A- som B-side

I første undersøgelsesrunde skulle gennemføres interview med såvel A- som B-siden med det formål at give en mere nuanceret beskrivelse af handlingsplanens gennemslagskraft samt for at vurdere graden af overensstemmelse.

- a. På grundlag af jeres læsning af rapporterne, hvad vurderer I så har været den væsentlige værdi ved at telefoninterviewe såvel A- som B-side?

Interview med sikkerhedsrepræsentant versus anden medarbejder

I første undersøgelsesrunde var hensigten på B-siden at interviewe sikkerhedsrepræsentanter, fordi man antog, at de bedre end menige medarbejdere var i stand til at besvare spørgsmål om arbejdsstedets forebyggende aktiviteter.

- a. Er det efter jeres mening muligt at udtale sig generelt om, hvorvidt man bør interviewe henholdsvis en sikkerhedsrepræsentant og en anden medarbejder?

Fordelingen af generelle og visionsspecifikke spørgsmål

Der er ikke i AT's udbudsmateriale nævnt ønsker om generelle spørgsmål på tværs af visionerne. I spørgeskemaerne er omkring halvdelen eller fle-

re spørgsmål generelle, dvs. ikke visionsspecifikke.

- a. Hvad er jeres vurdering af balancen mellem de to sæt af spørgsmål?
- b. Burde der have været relativt flere generelle spørgsmål eller flere visionsspecifikke spørgsmål?
- c. Hvad bidrager de generelle spørgsmål til ifølge jeres vurdering?

Spørgmålenes (i telefoninterviewene) relevans for de udvalgte temaer eller problemstillinger (ad spørgsmål 9)

I har været involveret i formuleringen af spørgsmålene, bl.a. ved at kommentere udkast til spørgeskemaer.

- a. Er det jeres vurdering, at spørgsmålene gennemgående afspejler de temaer, I har ønsket dækket?
- b. Er der væsentlige temaer, der burde have været belyst, men som ikke er blevet det?
- c. I hvor høj grad afspejler de visionsspecifikke spørgsmål gennemgående forebyggende aktiviteter?

Arbejdstilsynet angiver, at det ikke altid har været muligt at belyse områderne: viden, holdninger og handlinger på arbejdsstederne.

- a. Har det været et ønske fra starten af undersøgelsen?

Et sigte med undersøgelsen var, at den skulle muliggøre sammenligning over tid. Spørgsmålene i telefoninterviewene skulle derfor have en sådan præcision og entydighed, at det er muligt at gentage de samme spørgsmål i 2003.

- a. Vurderer I, at spørgsmålene gennemgående havde den ønskede præcision og entydighed?

Validering og supplerings af telefoninterviewene ved hjælp af arbejdsstedsbesøgene

Sigtet med at kombinere telefoninterview og arbejdspladsbesøg var, at arbejdsstedsbesøgene skulle validere samt supplere telefoninterviewene.

- a. Vurderer I, at det lykkedes at validere telefoninterviewene med besøgsdata?
- b. Vurderer I, at det lykkedes at supplere telefoninterviewene med besøgsdata?

Rapporternes praktiske relevans samt formidling

- a. Hvordan vurderer I muligheden for at udnytte de foreliggende resultater fra undersøgelsen (i de otte delrapporter) i indsatsen for at forbedre det forebyggende arbejdsmiljøarbejde på virksomhederne?
- b. Hvordan vurderer I anvendelsesværdien/brugbarheden af VOV1 som en statusmåling, som de efterfølgende undersøgelser skal sammenlignes med?
- c. Hvem er målgruppen for rapporterne?
- d. Hvordan vurderer I rapporternes formidlingsværdi?
- e. Har I ønske om at gennemføre anden formidling af resultaterne fra undersøgelsen end de otte rapporter og det, der foreligger skrevet i Overvågningsrapport 2002?
- f. Vi havde forventet en sammenfattende rapport, som det fremgår af kontrakten mellem AT og AMI. Hvad er grunden til, at den ikke foreligger?

Rapporternes og undersøgelsens forskningsmæssige kvalitet

- a. Lever rapporterne og undersøgelsen som helhed op til jeres forventninger til kvaliteten af sådanne undersøgelser?
- b. Evt. på hvilke punkter er der, efter jeres vurdering, især manglende kvalitet?

2. Spørgsmål til Overvågningsudvalget

i forbindelse med evalueringen af første overvågningsrunde af fremdriften i forebyggende aktiviteter på virksomhederne.

Formål med overvågningsrunden 2001

Der er to overordnede formål med VOV1: dels at udvikle og afprøve metode, dels at analysere og rapportere resultater som grundlag for, at aktørerne evt. kan justere deres indsats.

- a. Hvordan var vægtningen mellem disse to formål?

Udtræk af arbejdssteder til telefoninterview

Formålet med udtræk af en stikprøve af arbejdssteder var at etablere en første måling af forebyggende arbejdsmiljøaktiviteter på ca. 400 arbejdssteder pr. vision inden for branchegrupper, der er i fokus for arbejdsmiljøapparatets indsats. Det skulle endvidere tilstræbes at dække hver

branchegruppe med 50 arbejdssteder. Kravet var endvidere, at stikprøven var stratificeret efter arbejdsstedsstørrelse, således at man ved efterfølgende målinger kunne udtale sig sikkert om ændringer på små, mellemstore og store arbejdssteder.

- a. I hvilket omfang levede de faktiske udtræk af arbejdspladser efter jeres vurdering op til kravene?
- b. På hvilke punkter var der væsentlige mangler med konsekvenser for resultaternes generaliserbarhed til de udvalgte branchegrupper?

Interview med såvel A- som B-side

I første undersøgelsesrunde skulle gennemføres interview med såvel A- som B-siden med det formål at give en mere nuanceret beskrivelse af handlingsplanens gennemslagskraft samt for at vurdere graden af overensstemmelse.

- a. På grundlag af jeres læsning af rapporterne, hvad vurderer I så har været den væsentlige værdi ved at telefoninterviewe såvel A- som B-side?

Interview med sikkerhedsrepræsentant versus anden medarbejder

I første undersøgelsesrunde var hensigten på B-siden at interviewe sikkerhedsrepræsentanter, fordi man antog, at de bedre end en menig medarbejder var i stand til at besvare spørgsmål om arbejdsstedets forebyggende aktiviteter.

- a. Er det efter jeres mening muligt at udtale sig generelt om, hvorvidt man bør interviewe henholdsvis en sikkerhedsrepræsentant og en anden medarbejder?

Fordelingen af generelle og visionsspecifikke spørgsmål (ad spørgsmål 8)

Der er ikke i AT's udbudsmateriale nævnt ønsker om generelle spørgsmål på tværs af visionerne. I spørgeskemaerne er omkring halvdelen eller flere spørgsmål generelle, dvs. ikke visionsspecifikke.

- a. Hvad er jeres vurdering af balancen mellem de to sæt af spørgsmål?
- b. Burde der have været relativt flere generelle spørgsmål eller flere visionsspecifikke spørgsmål?
- c. Hvad bidrager de generelle spørgsmål til ifølge jeres vurdering?

Spørgsmålenes (i telefoninterviewene) relevans for de udvalgte temaer eller problemstillinger (ad spørgsmål 9)

I har været involveret i formuleringen af spørgsmålene, bl.a. ved at kommentere udkast til spørgeskemaer.

- a. Er det jeres vurdering, at spørgsmålene gennemgående afspejler de temaer, I har ønsket dækket?
- b. Er der væsentlige temaer, der burde have været belyst, men som ikke er blevet det?
- c. I hvor høj grad afspejler de visionsspecifikke spørgsmål gennemgående forebyggende aktiviteter?

Et sigte med undersøgelsen var, at den skulle muliggøre sammenligning over tid. Spørgsmålene i telefoninterviewene skulle derfor have en sådan præcision og entydighed, at det er muligt at gentage de samme spørgsmål i 2003.

- a. Vurderer I, at spørgsmålene gennemgående havde den ønskede præcision og entydighed?

Rapporternes praktiske relevans, samt formidling

- a. Hvordan vurderer I muligheden for at udnytte de foreliggende resultater fra undersøgelsen (i de otte delrapporter) i indsatsen for at forbedre det forebyggende arbejdsmiljøarbejde på virksomhederne?
- b. Hvordan vurderer I brugbarheden af undersøgelsen (VOV1) som en statusmåling, som de efterfølgende undersøgelser skal sammenlignes med?
- c. Hvem er målgruppen for rapporterne?
- d. Hvordan vurderer I rapporternes formidlingsværdi?
- e. Har I ønske om at gennemføre anden formidling af resultaterne fra undersøgelsen end de otte rapporter og det, der foreligger skrevet i Overvågningsrapport 2002?

3. Spørgsmål til de visionsansvarlige i AT på de enkelte visionsområder

i forbindelse med evalueringen af første overvågningsrunde af fremdriften i forebyggende aktiviteter på virksomhederne

Baggrund

Hvordan var du involveret i udarbejdelsen af spørgeskemaet til telefoninterview?

Spørgsmålenes (i telefoninterviewene) relevans

I hvilken grad afspejler spørgsmålene de temaer, der er relevante for visionen?

- a. Er der væsentlige temaer, der burde have været belyst, men som ikke er blevet det?
- b. I hvor høj grad afspejler de visionsspecifikke spørgsmål gennemgående forebyggende aktiviteter?

Et sigte med undersøgelsen var, at den skulle sammenlignes med en tilsvarende undersøgelse i 2003.

- a. Vurderer du, at spørgsmålene gennemgående muliggør sammenligning over tid?

Generel vurdering af de visionsspecifikke spørgsmål

Hvad er din gennemgående vurdering af spørgsmålenes præcision og entydighed?

Er der spørgsmål, der helt burde undlades, og nye spørgsmål, der burde medtages?

Fordelingen af generelle og visionsspecifikke spørgsmål

I spørgeskemaerne er omkring halvdelen eller flere spørgsmål generelle, dvs. ikke visionsspecifikke.

- a. Hvad er din vurdering af balancen mellem de to sæt af spørgsmål?
- b. Burde der have været relativt flere generelle spørgsmål eller flere visionsspecifikke spørgsmål?

4. Spørgsmål til kontrahenten for VOV1 ved AMI og CASA i forbindelse med evalueringen af første overvågningsrunde af fremdriften i forebyggende aktiviteter på virksomhederne

Udtræk af arbejdssteder til telefoninterview

- a. Hvordan er stratificeringen af arbejdssteder på arbejdsstedsstørrelse og branchegruppe foretaget?
- b. Hvordan er udtrukket af arbejdssteder til besøg foretaget?
- c. Hvordan vurderer I muligheden for at generalisere til brancher i de tilfælde, hvor svarprocenten er lav (fx under 60%)?
- d. Unge over 18 år og ulykker blev behandlet som én vision. Vil I anbefale, at det samme gøres i VOV2?

Filterspørgsmål

Der blev anvendt filterspørgsmål for nogle visioners vedkommende for at udelade virksomheder, som mod forventning ikke er relevante for den pågældende vision (kemi EGA og tunge løft, evt. børn og unge).

- a. Hvilke konsekvenser mener I, at forskellig anvendelse af filterspørgsmål har for at kunne generalisere til brancher?

Interview med såvel A- som B-side

- a. Hvilken informationsværdi mener I, det har givet at telefoninterviewe såvel A- som B-side?
- b. Hvordan vurderer I den aktuelle håndtering af uoverensstemmelser i besvarelsenerne?
- c. Hvilke problemer ser I ved, at man i nogle tilfælde kun interviewede enten A- eller B-siden? Og at B-siden systematisk har et større bortfald.

Interview med sikkerhedsrepræsentant versus anden medarbejder

- a. Hvordan vurderer I henholdsvis fordele og ulemper ved at bruge sikkerhedsrepræsentant som B-sidens informant (i stedet for en menig medarbejder)?

Validering og supplering af telefoninterviewene ved hjælp af arbejdsstedsbesøgene

Sigtet med at kombinere telefoninterview og arbejdsstedsbesøg var, at arbejdsstedsbesøgene skulle validere såvel som supplere telefoninterviewene.

- a. Vurderer I, at det lykkedes at validere telefoninterviewene med besøgsdata? Læring for VOV2?
- b. Hvordan blev data fra arbejdsstedsbesøgene anvendt til at supplere data fra telefonsurvey? Læring for VOV2?
- c. Er data fra arbejdsstedsbesøgene udnyttet i tilstrækkelig grad i visionsrapporterne?

Temaernes relevans

I AT's udbud og i rapporterne er fastlagt temaer (problemstillinger), som antages at være relevante for måling af udviklingen af de forebyggende aktiviteter på arbejdsstederne.

- a. Var nogle temaer lidet relevante efter jeres mening? Og er der behov for nye temaer til VOV2?
- b. Lykkedes det at måle forebyggende aktiviteter på de syv visionsområder?
- c. Er der behov for at supplere nye visionsspecifikke temaer eller temaer, som går på tværs af de syv visionsområder?
- d. Hvordan vurderer I afdækningen af sammenhængen mellem de visionsspecifikke temaer og det generelle arbejdsmiljøarbejde?

Kvantificering/kvalificering af besøgsdata

- a. I vurderingsskemaerne skulle foretages både en kvantitativ og kvalitativ opgørelse. Hvordan er de kvantitative og kvalitative opgørelses status i forhold til hinanden? Er det fx sådan, at den kvantitative opgørelse i praksis »overdømmer« den kvalitative?

Fordelingen af generelle og visionsspecifikke spørgsmål

- a. Er fordelingen af generelle og visionsspecifikke spørgsmål i såvel telefonspørgeskemaerne som i vurderingsskemaerne hensigtsmæssig i forhold til undersøgelsens formål?

Praktisk tilrettelæggelse af dataindsamling

- a. Var den praktiske gennemførelse af telefoninterviewene og arbejdsstedbesøgene samt kontakten med arbejdsstederne hensigtsmæssig i forhold til at opfylde undersøgelsens formål?
- b. Var træningen af besøgs konsulenterne tilstrækkelig?
- c. Hvordan håndteredes det forhold, at besøgs konsulenterne ved nogle arbejdsstedsbesøg stod over for at skulle udvælge arbejdsområder, der skulle være i fokus på et arbejdssted, hvor der kunne være en mangfoldighed af forskellige arbejdsområder?

Interpersonel variation ved opgørelserne fra arbejdsstedsbesøgene

- a. Var observatørvariationen mellem arbejdsstedsbesøgene truende for observationernes pålidelighed?
- b. Hvad blev gjort for at sikre ensartet praksis inden for de enkelte visioner, jf. fx ovennævnte om udvælgelse af arbejdsområder?

Rapporternes relevans og kvalitet

- a. Hvordan vurderer I den praktiske relevans af visionsrapporterne generelt? Er der nogle rapporter, som er særligt relevante?
- b. Hvordan vurderer I den forskningsmæssige kvalitet af visionsrapporterne generelt? Er der nogle rapporter, som er særligt vellykkede?
- c. Hvad skulle efter jeres mening især forbedres ved rapporterne, hvis I havde 15 ekstra arbejdsdage til hver rapport?
- d. Hvordan vurderer I relevansen og den forskningsmæssige kvalitet af den generelle rapport?

Hvilke anbefalinger har I til VOV2 (vi er opmærksomme på dem, I nævner i den generelle rapport)?

Særlige spørgsmål til hver visionsrapport

Det generelle arbejdsmiljøarbejde

- a. Hvorfor blev intentionen om en sammenfattende rapport på tværs af visionerne ikke realiseret?

Psykosociale risikofaktorer

- a. Undersøgelsen handler i betydelig grad om psykosociale eksponeringer på arbejdspladsen. Skete dette på bekostning af måling af forebyggende aktiviteter?
- b. På hvilket grundlag blev valget af de syv psykosociale risikofaktorer foretaget (ved telefoninterview)?

Kemi

- a. Mange spørgsmål i telefoninterview med tilbageskuende tidshorisont. Hvordan vurderer I pålideligheden af svar på sådanne spørgsmål?
- b. Vurderer I, at begrebet »kemikalier« vil give et for snævert betydningsrum for svarpersonerne, til forskel fra fx »kemiske produkter«?

Børn og unge

- a. Spørgeskemaet afgrænser børn og unge i fritidsjob – var denne afgrænsning konsekvent og ønsket i hele undersøgelsen?
- b. Børn og unges arbejde er et vanskeligt tilgængeligt område at forske i. Hvilke erfaringer og anbefalinger har I med henblik på VOV2?

Indeklima

- a. Forfatterne nævner, at spørgsmål om forekomst af problemer og udvikling i omfang er behæftet med usikkerhed, fordi man kun har spurgt to personer på en hel arbejdsplads. Hvorfor har man så stillet disse spørgsmål?

Støj

- a. Spørgsmålenes detaljeringsgrad i telefoninterview virker urealistisk for respondenterne, hvilket må give pålidelighedsproblemer. Hvad er jeres vurdering?

Tunge løft/EGA

- a. Undersøgelsen består i to delundersøgelser, hvor EGA-arbejdspladser (branche grupper) er underrepræsenteret i udtrækket. Hvilke konsekvenser har det haft for resultaterne ifølge jeres vurdering?

Ulykker

Ingen.

5. Spørgsmål til telefoninterviewere (og koordinatoren for telefoninterviewdelen) i forbindelse med evalueringen af første overvågningsrunde af frem- driften i forebyggende aktiviteter på virksomhederne

Baggrund

Konsulenter

Hvor mange interview gennemførte du alt i alt i undersøgelsen?

Hvor megen erfaring har du med at foretage telefoninterview på arbejds-
miljøområdet?

Koordinator

Hvad var din rolle i undersøgelsen?

Hvordan blev telefoninterviewere udvalgt til opgaven?

Interview med sikkerhedsrepræsentant versus anden medarbejder

I første undersøgelsesrunde var hensigten på B-siden at interviewe sikker-
hedsrepræsentanter, fordi man antog, at de var bedre i stand til at besvare
spørgsmål om arbejdsstedets forebyggende aktiviteter. Oprindeligt var det
Arbejdstilsynets tanke at gennemføre interview med flere arbejdstagere
på det enkelte arbejdssted frem for en repræsentant, men den model blev
ikke valgt i forbindelse med udviklingen af VOV1.

a. Hvad er, efter dine erfaringer, fordele og ulemper ved at interviewe
henholdsvis en sikkerhedsrepræsentant og en almindelig medarbejder?

Interview med ledelsesrepræsentant versus medarbejderrepræsentant

I første undersøgelsesrunde blev der i en række tilfælde kun interviewet
den ene part.

a. Hvad er fordele og ulemper ved, at der på nogle arbejdssteder kun er
interviewet en ledelsesrepræsentant eller kun en sikkerhedsrepræsen-
tant eller en anden medarbejderrepræsentant?

- b. Hvordan bedømmer du spørgsmålenes svarevne for henholdsvis ledelsesrepræsentant og sikkerhedsrepræsentant (med hensyn til eksponeringsspørgsmål)?

Telefoninterviewspørgsmålenes sproglige udformning

Den overordnede problematik er, om spørgsmålene er sprogligt formuleret på en sådan måde, at svarpersonerne vil opfatte dem på samme måde.

Hvad er din gennemgående vurdering af spørgsmålenes præcision og entydighed?

Hvordan var dine erfaringer med spørgsmålene i følgende henseender:

- a. Optræder der vanskelige sætningskonstruktioner?
- b. Optræder der svært forståelige ord for svarpersonerne?
- c. Optræder der ledende formuleringer?
- d. Forekommer der spørgsmål, der betød utidig afbrydelse under oplæsning?
- e. Forekommer der uklarhed i filterspørgsmål og filterinstruktioner?
- f. Forekommer uhensigtsmæssige spørgeføløb, fx overlap eller pludselige skift i emne eller fokus, der ikke er tydeligt markeret?
- g. Er der en logisk sammenhæng mellem spørgsmål og svarkategorier?

Praktisk tilrettelæggelse af dataindsamling

Var den praktiske gennemførelse af telefoninterviewene samt kontakten med arbejdsstederne hensigtsmæssig i forhold til at opfylde undersøgelsens formål?

Hvordan blev svarpersonerne orienteret om formålet med telefoninterviewene?

Hvilke reaktioner var der på spørgsmålene fra svarpersonerne på henholdsvis A- og B-siden?

Var det muligt at gennemføre interviewene inden for den afsatte tidsramme?

Koordinator

Hvor specifikke var adresseoplysninger om virksomhederne?

6. Spørgsmål til besøgs konsulenter

(og projektlederen for besøgsdelen)

i forbindelse med evalueringen af første overvågningsrunde af fremdriften i forebyggende aktiviteter på virksomhederne

Validering og supplerings af telefoninterviewene ved hjælp af arbejdsstedsbesøgene

Sigtet med at kombinere telefoninterview og arbejdspladsbesøg var, at arbejdsstedsbesøgene skulle validere såvel som supplere telefoninterviewene.

- a. Vurderer du, at det lykkedes at validere telefoninterviewene med besøgsdata? Blev data fra interviewguider anvendt ved arbejdsstedsbesøgene til vurdering på de kvalitative vurderingsskalaer i vurderingsskemaerne? Læring for VOV2?
- b. Vurderer du, at det lykkedes ved hjælp af arbejdsstedsbesøgene at supplere data fra telefonsurvey? Hvilke forhold var det især vigtigt at uddybe ved besøgene? Læring for VOV2?
- c. Hvad er i det hele taget fordele og ulemper ved at anvende arbejdsstedsbesøgene til at validere og supplere telefoninterviewene?

Temaernes relevans for udviklingen samt kvantificering/kvalificering af besøgsdata

Arbejdstilsynet havde i sit udbudsmateriale fastlagt en række temaer (problemstillinger), der skulle være relevante for at måle forebyggende aktiviteter på arbejdsstederne.

- a. Er de fastlagte visionsspecifikke temaer (problemstillinger) relevante for måling af udviklingen af de forebyggende aktiviteter på arbejdsstederne? Hvilke temaer er ikke relevante? Er der behov for nye temaer?
- b. Hvordan vurderer du relevansen af de generelle spørgsmål i vurderingsskemaerne?
- c. I vurderingsskemaerne skulle foretages både en kvantitativ og kvalitativ opgørelse. Hvordan er de kvantitative og kvalitative opgørelser anvendt i praksis af besøgs konsulenterne i forhold til hinanden? Er det fx sådan, at den kvantitative opgørelse »overdømmer« den kvalitative?

d. Hvordan vurderer du, at vurderingsskemaets kvantitative og kvalitative opgørelser i det hele taget tjener undersøgelsens formål?

Fordelingen af generelle og visionsspecifikke spørgsmål

Er fordelingen mellem generelle og visionsspecifikke spørgsmål hensigtsmæssig i forhold til undersøgelsens formål?

Praktisk tilrettelæggelse af dataindsamling

Var den praktiske gennemførelse af arbejdsstedsbesøgene samt kontakten med arbejdsstederne hensigtsmæssig i forhold til at opfylde undersøgelsens formål?

Var træningen af besøgs konsulenterne tilstrækkelig?

Hvordan håndteredes det forhold, at besøgs konsulenterne ved nogle arbejdsstedsbesøg stod over for at skulle udvælge arbejdsområder, der skulle være i fokus på et arbejdssted, hvor der kunne være en mangfoldighed af forskellige arbejdsområder?

Hvordan vurderer du i det hele taget, at det lykkedes at gennemføre arbejdsstedsbesøgene? Er der forhold, som man burde have håndteret anderledes?

Interpersonel variation ved opgørelserne fra arbejdsstedsbesøgene

Hvordan vurderer du betydningen af observatørvariationen mellem arbejdsstedsbesøgene med hensyn til observationernes pålidelighed?

Hvad blev gjort for at sikre ensartet praksis inden for de enkelte visioner, jf. fx ovennævnte om udvælgelse af arbejdsområder?

Hvilke anbefalinger har du i forhold til VOV2 både generelt og på dit visionsområde?

Evt. særlige spørgsmål for visionsområdet

7. Spørgsmål til kvalitetskontrollører
i forbindelse med evalueringen af første overvågningsrunde af fremdriften i forebyggende aktiviteter på virksomhederne

Hvordan foregik kvalitetskontrollen af VOV1?

Hvilket materiale vurderede du?

Af hvilken karakter var dine bemærkninger:

Hvor kritiske var de?

Stillede du forslag til ændringer?

Gav du dine bemærkninger mundtligt eller skriftligt?

Har dine bemærkninger haft konsekvenser for justering af materialerne?

Bilag 6

Vurdering af dataudtrækket til VOV1 og anbefalinger til det videre arbejde. Udarbejdet af Anders Holm, Sociologisk Institut, KU

Indledning

Dette notat indeholder en evaluering af udtræksproceduren i forbindelse med første fase af et projekt om fremdriften af arbejdsmiljøarbejdet. Endvidere indeholder notatet nogle anbefalinger til, hvorledes udtræksproceduren og analyserne af data kan forbedres i de fremtidige analyser.

Der er i forbindelse med undersøgelsen af fremdriften af arbejdsmiljøarbejdet defineret syv indsatsområder eller, som de også er benævnt, visionsområder. Disse områder er: børn og unge, kemi, ergonomiske arbejdsforhold og tunge løft, indeklima, det psykiske arbejdsmiljø, ulykker og støj. Punkterne a) til g) nedenfor henviser til evalueringsspørgsmålene, se bilag 2.

Evaluator er opmærksom på, at der ikke i udbudsmaterialet er stillet krav om statistisk styrke på brancheniveau.

Af dette notat fremgår det, at:

- a. der kræves mellem 180 og 700 observationer (virksomheder) for at kunne foretage sammenligninger mellem to strata, alt efter om man lægger vægt på statistisk signifikans eller styrke. Dvs. mellem 90 og 350 virksomheder i hvert stratum for at opnå de foreskrevne krav om statistisk signifikans og styrke. Det anbefales derfor at anvende et

- mindre krav til den statistiske styrke, fx 70%, hvilket vil medføre en stikprøve på ca. 150 observationer i hvert stratum.
- b. der ikke nogen steder i det materiale, der har dannet baggrund for dette notat, er vist en fordeling af virksomheder på branchegruppe og størrelse. Hvis strata er jævnt fordelt, er der ikke nogen grund til at stratificere, fordi en simpel tilfældig udvælgelse automatisk giver en jævn fordeling i dette tilfælde. Er fordelingen mellem strata derimod skæv, kan stratificering være nødvendig for at opnå tilstrækkelig statistisk styrke og signifikans. Vurdering af behovet for stratificering kræver således information om fordelingen i populationen, så man kan vurdere behovet for ekstra observationer i små strata.
 - c. der som nævnt ovenfor er defineret nogle visionsområder. Disse består hver især af en række udvalgte branchegrupper. Der er i forbindelse med udbuddet af den første fase af undersøgelsen udtrykt ønske om en stratificering på disse branchegrupper for hvert visionsområde. Derudover har der været diskussion af det betimelige i også at udføre en yderligere stratificering på størrelse af virksomhederne inden for de enkelte branchegrupper.
 - d. det imidlertid ikke af baggrundsmaterialet klart fremgår, hvorvidt der har været udført et udtræk stratificeret både på størrelse og branchegruppe eller kun på branchegruppe.
 - e. der er blevet foretaget et udtræk baseret i hovedsagen på to registre. Et der er virksomhedsbaseret, og et der er medarbejderbaseret. Efterfølgende er data fra det virksomhedsbaserede udtræk vægtet efter antal beskæftigede og er dermed sammenligneligt med det medarbejderbaserede udtræk. De enkelte strata er dermed repræsentative med hensyn til beskæftigelse.
 - f. der i det første udtræk, der her er evalueret, er tale om meget varierende bortfaldsprocenter. Men der er ikke foretaget en egentlig bortfaldsanalyse, og det er derfor vanskeligt at udtale sig om repræsentativiteten i det foreliggende udtræk.
 - g. der er indsamlet oplysninger om unge. Unge under 18 år er indsamlet i et udtræk baseret på uddannelsesinstitutioner og er dermed repræsentativt, mens unge på 18 år og derover er indsamlet via de stratificerede udtræk. Da strata ikke er repræsentativt, kan de to udtræk i princippet ikke sammenlignes, hvilket heller ikke har været meningen. Ydermere er oplysningerne om de unge på 18 år og derover er blevet indsamlet i forbindelse med ulykkesvisionen. I det omfang, at fremdriften for arbejdsmiljøet for de unge på 18 år og derover ikke er repræsenteret ved

de samme branchegrupper som ulykkesvisionen, er denne del af undersøgelsen naturligvis problematisk.

Resten af notatet indeholder to afsnit samt et appendiks. Det første afsnit indeholder dels en beskrivelse af kravene til udtræksprocedurerne for den første fase af undersøgelsen, dels en diskussion af det aktuelle udtræk og dets stratificering. Det andet afsnit indeholder forslag til den fremtidige udtræksprocedure og analyse i forbindelse med de kommende faser af projektet.

2. Udtræksprocedurerne ved 2001-undersøgelsen

Der er udført en telefoninterviewundersøgelse, og udtræksproceduren for denne skal diskuteres i dette afsnit.

2.1. Kravspecifikationerne til 2001-undersøgelsen

Dette afsnit bygger i hovedsagen på den indgåede kontrakt mellem Arbejdstilsynet og Arbejds miljøinstituttet, men også øvrige dokumenter indgår, herunder udbudsmaterialet til 2001-undersøgelsen.

Ifølge denne kontrakt ønskes et udtræk uden systematiske fejl. Der er aftalt et udtræk fra NAK-registret (se næste afsnit), hvorved der opnås repræsentation efter beskæftigelsesandele (dvs. virksomheder efter antal ansatte). I det oprindelige udbudsmateriale var der ønsket et udtræk fra CVR-registret (se næste afsnit), men dette krav er efterfølgende blevet fraveget til fordel for et udtræk fra NAK. Der ønskes 400 virksomheder inden for hvert af de syv visionsområder og en yderligere stratificering efter udvalgte branchegrupper. Det sidste sker i forbindelse med udførelsen af projektet. Data kan illustreres af nedenstående tabel:

Branchegruppe	Visionsområde 1			Visionsområde 7
Branchegruppe1	0-4 ansatte	5-19 ansatte	Mere end 19 ansatte		
.....					
Branchegruppe K					
I alt	400			...	400

Når der i tabellen også vises strata (celler) på virksomhedsstørrelser, er det fordi, der i processen omkring udtrækket er blevet diskuteret en yderligere stratificering på virksomhedsstørrelser.

Stikprøven skal altså stratificeres på branchegrupper inden for hvert visionsområde, hvoraf nogle af branchegrupperne er udvalgt specifikt inden for et visionsområde.

Der ønskes at kunne påvise ændringer på 10-15% (absolutte ændringer) på et 5%’s signifikansniveau med en styrke på 80%. Det opfattes i nærværende notat, som at der ønskes at kunne sammenligne forskelle mellem de enkelte strata i tabellen ovenfor, både mellem branchegrupper og over tid, altså mellem den allerede foretagne undersøgelse og kommende undersøgelser. Kravet om at kunne påvise en forskel på 10-15% på et 5%’s signifikansniveau medfører en samlet stikprøve i de to strata, der sammenlignes på mellem 180 og 400 observationer. Kravet om at kunne påvise en forskel på 10-15% med en styrke på 80% medfører en samlet stikprøve i de to strata, der sammenlignes, på mellem 400 og 820 observationer for VOV1 og VOV2 tilsammen. Hensynet til styrke bestemmer med andre ord stikprøvestørrelsen. I appendiks 1 er ovenstående beregninger nærmere gennemgået.

Hvis der i hvert visionsområde udtages 400, kan der altså slet ikke stratificeres, hverken på branchegrupper eller på virksomhedsstørrelser, hvis hvert stratum (celle i tabellen ovenfor) skal indeholde 300-ca. 800 observationer. Hvis kun signifikanskravet om at kunne påvise ændringer på min. 15% lægges til grund for bestemmelsen af stikprøven, kan der opereres med op til 4-5 branchegrupper inden for hvert visionsområde. Så opnås et antal observationer, der er nødvendig for at opnå et samlet antal observationer på ca. 180, når der sammenlignes to strata, hvis den ønskede signifikans skal opnås. Det anbefales derfor at ændre på kravet til styrken, til fx 70%. I så fald kræves 300 observationer i alt for at opnå den ønskede styrke eller 150 observationer i hvert stratum. Der kan i så fald sammenlignes 2-3 strata inden for hvert visionsområde med tilsvarende strata i andre visionsområder eller andre branchegrupper eller i senere undersøgelser.

Man skal dog også være opmærksom på, at man jo udmærket kan stratificere yderligere ved indsamling af stikprøven, men at der skal være

300 vægtede observationer i alt i de sammenligninger, man foretager, hvis den ønskede signifikans og styrke skal opnås.

2.2 Udtrækket til 2001-undersøgelsen

Dette afsnit bygger i hovedsagen på Bach et al. (2002). Der er tale om tre kilder til datamaterialet til den undersøgelse, der allerede er foretaget. Dels et udtræk fra NAK (den nationale arbejdsmiljøkohorte), dels CVR (det centrale virksomhedsregister på Danmarks Statistik) og dels en spørgeskemaundersøgelse blandt skolesøgende under 18, som var i beskæftigelse.

NAK er et personregister beliggende hos Arbejdsmiljøinstituttet. CVR er et virksomhedsregister for alle momsregistrerede virksomheder samt alle offentlige virksomheder, uanset momsregistrering. Dette register er baseret på arbejdsgivernes SE-numre. Endelig er spørgeskemaundersøgelsen blandt de skolesøgende under 18 år udført via et udtræk fra uddannelsesinstitutioner.

NAK er et personregister, og når der udtrækkes virksomheder fra dette register, bliver store virksomheder overrepræsenteret i forhold til populationen af virksomheder, idet der er større sandsynlighed for, at en tilfældigt udtrukket person er beskæftiget i en stor virksomhed end i en lille. Omvendt bliver udtrækket repræsentativt i forhold til populationen af beskæftigede.

CVR er et virksomhedsregister, og når der udtrækkes virksomheder fra dette register, opnås den modsatte situation, nemlig at udtrækket bliver repræsentativt i forhold til populationen af virksomheder, men ikke i forhold til populationen af beskæftigede.

Til at belyse det særlige visionsområde, der hedder »børn og unge«, skal udtages virksomheder med respondenter i aldersgruppen 13 til 24 år. Og sammenligninger inden for denne vision skal underlægges de samme krav til signifikans og styrke som de øvrige strata i undersøgelsen. Da NAK kun indeholder oplysninger om lønmodtagere over 18 år, opstår et særligt problem i forbindelse med udtagelse til dette stratum. Der udtages dels virksomheder fra NAK, dels oplysninger om 13-17-årige, der opnås via uddannelsesinstitutioner. Uvægtet kommer oplysningerne fra NAK til at afspejle beskæftigelsesfordelingen blandt de 18-24-årige inden for hvert

stratum. Oplysninger fra uddannelsesinstitutionerne er til at afspejle beskæftigelsessituationen blandt de 13-17-årige helt generelt, idet der ikke er mulighed for at udvælge på branchegruppe blandt disse respondenter. Dermed bliver materialet for de to aldersgrupper kun sammenligneligt, hvis der vægtes efter et eller andet kriterium.

For at opnå sammenlignelighed mellem de tre stikprøver er virksomhederne vægtet efter beskæftigelse.

Det fremgår ikke klart af noget af det udleverede materiale, præcist hvordan stikprøven er fordelt på strata. Det forekommer mangelfuldt, at der ikke er vist en tabel, der viser stikprøven, vægtet, såvel som uvægtet, fordelt på branchegrupper, virksomhedsstørrelser og visionsområder, jf. den illustrerede tabel ovenfor. Det er derfor svært helt præcist at vurdere, hvorvidt udtrækket opfylder krav til signifikans og styrke, således som de er beregnet i forrige afsnit. Men af Arbejdstilsynet (2000) fremgår det, at der inden for hvert visionsområde tilsyneladende opereres med mellem to og ni branchegrupper til brug for stratificeringen. Dvs., at hvis der ikke er stratificeret med hensyn til virksomhedsstørrelse, er der inden for alle visionsområder det nødvendige antal observationer for at opnå den ønskede statistiske styrke ved sammenligningerne.

2.2.1 Stikprøve og opnåelsesprocent

Til denne undersøgelse er i alt udtrukket 4.641 virksomheder, og heraf indgår 3.287 (70,8%) i undersøgelsen. Af de virksomheder, der indgik i undersøgelsen, var de 43% udtrukket via NAK, de 42% udtrukket via CVR og de sidste 15% via spørgeskemaundersøgelsen blandt skolesøgende børn under 18 år.

Svarprocenten svinger fra 44%³ ud af 81 udvalgte virksomheder inden for fjerkræslagterier til 95% for 145 udvalgte virksomheder inden for undervisning og forskning og 97% inden for 86 udvalgte virksomheder inden for pasning af børn. Generelt er bortfaldet størst inden for privat erhverv, især fremstilling, og mindst blandt offentlige virksomheder. Der burde have været foretaget en bortfaldsanalyse med henblik på at undersøge, hvorvidt bortfaldet er meget systematisk i de strata, hvor bortfaldet er størst. Evt. kunne der have været gennemført en bortfaldsanalyse for hele materialet.

2.2.2 Stratificering

De 4.641 udvalgte virksomheder blev udtrukket fordelt på delbranche-grupper, ca. 40 i alt, samt inden for de syv visionsområder. Det fremgår som tidligere nævnt ikke, hvorvidt der yderligere også er stratificeret efter virksomhedsstørrelse, men der nævnes strata med virksomheder med under 4 ansatte, med mellem 5 og 19 ansatte og med mere end 19 ansatte⁴. Da virksomhederne i nogle tilfælde er udtrukket efter antal beskæftigede (NAK og spørgeskemaundersøgelsen) og nogle gange efter antal virksomheder (CVR), skal data vægtes for at være sammenlignelige. Der er i undersøgelsen foretaget en vægtning af data fra CVR, således at virksomhederne fra dette register vægtes efter antal beskæftigede.

Den efterfølgende vægtning betyder, at det er muligt at sammenligne data inden for de forskellige strata og mellem de forskellige strata.

Det er derimod sværere at sammenligne over aggregerede strata. Hvis det fx ønskes at beregne en samlet andel, der svarer ja til et bestemt spørgsmål inden for et visionsområde, skal de forskellige strata vejes sammen efter et eller andet kriterium. Det er imidlertid ikke klart, hvad dette kriterium skal være. Et kriterium kunne være de enkelte stratas beskæftigelsesandele. Altså at virksomhederne vægtes efter den andel af beskæftigelsen, de har inden for de branchegrupper, der er udvalgt inden for det enkelte visionsområde. Men da de enkelte strata (branchegrupper) inden for hvert visionsområde ikke repræsenterer noget som helst, er det uklart, hvad de sammenvejede data udtrykker.

Endvidere skal der ved sammenvejning af indeks mellem strata tages hensyn til den såkaldte designeffekt. Denne effekt betyder, at usikkerheden på de estimerede parametre fra stikprøven, fx andele inden for forskellige visionsområder, er større end den, der fremgår af de almindelige beregninger for standardafvigelser, jf. Wolter (1985). Designeffekten skyldes, at variationen inden for strata typisk er mindre end den, der ville være opnået ved simpel tilfældig udvælgelse, idet observationer inden for strata typisk er mere homogene end observationer mellem strata. Derved bliver den samlede variation i en stratificeret stikprøve mindre end i en stikprøve udtaget efter simpel tilfældig udvælgelse. Den typiske designeffekt betyder, at de beregnede standardafvigelser er halvt så store, som de ville være ved

simpel tilfældig udvælgelse. Wolter (1985) anviser, hvorledes standardafvigelse i stratificerede stikprøver korrigeres.

3. Anbefalinger til det fremtidige arbejde

3.1 Udtræk

Det må anbefales, af hensyn til mulighederne for at vægte data, at der udtrækkes fra én datakilde, fx CVR. På den måde bliver den efterfølgende bearbejdning og dokumentation mere overskuelig. Herefter kan overvejelser om repræsentativitet med hensyn til beskæftigede eller virksomheder efterfølgende resultere i en stratificering af materialet, men denne stratificering og tilhørende vægtning kan foregå efter et ensartet grundlag.

3.2 Stratificering og stikprøvestørrelse

Hvis udtrækket stratificerer på fx virksomhedsstørrelse og branchegruppe, og man ønsker at vurdere en evt. ændring i arbejdsmiljøet på forskellige parametre, spiller stikprøvestørrelsen i de enkelte strata og på de forskellige målingstidspunkter selvfølgelig en væsentlig rolle. For at opnå den ønskede styrke i alle strata, bør det overvejes at sikre en større samlet stikprøve i de strata, hvor der indgår mange branchegrupper, mens visningsområder med et mindre antal branchegrupper kunne belyses med en mindre stikprøve. På den måde kan samme statistiske styrke opnås i alle stikprøvens strata.

Det kunne også overvejes at forøge den samlede stikprøvestørrelse. I strata, der i den første undersøgelse indeholder meget få virksomheder, kan en stratificering, hvor der opnås tilstrækkeligt antal virksomheder, i følgende undersøgelse opnå et samlet antal virksomheder, der er tilstrækkeligt til at opnå den ønskede signifikans.

Endelig kunne man også overveje at opstille en egentlig statistisk model for ændringerne på de enkelte parametre, fx en logistisk regressionsmodel. Modellen kunne indeholde forklarende faktorer for sammenhængen mellem de studerede variabler på den ene side, og branchegruppe og virksomhedsstørrelse på den anden. Hvis der også indgår andre virksomhedskarakteristika i undersøgelsen, kan disse også indgå i analysen. På denne måde kunne man opnå et klarere overblik over evt. systematiske effekter fra branchegruppe og virksomhedsstørrelse. Selv om nogle af disse effekter

muligvis går på tværs af hinanden vil den samlede statistiske model stadigvæk give information om, hvorvidt der er selvstændige effekter fra virksomhedsstørrelse og branchegruppe, men samtidig også information om, hvorledes den specielle effekt af virksomhedsstørrelse påvirker arbejdsmiljøarbejdet inden for hver branchegruppe. Ved at udføre en analyse som den ovenstående opnår man desuden, at de estimerede sammenhænge bliver ufølsomme over for stratificerings- og vægtningsproblemer. Man opnår konsistente estimater af effekter af branchegruppe- og virksomhedsstørrelse på arbejdsmiljøet, uanset om der vægtes eller ej, og uanset om stikprøven er stratificeret eller ej.

Appendiks 1

Det er nævnt ovenfor, at man ønsker at kunne påvise en ændring i de forskellige parametre for måling af arbejdsmiljøet på 10-15% med 5%’s signifikans og med en styrke på 80%. I det følgende diskuteres, hvilke krav det stiller til stikprøverne i to målinger, der ønskes påvist ændringer imellem.

Vi antager, at der er tale om andele, der støtter et bestemt synspunkt i de to målinger, fx at der er blevet mindre støj på arbejdspladsen. Nulhypotesen, som skal afvises med stor sandsynlighed, når der er tale om en ændring på mere en 10-15%, er, at der ikke er forskel på andelen i de to stikprøver. Vi har altså:

$$p_1 = p_2 = p_{H_0},$$

hvor p_1 og p_2 er andelene, der støtter det synspunkt, der undersøges i den første henholdsvis den anden undersøgelse.

Variansen på estimerterne af de to andele under nulhypotesen om, at de ikke er forskellige, er:

$$\text{var}(\hat{p} | H_0) = \frac{p_{H_0}(1 - p_{H_0})}{n_j}; j = 1, 2,$$

hvor n_j er antallet af observationer i den j 'te undersøgelse i det betragtede segment. Hermed bliver variansen på forskellen mellem de to estimerter:

$$\begin{aligned} \text{var}(\hat{p}_2 - \hat{p}_1 | H_0) &= \text{var}(\hat{p}_2 | H_0) + \text{var}(\hat{p}_1 | H_0) = \frac{p_{H_0}(1-p_{H_0})}{n_2} + \frac{p_{H_0}(1-p_{H_0})}{n_1} \\ &= p_{H_0}(1-p_{H_0}) \left(\frac{1}{n_2} + \frac{1}{n_1} \right) < \frac{p_{H_0}(1-p_{H_0})}{n/4}; n = n_2 + n_1 \end{aligned}$$

Beviset for den sidste ulighed fremgår af appendiks 2. Ved at bruge variansudtrykket til højre for ulighedstegnet opnås et variansskøn, der er større end det sande. Det vil derfor lede til stikprøver, der er større end nødvendigt for at opnå den ønskede styrke. Men det er approksimativt lig den sande varians, når stikprøverne i de to strata er nogenlunde ens og er en rimelige approksimation af den sande varians, så længe hver stikprøve er over 10, og den ene stikprøve ikke er dobbelt så stor som den anden. I sådanne situationer opnås der således ikke væsentligt forskellige stikprøvestørrelser, uanset hvilket udtryk der anvendes. Ved at bruge udtrykket til højre opnås endvidere et mere simpelt udtryk for de ønskede stikprøvestørrelser, end hvis de korrekte udtryk anvendes.

Det fremgår også af formlerne for variansen på forskellen i andele, at denne opnår sit maksimum ved $p_{H_0} = 0,5$. For at opnå så konservativ en vurdering af stikprøvestørrelse som mulig anvendes denne værdi i det følgende, også når der beregnes styrke ved test af hypotesen om, at de to andele ikke er ens⁵.

For at en evt. forskel mellem estimerne skal være signifikant (med 95% sandsynlighed for, at forskellen faktisk skyldes, at andelen ikke er ens), skal forholdet mellem den observerede forskel og standardafvigelsen (kvadratroden af variansen) på forskellen være approksimativt 2, dvs.:

$$\frac{\hat{p}_2 - \hat{p}_1}{\sqrt{\frac{4 \cdot 0,25}{n_2 + n_1}}} = \frac{\Delta \hat{p}}{\sqrt{\frac{4 \cdot 0,25}{n_2 + n_1}}} > 2,$$

hvor $\Delta \hat{p} = \hat{p}_2 - \hat{p}_1$. Dvs. hvis $\Delta \hat{p} = 0,15$ får vi:

$$\frac{0,15}{\sqrt{\frac{4 \cdot 0,25}{n_2 + n_1}}} > 2 \Rightarrow (n_2 + n_1) > 178,$$

dvs. n_1 og n_2 skal hver være på ca. 90. Den samlede stikprøve i de to strata skal tilsammen være på minimum 180 respondenter, før der kan påvises forskelle i andele på minimum 0,15. Hvis der skal kunne påvises en forskel på 0,1, skal stikprøverne tilsammen være på minimum 400 observationer.

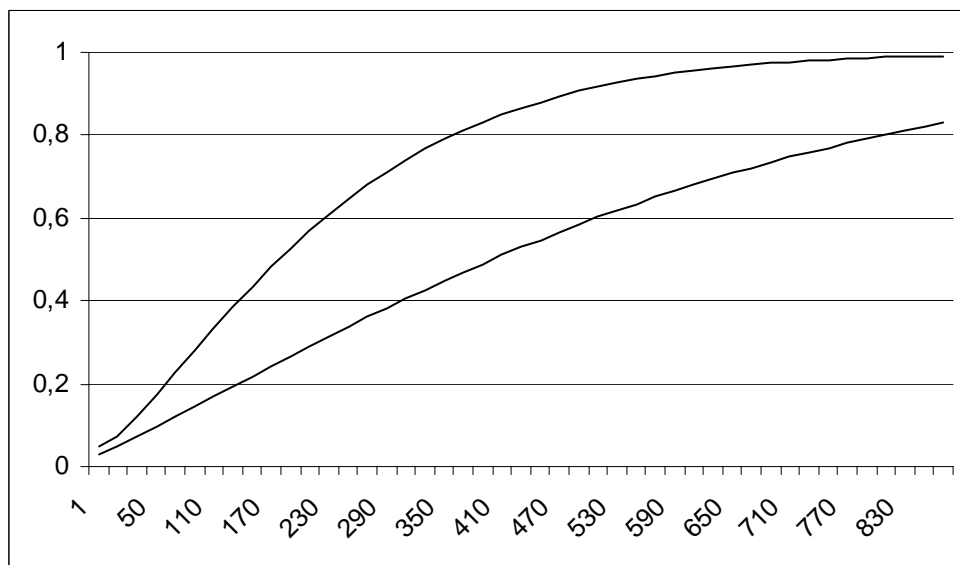
Hvis stikprøven i de to strata tilsammen er $n = n_2 + n_1$, den sande forskel på de to andele, er Δp , og der er observeret en forskel på $\Delta \hat{p}_{H_0:afvises}$ svarende til, at nulhypotesen lige akkurat afvises⁶, opnås en styrke på:

$$P\left(|\Delta \hat{p}| \geq \Delta p_{H_0:afvises} \mid \Delta \hat{p} \sim N\left(\Delta p; \sqrt{\frac{4 \cdot 0,25}{n}}\right)\right) = 1 - \Phi\left(\frac{\Delta p_{H_0:afvises} - \Delta p}{\sqrt{\frac{4 \cdot 0,25}{n}}}\right) + \Phi\left(\frac{-(\Delta p_{H_0:afvises} + \Delta p)}{\sqrt{\frac{4 \cdot 0,25}{n}}}\right)$$

I figur A1 vises styrken ved forskellige stikprøvestørrelser, hvis den sande forskel mellem de to andele er på 0,15, og hvis forskellen er på 0,10.

Figur A1 Styrken ved test af forskelle i andele mellem to stikprøver

A1



Den øverste kurve er styrken, hvis den sande forskel er 0,15, og den nederste, hvis forskellen er 0,1. X-aksen er antallet af observationer. Det ses, at en styrke på 80% i det første tilfælde opnås ved en stikprøve på ca. 400 observationer, mens den samme styrke opnås ved ca. 820 observationer i det andet tilfælde.

Appendiks 2

Sæt: $n = n_2 + n_1; n_2, n_1 > 0$. Så fås:

$$p_{H_0}(1-p_{H_0})\left(\frac{1}{n_2} + \frac{1}{n_1}\right) \leq \frac{p_{H_0}(1-p_{H_0})}{n/4} \Leftrightarrow \left(\frac{1}{n_2} + \frac{1}{n_1}\right) \leq \frac{4}{n_2 + n_1} \Leftrightarrow$$

$$\frac{n_2}{n_2} + \frac{n_2}{n_1} + \frac{n_1}{n_2} + \frac{n_1}{n_1} \leq 4 \Leftrightarrow \frac{n_2}{n_1} + \frac{n_1}{n_2} - 2 \leq 0 \Leftrightarrow$$

$$n_2^2 + n_1^2 - 2n_2n_1 \leq 0 \Leftrightarrow (n_2 - n_1)^2 \leq 0$$

Referencer

Arbejdstilsynet (2003): Foreløbig fastlæggelse og præcisering af evalueringsspørgsmål.

Arbejdstilsynet (2002): Overvågning 2001, Bilagsrapport – Del 1 (inkl. Bilag A-F).

Arbejdstilsynet (2001): Overvågning, bilagsrapport – Del 1 (inkl. Bilag A-F).

Arbejdstilsynet og Arbejds miljøinstituttet (2000): Kontrakt mellem Arbejdstilsynet og Arbejds miljøinstituttet.

Arbejdstilsynet (2000): Udbud af overvågning af fremdriften i 2005-handlingsplanen.

Arbejdstilsynet (2000): Udbud af dataindsamlingsdelen i projektet »Overvågning af fremdriften i 2005-handlingsplanen«.

Bach, Elsa; Peter Hasle og Hans Jørgen Limborg (2002): Måling af fremdriften i arbejdsmiljøarbejdet: Det generelle arbejdsmiljøarbejde. Arbejds miljøinstituttet. 1. udkast.

COWI A/S (2003): Overvågning af fremdriften i arbejdsmiljøarbejdet i virksomhederne. Projektbeskrivelse. COWI A/S.

Wolter, K.M. (1985): Introduction to variance estimation. Springer, New York.

Bilag 7

Sammenfatning af sproglig evaluering af spørgeskemaer til VOV1

Udarbejdet af Hanne-Pernille Stax, Institut for Almen og Anvendt Sprogvidenskab, KU

Sprogkonsulenten har gennemgået de syv spørgeskemaer, som telefoninterviewerne har anvendt. Hun har i de oprindelige skemaer indsat markeringer og koder for problemer, fremsat forslag til konkrete løsninger og for hvert spørgsmål vurderet, om det vil være muligt at anvende samme spørgsmål i VOV2, samt om sammenligning af data vil være mulig.

På grund af omfanget (de mange sider) bringes i dette bilag alene sprogkonsulentens generelle bemærkninger samt opsummering af hendes vurdering af de enkelte visionsskemaer.

En sproglig evaluering er i sagens natur en kvalitativ indsats

Ved vurderingen af de enkelte spørgsmål anvendes en række teoretiske kriterier for »god spørgeteknik« samt erfaringer fra empirisk forskning i effekter af forskellige udformninger af spørgeskemaspørgsmål. Hvert enkelt spørgsmål tjekkes bl.a. for:

- uklare formuleringer og flertydighed
- vanskelig sætningskonstruktion
- svært forståelige ord – i forhold til målgruppen
- ledende formuleringer
- spørgsmålsdesign, der giver øget risiko for utidig afbrydelse under oplæsning

- inkonsistens mellem spørgsmål og svarformuleringer
- uklarhed i filterspørgsmål og filterinstruktioner
- risiko for respons-bias grundet konteksteffekter
- uhensigtsmæssige spørgeforløb, fx ennemæssige overlap eller pludselige skift i emne eller fokus, der ikke er tydeligt markeret

Hvert enkelt spørgsmål vurderes i den sammenhæng, det indgår i. Og det betyder, at det ikke altid er vurderet hensigtsmæssigt at foreslå den samme løsningsmodel til at løse tilsyneladende parallelle problemtyper i forskellige spørgsmål.

Hvert enkelt spørgsmål er ved evalueringen forsynet med to koder. En kode, der angiver den sproglige kvalitet af spørgsmålet i VOV1. Og en kode, der angiver hvorvidt, de foreslåede ændringer vil gøre sammenligning mellem VOV1 og VOV2 problematisk. Se beskrivelse af koderne nedenfor.

Disse koder er anvendt for at gøre det muligt at danne sig et hurtigt overblik over den sproglige kvalitet af spørgeskemainstrumenterne i VOV1 samt af de mulige negative konsekvenser af at forsøge at forbedre den sproglige kvalitet af spørgeskemainstrumenterne til VOV2.

Koderne er retningsgivende, og der vil være tilfælde, hvor to tilsyneladende parallelle spørgsmål – eller parallelle ændringsforslag – er kodet lidt forskelligt. Fx er spørgsmål, der indeholder parenteser og skråstreger i VOV1 generelt tildelt koden ÆA (ændring anbefales). Koden anvendes, fordi parenteser og skråstreger vanskeliggør interviewerens oplæsning og kan resultere i, at forskellige respondenter præsenteres for forskellige versioner. I nogle tilfælde er sådanne spørgsmål dog tildelt koden ÆM (ændring mulig), idet det er vurderet, at brug af parentes eller skråstreg i enkelte tilfælde ikke vil give anledning til problemer for interviewer eller respondent. Koderne angiver en tentativ vurdering, men den sproglige evalueringens substans er bemærkningerne til de enkelte spørgsmål samt de foreslåede omformuleringer.

De foreslåede omformuleringer angiver ikke endegyldige løsninger på de påpegede problemer. I samspil med karakteristikken af de afdækkede problemer kan de fungere som støtte, når skemadesignereren skal udforme de endelige spørgsmål.

I sagens natur er den intention som spørgeren oprindeligt havde med hvert enkelt spørgsmål, ikke kendt for sprogkonsulenten. De foreslåede alternative formuleringer repræsenterer således konsulentens bedste bud på hensigtsmæssig formulering af det, som antages at have været intentionen. I tilfælde, hvor et oprindeligt spørgsmål er uklart eller tvetydigt, kan en foreslået alternativ formulering naturligvis godt ramme ved siden af det, der var den oprindelige intention med spørgsmålet.

De afdækkede problemtyper

Hovedparten af de afdækkede problemer fordeler sig på nedennævnte seks problemkategorier:

Utilstrækkelige filtergange

I mange tilfælde sikrer filtergangene i VOV1 ikke at respondenter ledes uden om irrelevante spørgsmål. Det giver et sjusket og ufærdigt indtryk, og det tager unødigt lang tid, når der fx stilles spørgsmål, som forudsætter gennemført APV, til respondenter, der har svaret, at der ikke er gennemført APV på arbejdspladsen (se fx spm 40 & 42), eller når der stilles spørgsmål, som forudsætter sikkerhedsorganisation, til respondenter på arbejdspladser uden sådanne (se fx 45 & 46). Det virker ugennemtænkt at disse spørgsmål stilles til alle, også selv om spørgsmålene har en svarmulighed, hvor det kan registreres, at de er irrelevante.

Uhensigtsmæssig udformning af spørgsmål til oplæsning i interview

Der er mange parenteser og skråstreger i spørgsmålsformuleringerne. Det betyder, at de enkelte interviewere må træffe mange valg, som designerne skulle have truffet for dem. Konsekvenser er unødvendig belastning af intervieweren. Og – mere problematisk – en stor risiko for, at forskellige interviewere vil oplæse de »ufærdige« spørgsmål forskelligt. Dermed kan reliabiliteten af nogle svar i VOV1 anfægtes.

Centrale specifikationer er ofte placeret i instruktion til interviewerne og ikke indarbejdet i spørgsmålsformuleringerne. Konsekvenser som ovenfor.

Unødvendigt komplekse sætningskonstruktioner

Komprimerede skriftsproglige formuleringer vanskeliggør oplæsning og lytteforståelse. I en del tilfælde er de foreslåede alternativer lidt længere end i VOV1. Denne forlængelse vil dog forventeligt opvejes af tidsbesparelse i VOV2 grundet færre oplæsningsproblemer og færre anmodninger om gentagelse af komplekse formuleringer.

Unødvendigt »teknisk« ordvalg

Lange ord, som er sammensat til lejligheden, og som vanskeliggør oplæsning og lytteforståelse. Det anbefales hyppigt at bryde sammensatte ord op, hvis de ikke er almindeligt anvendte.

Inkonsistent begrebsanvendelse omkring centrale størrelser

Fx skiftes der hyppigt mellem »virksomhed«, »arbejdsplads«, »den lokale arbejdsplads« og »arbejdssted«. Det er unødigt forvirrende, og respondenter fra arbejdspladser, der er underafdelinger af større virksomheder, kan fx komme i tvivl om, hvilken enhed der spørges til. Det anbefales generelt af anvende begrebet »arbejdsplads«.

Der skiftes også hyppigt mellem »ansatte«, »medarbejdere«, »arbejdstagere« og »de beskæftigede«.

Urealistisk respondentopgave

Spørgsmålene stiller urealistiske krav til respondenterne ang. opgaveforståelse og svarproduktion. Fx fordrer en del spørgsmål, at respondenterne skal sammenligne antal ansatte med »X arbejde« eller »Y problem« med antal medarbejdere med samme arbejde/problem for tre år siden. Og dertil inkludere hhv. ekskludere undergrupper af medarbejdere i disse optællinger.

Koder for problemomfang og sammenlignelighed efter ændring

Koderne er generelt tildelt spørgsmål for spørgsmål. I nogle tilfælde vil en ændring i et spørgsmål kunne påvirke muligheden for at sammenligne

efterfølgende spørgsmål i VOV1 og VOV2. Dette er der ikke taget hensyn til ved tildelingen af koder.

- ÆM Ændring mulig. Her kan formuleringen med fordel ændres. Koden anvendes fx ved formuleringer, der ikke er direkte tvetydige eller uklare, men fx unødvendigt komplekse og dermed svære at oplæse og forstå. Spørgsmålene *kan* bibeholdes i den oprindelige version, men de foreslåede ændringer vil fx hyppigt kunne modvirke, at spørgsmålene skal gentages eller forklares af interviewerens, og dermed kan der spares tid, penge og irritation.
- ÆA Ændring anbefales. Koden anvendes, hvor formuleringer er så uhensigtsmæssige, at de bør ændres. Fx anvendes der i mange tilfælde parentes og skråstreger i spørgsmål, der skal læses op. Her må de enkelte interviewere træffe et valg. Det er mere hensigtsmæssigt, at skemadesignerens træffer et valg én gang for alle. Interviewerens arbejde lettes, og alle respondenter kan tilbydes samme formulering. Koden anvendes ligeledes i tilfælde, hvor centrale specifikationer mangler eller er placeret i en kursiveret instruktion til interviewerens.
- Koden anvendes også ved spørgsmålsformuleringer, der kan være uklare eller tvetydige. I disse tilfælde vil det foreslåede alternativ repræsentere en formulering af, hvad sprogkonsulentens antager må have været spørgerens intention.
- ÆU Udeladelse eller erstatning anbefales. Koden anvendes i tilfælde, hvor det er uklart, hvad spørgsmålet – og svarene – belyser. Den anvendes også ved spørgsmål, som præsenterer respondenterne for en urealistisk svær svaropgave. Der er hyppigt ikke udformet alternative formuleringer. Enten fordi intentionen er uklar eller fordi problemerne ikke er af sproglig karakter.
- % Der er ikke foreslået nogle ændringer. Spørgsmålet anbefales bibeholdt som det er.

- SM Sammenligning mulig. Her antages de foreslåede ændringer ikke at påvirke sammenligningsmuligheden mellem VOV1 og VOV2.
- SP Sammenligningsmulighed evt. problematisk. Her kan det være problematisk at sammenligne svar på den oprindelige og den ændrede version, typisk fordi den oprindelige version kunne give anledning til flere forskellige forståelser. Koden er anvendt alle steder, hvor der kan være problemer. I en del tilfælde vil de forventeligt være små.
- SU Sammenligning umulig. Her antages det, at den foreslåede ændring vil gøre sammenligning umulig. Koden anvendes primært, hvor sammenligning er en logisk umulighed, fx når spørgsmål som i VOV1 spurgte til »antal medarbejdere med EGA« foreslås erstattet med spørgsmål, der spørger til »andel medarbejdere med EGA«.

Om omfang og distribution af afdækkede problemkilder

Opgaven er i udgangspunktet grebet an som en sædvanlig sproglig kvalitetssikring af et sæt spørgeskemaer – gennemført *inden* de bringes i anvendelse. Det betyder, at der er foreslået alternative formuleringer i *alle* tilfælde, hvor sådanne forekommer at være mere hensigtsmæssige end de anvendte formuleringer. Denne vurdering er foretaget, desuagtet, at skemaerne *har* været anvendt, og at undersøgelsen påtænkes gentaget med sigte på sammenligning mellem VOV1 og VOV2.

Som det vil fremgå, er der foreslået alternative formuleringer af stort set alle spørgsmål i materialet. Det er ikke uhørt, at en sproglig kvalitetsvurdering giver anledning til forslag om omformulering af mange spørgsmål i et spørgeskema. De alternative formuleringer indikerer, at spørgsmålene *kunne* formuleres mere hensigtsmæssigt – hvis der ikke var andre hensyn at tage end de sproglige. I mange tilfælde (ved kode ÆM) er en ændring ikke *nødvendig* for at sikre de enkelte svares kvalitet. Men de foreslåede ændringer forventes at kunne lette interviewets forløb og dermed resultere i kortere interview.

De visionsspecifikke skemaer er bemærkelsesværdigt forskellige både med hensyn til opbygning og sproglig kvalitet.

Opbygning af skemaer

En del skemaer arbejder fx med en tilbageskuende tidshorisont, som skal afdække fremdriften frem til interviewtidspunktet på den enkelte arbejdsplads. Andre skemaer anlægger ikke dette tilbageskuende perspektiv. Respondentopgaven er generelt ret vanskelig i de spørgsmål, der har en tilbageskuende tidshorisont, og det bør overvejes, om disse spørgsmål skal bibeholdes.

Nogle skemaer belyser omfanget af »X arbejdstype« eller »Y problem« ved at spørge til *antal* medarbejdere med »X arbejde« eller »Y problem«. I andre skemaer spørges der til, hvor stor en *andel* af medarbejderne, der har »X arbejde eller »Y problem«. I de tilfælde, hvor der ønskes information om udviklingen i omfanget af »X arbejde« eller »Y problem« anbefales det, at der spørges til *andel* af medarbejderne, frem for det absolute antal. Dermed vil svaret forventeligt være relativt meningsfuldt – også selv om der evt. har været en udvikling i det *samlede* antal medarbejdere på arbejdspladsen inden for det tidsrum, som sammenligningen angår. Alternativt kan der spørges til andel + ændringer i »problemets omfang« e.l.

Sproglig kvalitet

Der er stor forskel på, hvor ofte de forskellige koder for sproglig kvalitet er anvendt i de enkelte skemaer. Og der er også forskel på, hvordan de foreslåede eller anbefalede ændringer forventes at kunne påvirke muligheden for at sammenligne svarene fra VOV1 og VOV2.

I skemaet om **Indeklima** er hovedparten af spørgsmålene tildelt koden ÆM (ændring mulig). Her *kan* de fleste spørgsmål altså bevares. De fleste af de små foreslåede ændringer forventes ikke at ville påvirke muligheden for at sammenligne svar mellem VOV1 og VOV2.

I skemaet om **Arbejdsulykker** er de fleste spørgsmål tildelt koden ÆA (ændring anbefales), men i de fleste tilfælde vil de anbefalede ændringer forventeligt ikke gøre sammenligning mellem VOV1 og VOV2 problematisk.

I skemaerne om **Kemiske produkter** og **Børn & unge** anvendes koden ÆA også meget hyppigt, og her vil de anbefalede ændringer hyppigt have konsekvenser for muligheden for at sammenligne svar fra VOV1 og VOV2.

Skemaet om **Støj** er materialets svageste. Mange spørgsmål er tvetydige, uklare og svært forståelige, og svaropgaven er hyppigt urealistisk svær. Dertil er mange spørgsmål placeret uhensigtsmæssigt i det overordnede forløb. Der er anbefalet så mange ændringer og omrokeringer, at det er besluttet ikke at kode de enkelte spørgsmål hver for sig.

Skemaerne om **Psykisk arbejdsmiljø**, **Ergonomi** og de **53 generelle spørgsmål** placerer sig et sted midt i feltet.

Særlige problemer

Nogle af de visionsspecifikke skemaer indledes med filterspørgsmål eller centrale begrebsdefinitioner, som kan give anledning til uklarhed. Disse uklarheder kan have implikationer for »filtreringen« af respondenter for alle efterfølgende spørgsmål. Fx indledes skemaet om kemiske produkter med en spørgsmålsformulering, der ikke entydigt specificerer, om de efterfølgende spørgsmål handler om »kemikalier« eller – mere bredt – om »kemiske produkter« (som fx kemikalier, maling og rengøringsmidler). Denne uklarhed kan betyde, at relevante respondenter, som ikke forstår »kemikalier« bredt, ikke besvarer de efterfølgende spørgsmål. Det samme gælder begrebet »fritidsjob« i de første spørgsmål i skemaet om børn og unge.

Hvis disse problemer afhjælpes i VOV2, sådan at uklarhed undgås, så kan konsekvensen blive, at de efterfølgende spørgsmål besvares af en anden respondentkreds end i VOV1. Dermed kan sammenligning af svar på alle efterfølgende spørgsmål blive problematisk.

Skemaet om støj er karakteriseret ved urealistiske krav til respondenter. Her skal respondenterne beregne antal og udvikling i antal medarbejdere, der udsættes for støj på forskellige niveauer. Definitionerne på støj på forskellige niveauer samt fordringerne om at inkludere og ekskludere bestemte grupper i optællingen vil næppe generere valide svar. I dette skema er problemerne ikke primært af sproglig karakter, hvorfor der hyppigt ikke er foreslået alternative formuleringer.

Belysning af problemstillinger for de enkelte visioner i oprindeligt udbud

Afgørelse af, hvorvidt de enkelte problemstillinger er tilstrækkeligt belyst ved de stillede spørgsmål er ikke en rent sproglig opgave. Opgørelsen for de enkelte visioner er således kun indikativ og ikke baseret på en grundig granskning af relationer mellem spørgeskemaspørgsmål og problemstillinger.

Det generelle skema (spørgsmål 1-53)

Distribution af koder

Omtrent ligelig fordeling mellem spørgsmål, der kan forbedres (ÆM), og spørgsmål, som det *anbefales* at forbedre (ÆA). I en del tilfælde, hvor koden ÆA (ændring anbefales) er anvendt, vil den foreslåede ændring forventeligt ikke umuliggøre sammenligning mellem VOV1-version og den foreslåede version.

Særlige karakteristika og problemkilder

Der er afdækket en del problemer, hvor spørgsmål forudsætter forhold, som ikke er fastslået i forudgående spørgsmål. Fx spørgsmål, der forudsætter flere »sikkerhedsgrupper«, spørgsmål der forudsætter APV eller spørgsmål, der forudsætter, at der er kortlagt problemer. Der er foreslået en del nye filterspørgsmål samt omplacering af oprindelige spørgsmål for at afhjælpe disse problemer.

Der er afdækket en del problemer, hvor spørgsmål forudsætter forhold, som ikke er fastslået i forudgående spørgsmål. Fx spørgsmål, der forudsætter flere »sikkerhedsgrupper«, spørgsmål der forudsætter APV eller spørgsmål, der forudsætter, at der er kortlagt problemer. Der er foreslået en del nye filterspørgsmål samt omplacering af oprindelige spørgsmål for at afhjælpe disse problemer.

Kodeoversigt

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå				
				27	ÆM/SM		
				28	ÆM/SM		
				29		ÆA/SP	
1	ÆM/SM			30			Æ?
2	ÆM/SM			31		ÆA/SM	
3	ÆM/SM			32		ÆA/SM	
4	ÆM/SM			33		ÆA/SP	
5		ÆA/SP		34	%		
6	ÆM/SM			35	%		
7		ÆA/SM		36	%		
8		ÆA/SM		37		ÆA/SP	
9		ÆA/SP		38	ÆM/SM		
10	ÆM/SP			39	%		
11		ÆA/SP		40		ÆA/SP	
12	ÆM/SM			41		ÆA/SU	
13		ÆA/SP		42	%		
14		ÆA/SP		43	%		
15		ÆA/SP		44	ÆM/SM		
16	ÆM/SM			45	ÆM/SM		
17			Æ?	46	ÆM/SM		
18		ÆA/SP		47	ÆM/SM		
19		ÆA/SM		48	ÆM/SM		
20		ÆA/SP		49		ÆA/SU	
21		ÆA/SP		50	ÆM/SM		
22	ÆM/SM			50a			ÆU
23		ÆA/SM		50b			ÆU
24		ÆA/SP		51	ÆM/SP		
25	ÆM/SM			52	ÆM/SP		
26	ÆM/SM			53	ÆM/SM		

Foreslået rækkefølge af spm 1-53

1-6, 8-10, 12, 45, 7, 11, (13), 46, ++, (13), 14-16, 18, 17, 20, 19, ++, 21, 22-36, 42-44, ++, 40, 41, 37-39, ++, 47-53

Arbejdsulykker

Distribution af koder

Spørgsmålene er kodet blok for blok og ikke udsagn for udsagn. Alle blokke er tildelt koden ÆA (ændring anbefales). Der er mange formuleringer, der bør ændres for at lette interaktionen. Men i de fleste tilfælde forventes de foreslåede ændringer ikke at påvirke muligheden for sammenligning. Dog er der en del udsagn i spm. 56, der er tvetydige, hvorfor sammenligning mellem VOV1 og de foreslåede omformuleringer kan være problematisk.

Særlige karakteristika og problemkilder

Der er ingen spørgsmål med tilbageskuende tidshorisont til belysning af evt. fremdrift på den enkelte arbejdsplads.

Alle spørgsmål er udformet som udsagn med en 4-punkts svarskala. Det anbefales, at alle svarmuligheder præsenteres før hver ny blok af udsagn. Det vil forventeligt lette den efterfølgende interaktion og sikre, at alle respondenter producerer svar, der er relative til samme svarskala.

Belysning af problemstillinger i oprindeligt udbud

De fleste spørgsmål belyser respondentens vurdering af årsager til ulykker + holdninger til sikkerhed på arbejdspladsen. En del problemstillinger belyses ikke i ulykkeskemaet, men i generelt skema. Men her er det p.t. ikke specificeret, at der skal være fokus på sikkerhed.

1. Sikkerhed i forhold til APV og ledelsesbeslutninger	Belyses i spm. 34-44. Men p.t. UDEN fokus på sikkerhed
2. Sikkerhedsinspektioner med ledelsesdeltagelse + ansatte	?
3. Sikkerhedsaspekt i instruktion + efterlevelse	54f, 55a,b,c,e, (56m)
4. Holdninger til sikkerhed/ulykkesårsager	54a,b,c,e,g,h (56 alle)
5. Opfølgning på ulykker + nær-ved-ulykker	13-15, 26-28, 55e
6. Ansattes krav til sikkerhed	57f,g, 58c
7. Ledelsens hensyn til sikkerhed ved planlægning mv.	57a, 58d

Kodeoversigt

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå		
				55	ÆA-SM
				56	ÆA-SP
Tekst 5		ÆA-SM		57	ÆA-SM
54		ÆA-SM		58	ÆA-SM
Tekst 6		ÆA-SM		59	ÆA-SM

Børn og unge

Distribution af koder

Koden ÆA (ændring anbefales) er anvendt hyppigt grundet formuleringer, der er uklare eller tvetydige. Fx uklarhed ang. fritidsjob/job og formuleringer, der spørger til »de unge«, og som forudsætter, at »de unge« er en homogen gruppe med samme arbejde/arbejdstid/vilkår. I mange tilfælde vil afklaring af spørgsmålsformulering kunne betyde, at sammenligning mellem VOV1 og VOV2 bliver problematisk.

Særlige karakteristika og problemkilder

Der er ingen spørgsmål med tilbageskuende tidshorisont til belysning af evt. fremdrift på den konkrete arbejdsplads (bortset fra spm. 55).

Det er problematisk, at skemaet indledes med en uklar distinktion mellem børn og unge i fritidsjob og evt. unge ansatte, som har arbejdet som deres hovedbeskæftigelse.

Belysning af problemstillinger i oprindeligt udbud

1. Kendskab til arbejdsmiljølovgivning 70? (forudsættes)
2. Kendskab til arbejdsmiljømæssige risikofaktorer/effekter 63?, 69? (for unge på den konkrete arbejdsplads)
3. Kendskab til sikkerheds- og sundhedssystemet
4. Arbejdstidens længde og placering 60, 61, 62b,c
5. Løft, fysisk belastning og EGA 58g,h,j
6. Arbejde med maskintyper 58a,b
7. Ulykkesrisici, kemi + eksem 58c,d,e,l,
8. Tekniske hjælpemidler 62g
9. Trafiksikkerhed 58k, 62i
10. Introduktionsprogrammer for unge 62a
11. Information af forældre (64) ikke fokus på risiko-
XX: Alene, med penge 58i, 59, 62d,j
XX: Støj 58m
XX: Sikkerhedsudstyr 62h,i
XX: Episoder 65, 66, 67, 68

Kodeoversigt

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå		
				61	ÆA/SP
				62	ÆA/SP
Tekst 5		ÆA/SP		63	ÆM/SM
54		ÆA/SP		64	ÆM/SM
55		ÆA/SP		65	ÆA/SP
56		ÆA/SP		66	ÆA/SP
57		ÆA/SP		67	ÆM/SM
58		ÆA/SP		68	ÆM/SM
59		ÆA/SP		69	ÆM/SM
60		ÆA/SP		70	ÆA/SP

Ergonomi

Distribution af koder

Her er over halvdelen af spørgsmålene tildelt koden ÆA (ændring anbefales). I mange tilfælde vil de foreslåede ændringer forventeligt betyde, at VOV1 og VOV2 bliver usammenlignelige.

Særlige karakteristika og problemkilder

Der stilles spørgsmål med tilbageskuende tidshorisont.

Her spørges til *antal* medarbejdere med EGA/tungt arbejde før og nu.

Det anbefales at spørge til *andel* + evt. omfang af EGA/tungt arbejde

Belysning af problemstillinger i oprindeligt udbud

1. EGA/tunge løft i APV (59), (74) Kortlagt - ikke APV
2. Forebyggelse af skader fra EGA og tunge løft (58), 62, (63), 64, 65, 67, (72), 77, (79), (80), 81,
3. Viden om risiko og forebyggelse 73,
4. Beskrivelse af jobfunktioner mht forebyggelse (56), (68), (70)
5. Effekt af sikkerhedsarbejde (63), (79), 80,

Kodeoversigt

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå		
				67	ÆA/SM
				68	ÆA/SU/SM
				69	ÆA/SM
54		ÆA/SM		70	ÆA/SU/SP
55		ÆA/SP		71	ÆA/SP/SM
56		ÆA/SU/SP		72	ÆM/SP
57		ÆA/SU/SP		73	ÆA/SM
58	ÆM/SP			74	ÆA/SP
59		ÆA/SP		75	ÆM/SM
60	ÆM/SM			76	ÆM/SM
61	ÆM/SM			77	ÆM/SP
62		ÆA/SP		78	ÆM/SM
63		ÆA/SP		79	ÆA/SP
64	ÆM/SM			80	ÆA/SP
65		ÆA/SP		81	ÆM/SM
66		ÆM/SU/SM		82	ÆM/SM

Parallele spørgsmål

55	56	57	58	59	60	61	62	63	64	65	
69	70	71	72	74	75	76	77	79	80	82	

Indeklima

Distribution af koder

Hovedparten af spørgsmålene er tildelt koden ÆM (ændring mulig). Her *kan* de fleste spørgsmål altså bevares. De fleste af de små foreslåede ændringer forventes ikke at ville påvirke muligheden for at sammenligne svar mellem VOV1 og VOV1.

Særlige karakteristika og problemkilder

Der stilles spørgsmål med tilbageskuende tidshorisont. Der stilles spørgsmål til andel af medarbejdere med »X problem« + ændringer i problemets omfang.

Belysning af problemstillinger i oprindeligt udbud

1. Kortlægning og opfølgning 56, 58, 59, 62, 64, 65, 68, 70,71,74,76,77, 79, 81, 82, 85, 86
2. Produktoplysninger (97)
3. Hensyn ved køb af udstyr (96), 98
4. Vedligeholdelse 94, 95
5. Instruktion i klimaanlæg 87,88

Kodeoversigt

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå		
				76	ÆM/SM
				77	ÆM/SM
54	ÆM/SM			78	ÆM/SM
55		ÆA/SP		79	ÆM/SM
56		ÆA/SM		80	ÆM/SM
57	ÆM/SM			81	ÆA/SP
58	ÆM/SM			82	ÆM/SM
59	ÆM/SM			83	ÆA/SM
60	ÆM/SM			84	ÆA/SP
61		ÆA/SP		85	ÆA/SP
62	ÆM/SM			86	ÆA/SP
63	ÆM/SM			87	ÆM/SM
64		ÆA/SM		88	ÆM/SM
65	ÆM/SM			89	ÆM/SM
66	ÆM/SM			90	ÆM/SM
67		ÆA/SP		91	ÆM/SM
68	ÆM/SM			92	ÆM/SM
69	ÆM/SM			93	ÆA-SM
70	ÆM/SM			94	ÆM/SM
71	ÆM/SM			95	ÆM/SM
72	ÆM/SM			96	ÆM/SM
73		ÆA/SP		97	ÆA/SM
74	ÆM/SM			98	ÆM/SM
75	ÆM/SM				

Parallele spørgsmål

Varme mv.	54	55	56	57	58	59
Luftkvalitet	60	61	62	63	64	65
Lyd	66	67	68	69	70	71
Lys	72	73	74	75	76	77
Fugt	78		79	80	81	82

Kemiske stoffer

Distribution af koder

Stor overvægt af spørgsmål, der er tildelt koden ÆA (ændring anbefales). Der er mange uklare og tvetydige formuleringer.

Særlige karakteristika og problemkilder

Der stilles spørgsmål med tilbageskuende tidshorisont. Her spørges til *antal* medarbejdere med kemisk belastning før og nu. Det anbefales at spørge til *andel*, + evt. blot til, om problemet er blevet større eller mindre.

Mange spørgsmål har meget stor fordring om detaljeringsgrad. Fx spm. 72 og 80 om antal arbejdssteder med belastning/foranstaltninger samt spørgsmålene om, hvor mange ansatte der er udsat for x, y eller z nu og for tre år siden. Det belyses ikke, hvor mange ansatte der var i alt for tre år siden, og dermed kan svarene ikke belyse udvikling i belastning.

»Kemikalier« i spm. 54 er et uheldigt begreb, HVIS der ønskes oplysninger om fx maling, rengøringsmidler o.l. Det anbefales generelt at bruge formuleringen »kemiske produkter«.

Det er muligt, at formuleringen »kemikalier« i spm. 54 kan have filteret relevante respondenter fra i VOV1. HVIS spm. 54 ændres til mere inklusiv formulering i VOV2, vil sammenligning mellem VOV1 og VOV2 kunne blive problematisk.

Der er ikke taget hensyn til dette overordnede problem i kodning for sammenlignelighed. Koderne for sammenlignelighed er tildelt spørgsmål for spørgsmål ved en isoleret vurdering af VOV1-version over for ny foreslået version.

Der mangler filtergange mange steder. Der er således blevet stillet mange redundante spørgsmål i VOV1. Dertil er det i nogle tilfælde uklart, om et »nej« betyder »nej« til spørgsmål eller forudsætning. Fx spm. 89 om

eksterne konsulenter ved målinger. Her kan »nej« betyde »ingen konsulenter« eller »ingen målinger«.

Belysning af problemstillinger i oprindeligt udbud

- | | |
|---------------------------------|-----------------------------------|
| 1. Kemi og APV | 71, 79, 82, 87 |
| 2. Dansk brugsanvisning | (64), 65, |
| 3. Tilgængelig brugsanvisning | 66, 67, 68 |
| 4. Medarbejderes viden om kemi | |
| 5. Kemi og sikkerhed i oplæring | 69 |
| 6. Jævnlig undervisning | (70) |
| 7. Arbejdshygienisk måling | 88, 89 |
| 8. Sanktioner ved brud | 90 (vist næsten ordret fra udbud) |

Kodeoversigt

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå			
			71		ÆA/SP	
			72			ÆU
			73		ÆA/SU/SM	
			74		ÆA/SU/SM	
Tekst 5		ÆA/SM	75		ÆA/SU/SM	
54		ÆA/SP	76		ÆA/SU/SM	
55		ÆA/SP	77		ÆA/SU/SM	
56		ÆA/SP	78		ÆA/SU/SM	
57		ÆA/SP	79		ÆA/SP	
58		ÆA/SP				
59		ÆA/SP	80			ÆU
69		ÆA/SP	81	%		
61		ÆA/SP	82		ÆA/SP	
62		ÆA/SP	83		ÆA/SP	
63	ÆM/SM		84		ÆA/SP	
64	ÆM/SM		85		ÆA/SM/SP	
65	ÆM/SM		86		ÆA/SP	
66	ÆM/SM		87		ÆA/SP	
67		ÆA/SP	88	ÆM/SM		
68		ÆA/SP	89	ÆA/SP		
69		ÆA/SM	90		ÆA/SP	
70	ÆM/SM					

Parallele spørgsmål

	Antal	Antal 3 år	Mængde 3 år	Ansatte	Ansatte 3 år
Mærket		55	56	73	74
Kræft	57	85	59	75	76
Organisk	60	61	62	77	78

Psykisk arbejdsmiljø

Distribution af koder

Ligelig fordeling mellem spørgsmål, der kan bibeholdes (ÆM), og spørgsmål, hvor ændring anbefales (ÆA). De fleste af de spørgsmål, som

bør ændres, er tvetydige eller uklare i VOV1, således at sammenligning mellem VOV1 og foreslået alternativ kan blive problematisk.

Særlige karakteristika og problemkilder

Der er et par spørgsmål med tilbageskuende tidshorisont om prioritering af psykisk arbejdsmiljø.

Spørgsmål om udbredelse af problemer spørger til andel af medarbejdere og ikke antal.

Tvetydige filterspørgsmål: Fx er spm. 79 – som er et filterspørgsmål – tvetydigt i VOV1. Afklaring af tvetydighed (igangværende vs. initiativer overhovedet) vil betyde en anderledes filtrering af respondenterne end i VOV1, hvorfor sammenligning mellem svar på spm. 79 kan blive problematisk.

Belysning af problemstillinger i oprindeligt udbud

- | | |
|---|--|
| 1. Psyk og APV | 69 |
| 2. Viden om Psyk | 76 |
| 3. Væsentlige problemer, forebyggelse, effekt | 61-68, 72, (79, 80 - ikke konkret), 84, 85, 86 (56,57) |
| 4. Ændret praksis i forhold til psyk | |
| 5. Behandling af konkrete emner | Fokus på »psykiske arbejdsproblemer«
Generelt. Meget lidt konkret |
| 6. Behandlings form | (82,83) |

Spm.nr.	Ændring mulig	Ændring anbefales	Udgå			
				70	ÆM-SM	
				71		ÆA-SP
Tekst 5		ÆA-SM		72	ÆM-SM	
54	ÆM-SM			Tekst 7	ÆM-SM	
55	ÆM-SM			73		ÆA-SP
56		ÆA-SP		74		ÆA-SP
57		ÆA-SP		75		ÆA-SP
58	ÆM-SM			76		ÆA-SM
59	ÆM-SM			77		ÆA-SP
Tekst 6	ÆM-SM			78		ÆA-SM
60		ÆA-SP		79		ÆA-SP
61		ÆA-SM		80	ÆM-SP	
62	ÆM-SM			81	ÆM-SP	
63		ÆA-SP		82	ÆM-SP	
64		ÆA-SM		83		ÆA-SP
65		ÆA-SM		Tekst 8	ÆM-SM	
66	ÆM-SM			84		ÆA-SP
67		ÆA-SM		85	ÆM-SP	
68	ÆM-SM			86	%-SP	
69	ÆM-SM			87	%-SP	

Støj

Distribution af koder

Spørgsmålene er ikke tildelt koder. Skemaet er som helhed ikke velfungerende (se nedenfor), og de fleste spørgsmål bør omformuleres eller evt. erstattes af helt nye spørgsmål og spørgeforløb.

Særlige karakteristika og problemkilder

Mange ulogiske forløb, hvor det er sandsynligt, at respondenter vil svare med løse gæt, frem for at foretage de beregninger (subtraktioner), som spørgsmålene fordrer. Det anbefales at lette respondentopgaven generelt og anlægge et detaljeringsniveau, som er realistisk.

Mange komplekse spørgsmålsformuleringer, der let oplæses forkert og let misforstås.

Som fx: »Når indkøb af støjende maskiner besluttes, anvendes leverandøroplysninger angående støj?«

Belysning af problemstillinger i oprindeligt udbud

Ikke vurderet grundet ovenstående forhold.

Summary

Evaluation of the project »Monitoring Preventive Work Safety and Health Measures at Workplace Level«

Issued January 2005

by Olaf Rieper and Leena Eskelinen

This report contains the results of an evaluation of the first wave of »Monitoring Preventive Work Safety and Health Measures at Workplace Level« (Virksomhedsovervågning – VOV1) which the Danish Working Environment Authority initiated in 2000 as part of the Government's action programme for a clean working environment by 2005. The purpose of the monitoring is to describe and analyse the development in enterprises' preventive activities with regard to the seven focus areas included in the action programme. These focus areas cover such aspects as chemicals, child and youth labour, heavy lifting, indoor climate, monotonous and repetitive work, noise, psychosocial working environment and work accidents.

The first wave of the project was carried out in 2000-2002. It was divided into three parts dealing with development and validation of a method of monitoring the preventive activities in the enterprises as well as analysis and reporting, data collection from telephone interviews with approx. 4,600 enterprises (with interview of both a manager and an employee from each enterprise) and observations made at the worksite of approx. 10 per

cent of the enterprises that participated in the telephone interviews. As the first wave of the project, VOV1 was also to serve as a point of reference (status measurement) for future studies of the enterprises' preventive activities.

The purpose of the evaluation of VOV1 is to ensure adjustment of the design and methods for the second wave of the project (VOV2), as well as to evaluate to what extent VOV1 has achieved its goal. The evaluation is an integrated part of the second wave of the project, which is to be carried out in the autumn of 2003. The evaluation is based on peer review of the seven focus area reports, including design, sampling and data collection methods. The evaluation also included interviews with the persons who ordered and carried out the study.

It is the assessment of the evaluator that the basic design, which includes telephone interviews with a manager and an employee at a large number of workplaces supplemented with observations at the worksites, is appropriate. However, a number of adjustments and changes need to be made to the sampling procedures, questionnaires and observation forms. This will limit the possibilities to compare VOV1 with future waves of the project.

Regarding whether VOV1 has achieved its goal, it is the assessment of the evaluator that it was only partially achieved. As justification for this assessment, the evaluator draws attention to lack of documentation of the data, the methods, the data collection and flaws in the analyses of the data collected on the seven focus areas.

VOV1 constitutes a platform and an experience base for future monitoring projects, but the possibility of using VOV1 as a point of reference is assessed to be limited. Thus, there is a need for clarification and revision of a number of design and method-related aspects regarding VOV1 with a view to developing a reliable and valid methodology appropriate to future studies of the development of preventive work safety and health measures at workplace level.

Noter

1. Det fremgår ikke helt klart af det udleverede materiale, hvorvidt der faktisk er blevet stratificeret på virksomhedsstørrelse. I den forbindelse skal det nævnes, at der i materialet fra de udførende institutioner burde have været en samlet præsentation af de indsamlede data med vægtede og uvægtede antal virksomheder fordelt på samtlige faktisk anvendte strata.
2. Spørgsmålene blev ved den kollegiale vurdering opdelt i fire kategorier (jf. bilag 3): *faktuelle spørgsmål* (dvs. spørgsmål, som kun indeholder lidt skøn fra svarpersonens side, og hvor genstanden kan observeres); *skønsmæssige spørgsmål* (genstanden kan i princippet observeres, men udstrækning i rum eller tid gør, at skønnet ved besvarelsen bliver stort); *vurderingsmæssige spørgsmål* (spørgsmålet drejer sig om årsagsvirknings-forhold, der er svært tilgængelig for observation og spørgsmål, som indeholder et stort element af fortolkning/vurdering); *normative/holdningsmæssige spørgsmål* (genstanden er svarpersonens indstilling, ønsker, prioriteringer o.l.), som ikke kan gøres til genstand for observation andet end indirekte).
3. Det fremgår ikke af tabel 4.1 i Bach et al. (2002), hvorvidt der er tale om vægtede eller uvægtede tal.
4. Det fremgår ikke helt klart af det udleverede materiale, hvorvidt der faktisk er blevet stratificeret på virksomhedsstørrelse. I den forbindelse skal det nævnes, at der i materialet fra de udførende institutioner burde have været en samlet præsentation af de indsamlede data med vægtede og uvægtede antal virksomheder fordelt på samtlige faktisk anvendte strata.
5. Det kan naturligvis ikke lade sig gøre, at begge andele i virkeligheden er 0,5 og samtidig er forskellige. Men antagelsen giver anledning til det mest usikre skøn og dermed en overgrænse for krav til stikprøvernes størrelse.
6. Hvis fx $n = 179$ er $\Delta \hat{p}_{H_0, \text{afvises}} = 0,15$.