

Udviklingskontrakter mellem Forskningsministeriet og Universiteter

Belyst med inddragelse af erfaringer fra Benchmarking og Evalueringer

**Karen Siune
Bettina Damm**

**Notatet er blevet til foranlediget af henvendelse
fra forskningsminister Jan Trøjborg**

**Notat fra
Analyseinstitut for Forskning
1998/3**

Indholdsfortegnelse

1. Indledning	4
2. Definitioner på benchmarking.....	6
3. Typer af benchmarking	9
4. Anvendelse af benchmarking	12
4.1 Eksempler fra private danske virksomheder	14
4.2 Benchmarking anvendt i Forskning og Udvikling	16
4.3 Danske eksempler på benchmarking anvendt i Forskning og Udvikling.....	16
4.4 Benchmarking anvendt i uddannelsessektoren.....	18
5. Resultatstyring og Kriterier	19
5.1 Kvalitet og resultatstyring i den Offentlige sektor i Danmark.....	19
5.2. Kriterier for de videregående uddannelser.....	22
6. Overvejelser over kriterier til kontrakter for universiteter	24
6.1 Forskningsprojekter	25
6.2 Produktivitet	26
6.3 Formidling	26
6.4 Samarbejde med omverdenen.....	26
6.5 Rådgivning på forskellige niveauer	27
6.6 Deltagelse i bedømmelsesudvalg, internt/eksternt	27
6.7 Tidsskriftsredaktionsarbejde, nationalt/internationalt.....	27
6.8 Afholdelse af konferencer/deltagelsesomfang	27
6.9 Effektivitet (produktion pr. given enhed)	27
6.10 Innovationsevne	27
7. Metodeovervejelser	29
8. Problemer med benchmarking i forhold til udviklingskontrakter	30
9. Sammenfatning vedr. benchmarking i forhold til udviklingskontrakter	32
10. Forskningsministeriets udspil til målsætninger	32

1. Indledning

Mål og resultatstyring er den fælles overskrift for en række initiativer, der de sidste år er taget for at sikre effektiv ledelse i staten. Blandt de centrale redskaber, som indtil nu er anvendt, er virksomhedsregnskaber, resultatkontrakter, direktørkontrakter og øget brug af resultatløn. Fælles for disse redskaber er, at de bygger på et princip om, at god ledelse kræver, at der opstilles resultatmål samt foretages løbende resultatmålinger.

Denne stigende fokusering på mål og resultatstyring er et led i den stigende interesse for kvalitet såvel nationalt som internationalt. Et af de steder, hvor man har taget fat på emnet, er i Undervisningsministeriet under overskriften "Kvalitet der kan ses". Motivationen for dette projekt er, foruden den generelle tendens for kvalitet i såvel det offentlige som i den private sektor, at viden vil spille en afgørende rolle i fremtidens samfund. Derfor er der stor interesse for uddannelse, og stor interesse for hvad vi får ud af at uddanne os. Samtidig er der fokus på internationalt sammenlignelige færdighedsundersøgelser. For at øge sammenligneligheden i kvalitetsvurderinger mellem uddannelsesområderne har ministeriet udarbejdet et grundlag for en fælles metode. Den foreslåede kvalitetssystematik skal være et redskab til at evaluere og udvikle kvalitet, ligesom den skal videregive informationer, der kan hjælpe til at vurdere, hvor en fremtidig indsats giver bedst kvalitet i forhold til de givne ressourcer. Systemet skal i følge udmeldinger fra Undervisningsministeriet tjene som inspiration til udvikling og dynamik på den enkelt institution/uddannelse, således at det bidrager til de visioner, der sikre engagement og fremtidssikring.

Forskningsministeriet er et andet sted, hvor man har taget fat på emnet i forbindelse med beskrivelsen af de overordnede rammer for arbejdet med udviklingskontrakter. Forskning er en international aktivitet, der er vanskelig at måle kvantitativt og statistisk, så snart man går udover de traditionelle mål i form af penge bevilget til forskning, antal timer/årsværk brugt til forskning og antal publikationer. Evalueringer er traditionelt måden, man har vurderet forskningsmiljøers kvalitet på, det være sig i form af interne evalueringer eller som oftest i Danmark i form af eksterne evalueringer med inddragelse af internationale evalueringsudvalg. Indgåelse af kontrakter mellem universiteter og ministeriet/ministerier er noget nyt og uprøvet i Danmark. Forskningsministeriet har derimod etableret en erfaring med indgåelse af udviklingskontrakter med diverse sektorforskningsinstitutter. Forskningsministeriet har i erkendelse heraf i diverse skrivelser til universiteterne henvist til, at der vil være et behov for et udviklingsarbejde før selve arbejdet med at indgå kontrakter med de enkelte institutioner kan begynde. Udviklingsarbejdet skal ifølge ministeriet især koncentrere sig om følgende:

- Hvilke typer mål ønsker ministeriet medtaget i kontrakterne
- Hvordan formuleres målene, således at målopfyldelsen kan registreres
- Hvilke målemetoder skal udvikles.

Dette notat gennemgår forskellige typer af tilgange til at løse de problemer, der knytter sig til de nævnte tre problemstillinger. Evalueringer har været den traditionelle metode til at vurdere kvalitet af forskning, og notatet forholder sig til en række af de kriterier, der har været inddraget i danske og udenlandske evalueringer af forskning; men i forlængelse af den principielle og metodiske diskussion om kvalitet, mål, udviklingskontrakter og resultatstyring, en diskussion der for tiden foregår mange andre steder i Europa end i Danmark, finder man i dag ofte begrebet benchmarking. I dette notat er der i de følgende afsnit en gennemgang af benchmarking; hvad dækker begrebet og

på hvilken måde er det relevant at inddrage benchmarking i forbindelse med udviklingskontrakter? Benchmarking kan kort beskrives som en metode, der kan anvendes til systematisk at søge efter viden hos sig selv og andre, med det formål at lære nye og bedre måder at løse opgaven på. Den nye viden, der opnås gennem benchmarking, anses for yderst velegnet til at tilføre en organisation nye synspunkter og overvejelser til den stadige udvikling af den strategiske vision og det fælles fremtidsbillede, der skal være gældende for den enkelte virksomhed, institution eller uddannelsessted.

Benchmarking indeholder to væsentlige procesforløb. Først analyse og systematisk sammenligning af arbejdstilrettelæggelse og resultater, og derefter en refleksion, som bygger på læring og inspiration til at indrette arbejdstilrettelæggelsen og ressourceanvendelsen på den mest effektive måde i forhold til de ønskede mål og resultater.

Overvejelser over benchmarking kan indgå som et led i arbejdet med at opstille udviklingskontrakter, ligesom principper kendt fra benchmarking kan indgå i arbejdet med at opfylde en udviklingskontrakt, men da benchmarking især er praktiseret i private virksomheder, kræver det særlig opmærksomhed omkring de målsætninger og deraf tilknyttede problemer, som gælder for offentlige virksomheder, herunder universiteter, førend man går i gang med at anvende benchmarking kriterier i forbindelse med udviklingskontrakter.

I dette notats første del vil benchmarking blive præsenteret vha. forskellige definitioner. Derefter beskrives eksempler på forskellige anvendelser af benchmarking, hvoraf en del er taget fra det private erhvervsliv. Derefter afgrænses der til at se på anvendelsen af benchmarking i den offentlige sektor specielt med fokus på benchmarking af uddannelse og derefter med fokus på forskning & udvikling. Undervejs kombineres de tankegange, som man finder i en række nuværende projekter omkring resultatstyring og kvalitet, med anvendelsen af benchmarking.

Notatets gennemgang af benchmarking slutter med en sammenfatning af problemer og muligheder ved at anvende en benchmarking-tilgang til udviklingskontrakter. I det efterfølgende afsnit er der fokuseret på kriterier, der traditionelt indgår i evalueringer af forskningsmiljøer. En del af disse kunne umiddelbart tænkes indarbejdet i udviklingskontrakter for universiteter o.l. med respekt for de opgaver, som netop er henlagt til denne type af institutioner. Dermed kan det hævdes, at man er nået et skridt længere på vejen i besvarelsen af spørgsmålene fra ministeriet: 1) hvilke typer af mål, 2) hvordan formuleres målene, således at målopfyldelsen kan registreres, mens punkt 3) hvilke målemetoder skal udvikles for at opnå det for ministeriet implicite mål at måle kvalitetsforbedringer, står tilbage med mindre man vælger at anvende de fra evalueringer kendte metoder med deltagelse af evaluatorene.

2. Definitioner på benchmarking

Slår man op i ordbogen fra B. Kjærulff Nielsen: "Engelsk-dansk ordbog: Gyldendals store røde ordbøger" finder man følgende forklaring:

Benchmark:

1. {nivelements}fikspunkt (hvorfra målinger kan foretages):

stake with ~marks: kotepæl

2. (fig.)

a) referencepunkt, kriterium, standard

spec: edb: standardiseret problem brugt som sammenligningsgrundlag for datamat-system eller ~program

b) economic ~marks = økonomiske basistal for visse vigtige faktorer (fx: olie, stål, lånerenter)

Benchmark bond: toneangivende obligation.

Den leksikalske betydning af "benchmark": fikspunkt (hvorfra målinger kan foretages); udgangspunkt; sammenligningsgrundlag; norm; rettesnor.

I det efterfølgende afsnit vil der blive fremstillet både danske og udenlandske definitioner på benchmarking. Da der er citeret direkte fra original teksterne vil afsnittet være præget af en blanding af dansk og engelsk.

- Origins of benchmarking and its meaning. "A bench mark is a mark whose height, relative to ordnance datum has been determined by levelling". We have established that a bench mark is extremely worthy of emulation. "Benchmarking" is a word which appeared as it were over the Total Quality horizon two or so years ago.¹
- Benchmark - A surveyor's mark indicating a point in line of levels; a standard or point of reference. (Benchmarking has become synonymous with the means of identifying the benchmarking).¹
- Benchmarking today: **"Benchmarking is the continuous process of measuring products, services and processes against the strongest competitors or those renowned as world leaders in their field"**.¹

Benchmarking can be described as:

' the continuous process of measuring products, services and practices against leaders, allowing the identification of best practices which will lead to sustained and superior performance.'

More simply benchmarking is:

'finding and implementing best practice'.²

"Benchmarking are measures or standards, benchmarking is the process of understanding key issues within your own organisation and learning from others who are better".²

¹ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

² "Benchmarking for continuous improvement in the public sector" by John Bullivant (1994).

- **Benchmarking is about improving competitive position and using best practice to simulate radical innovation rather than seeking minor, incremental improvement on historic performance.**³

In our view, benchmarking is about taking action, not collecting data and comparing it euphemistically or pessimistically.³

A loose definition is probably the best to use:

Broadly speaking benchmarking is the search for best practice, and the subsequent translation of this best practice into use in the organisation. The whole approach often uses comparative performance statistics to identify either areas for improvement or the better performers amongst a group of benchmarkers.⁴

Benchmarking is:²

- innovative
- pareto based
- customer focused
- success orientated
- comparison with the best
- an improvement process
- empowerment
- sharing and learning
- breakthrough thinking
- about implementing
- developing healthy services
- energising

Benchmarking is not:²

- new
- cost cutting
- just competitive analysis
- number crunching
- site tours
- just copying or catch-up
- spying or espionage
- lone ranger activity
- quick and easy
- stealing ideas
- concerned with average performance⁵.

³ "Benchmarking customer service" by Glen Peters (1994).

⁴ "Benchmarking for best value in local authorities" by Mary Bowerman and Amande Stephens (1997).

⁵ "Benchmarking for continuous improvement in the public sector" by John Bullivant (1994).

In summary, benchmarking is very much an opportunity for an organisation to learn from the experience of others.⁶

To conclude, benchmarking is not a one-off analysis. It provides insights into a business and challenges its usual methodology. Implementation of the results and findings on the willingness to change and to adapt to new ways of doing business.⁶

Dansk Industri's definition:

Benchmarking er et ledelsesværktøj, hvor virksomhedens ledelse, ved involvering af medarbejdere systematisk måler og evaluerer egne resultater og processer i forhold til førende virksomheders resultater og processer, for her igennem at forbedre egne resultater og processer markant.

Gennemgående betragtes benchmarking som et værktøj til at opnå de mål, man har sat sig, og benchmarking skal ikke opfattes som et mål i sig selv. Der er uden tvivl tale om en dynamisk proces, som fortsætter "uendeligt", og som skal sikre kontinuerede forbedringer og kontinuert indlæring. De følgende citater er medtaget for at vise hvad benchmarking kan være et værktøj til:

- Benchmarking is the tool for educating, communication, for giving people the freedom to ask questions, empowering them, and building trust and loyalty⁷.

I en del af litteraturen finder man værktøjet benchmarking inddelt i 2 komponenter, en viden (informations) del og en evne/ressource/kompetence del.

- Stage 1: **the Demand**. (scanning the business .. find level of demand). Benchmarking can be a very powerful tool in building a good understanding of market conditions (customer, global markets, shareholder, environment, technology and time/speed => Intelligence building benchmarking).
- Stage 2: **the Offering**: Benchmarking can be a very powerful tool in building a capability of delivering goods and services. Capability represent levels off energy, strength and competencies that any organisation should be able to control and harness effectively. (creativity and innovation, teamwork, streamlined processes, technology, measurement and a culture of continuous improvement).⁷

Samtidig finder vi inddelingen af benchmarking i 2 niveauer, den strategiske og den operationelle benchmarking. Det der er vigtigt på det strategisk plan er, at definere hvad vi vil (mål), hvad vi kan (evner/kompetencer) og så anvende benchmarking til at definere det "gap" (den forskel) der forekommer, samt til at prioritere hvad der skal satses på først, altså hvordan vi skal prioritere vores ressourcer. For at kunne anvende benchmarking til at definere dette "gap" kræves kendskab til en detaljeret forståelse for, hvordan resultaterne er opnået, og hvilke processer, praksis og metoder mv. som leverer hvilket resultat. Hele denne forståelse af eget system er nødvendig før sammenligninger med andre kan foretages. Det er på det operationelle plan, at vi skal finde forståelsen for de bedste praksiser eller processer som hjælper andre med at opnå verdensklasse udformning (performance). Der findes altså mange typer af benchmarking.

⁶ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

⁷ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

3. Typer af benchmarking

Nedenfor er listet et eksempel på nogle af dem hentet fra Zairi & Leonard's bog:

- Konkurrencedygtig benchmarking
- Funktionel benchmarking (sammenligne funktioner ex. distribution, service etc.)
- Intern benchmarking (intern søgning efter et startpunkt - her forekommer ingen sammenligning med ekstern virksomhed).
- Generisk benchmarking (fokus på multifunktionelle virksomhedsprocesser - kontinuert bestræbelser på at sammenligne funktioner og processer med dem, som er "best in class", lige meget hvor de findes).⁷

Kriterier brugt til benchmarking:

- Strategisk fokus (fastsættelse af mål og udvide målene)
- Operationel fokus (processen og den udformning)
- Kundefokus (markedsdrevet strategier - værdier for slutkunden)
- Procesfokus (styrker i organisation)
- Sammenkædet med TQM (intern metode til problemløsning og kvalitets forbedring)
- Kontinuert forbedring
- Kontinuert indlæring (nyhed og innovation - "gøre de rigtige ting rigtigt")⁸

Benchmarking er ofte kædet sammen med kvalitet. Ofte finder vi begrebet benchmarking i forbindelse med TQM (total quality management), som er navnet på et koncept eller opfattelse, der kommer fra Japan, og som drejer sig om kvalitet ikke bare i forbindelse med produktionen, men i alt hvad man foretager sig samt bevidstheden og forståelse af vigtigheden af kvalitet. Kundetilfredshed og kontinuerede forbedringer er nogle af hovedelementerne i TQM. Benchmarking er kun en del af TQM.

Definition på TQM: "TQM is a concept which embraces the whole organisation, not just the manufacturing of the product but also design, delivery, billing and service. TQM is a commitment to a company-wide culture where everyone is clear of the direction and objectives of the organisation and the work in support of focused on customer satisfaction and continuous improvement. However there is no blueprint for TQM and organisations must create their own approach to TQM perhaps by learning from others and by selecting from the wide range of quality and management practices and tools."⁹

Som det er fremgået tidligere er benchmarking ikke noget nyt, komponenterne i benchmarking er nærmere et skelet til fokusering på emner, som er signifikante for virksomhedens overlevelse og udvikling. Benchmarking er blevet et buzz-word i 90'erne for noget, man har anvendt i mange år. Det betyder, at mange i dag anvender benchmarking uden at være bevidst om det.

Benchmarking er et vigtigt værktøj til at assistere ledelsen af ændringer (management of change), men benchmarking kræver/forudsætter en (commitment) forståelse for 'at gøre ting bedre', og det kræver en organisation, som allerede har undersøgt dens interne processer og deres forhold til virksomhedens mål, kundetilfredshed osv. Dataindsamling er kun en del af succesfuld benchmar-

⁸ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

⁹ "Benchmarking for continuous improvement in the public sector" by John Bullivant (1994).

king. Det, som er mest lige så vigtigt, er forståelsen af hvordan din egen organisation fungerer. Analysen af processen og dens fejlpunkter vil vise, om dataindsamlingen måler de rette ting. Benchmarking hjælper til at fokusere på ændringsprocessen og sikre ved sammenligning med andre, at de nye mål er relevante og opnåelige og at de foreslåede måder at gøre tingene på virker. Benchmarking hjælper med at definere aktiviteter på procesniveau. På dette niveau kan sammenligninger lettere laves mellem industri sektorer og på tværs af organisationer eller internationale grænser. Benchmarking betyder ikke bare at kigge på praksisovre hos naboen eller hos samme autoritet. At være den førende i klassen betyder ikke nødvendigvis, at man er den bedste i klassen. Benchmarking er en proces til kontinuerlige forbedringer og betyder at identificere og opnå 'den bedste af den bedste' status for ethvert af de valgte områder.

Kort sagt: nøglen til succesfuld benchmarking ligger i følge John Bullivant i forståelsen af hvad organisationen står for ved at kortlægge processen og måle udformningen/præstationen, samt i at søge den højeste udformning og sikre dens støtte. Endelig betyder det effektiv projektmanagement til implementering af de nødvendige ændringer. Hele metoden til succesfuld benchmarking forudsætter/kræver commitment og anvendelse.

I Dansk Industri's rapport om benchmarking finder man en tilsvarende opfattelse af hvilke betingelser, som er vigtige for at opnå succesfuld benchmarking.

Betingelser for succesfuld benchmarking:

Kend dit eget præstationsniveau og arbejdsproces og dokumenter disse.
Find en eller flere virksomheder som med garanti er meget bedre end dig på det felt, du ønsker at forbedre og få deres tilsagn om medvirken.
Kombiner det bedste fra de bedste - det betyder lad dig inspirere til nytænkning.
Opstil mål for procesændringer der har betydning for virksomhedens konkurrenceevne.
Afsæt nødvendige ressourcer til implementering af den nye adfærd.
Inddrag medarbejderne fra begyndelsen af processen.

Det er vigtigt at gøre sig klart, at en ting som commitment ikke er noget man får ved at bede om det, commitment er noget, man gør sig fortjent til gennem seriøst arbejde.

Xeron er nævnt som eksempel i næsten alle dele af litteraturen (specielt i den udenlandske litteratur). De startede allerede i 1979 med studier og målinger af konkurrenter, 'best practices' mv. Alt så hele den proces vi i dag kalder benchmarking. Deres metoder er i dag foregangsbillede for mange. Der henvises i litteraturen også til andre metoder, men fælles for dem er en planlægningsfase, en analysefase og en handlingsfase.

Nedenstående eksempel viser Xeron 12 Step til succesfuld benchmarking, og under det eksempel er Vaziri's måde vist.¹⁰

¹⁰ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

Figur 1. Rank Xeron's metode

Plan

- Step 1: Select Subject Area
- Step 2: Define Process to Benchmark
- Step 3: Identify Potential Benchmarking Partners
- Step 4: Identify Data Required

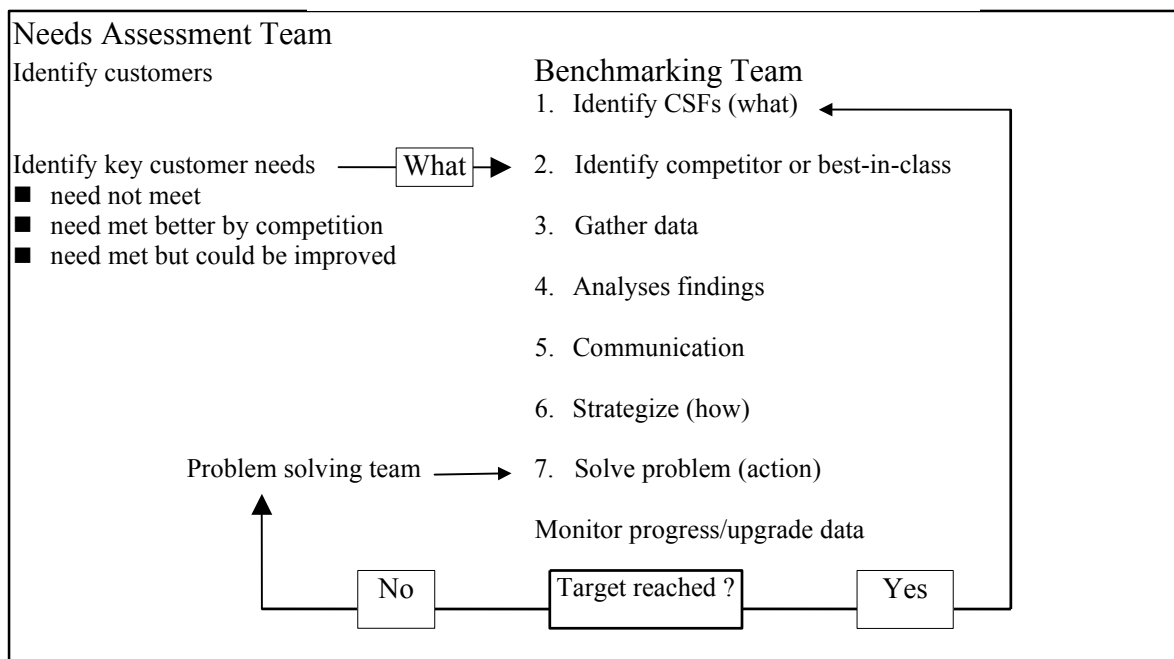
Analysis

- Step 5: Collect the Data and Select Benchmark Partners
- Step 6: Determine the Gap Compared to Benchmark
- Step 7: Establish Difference in Process
- Step 8: Target Future Performance

Action

- Step 9: Communication and Commitment
- Step 10: Adjust Targets and Develop Corrective Improvement Plan
- Step 11: Implementation and Monitoring
- Step 12: Review Progress and Recalibrate

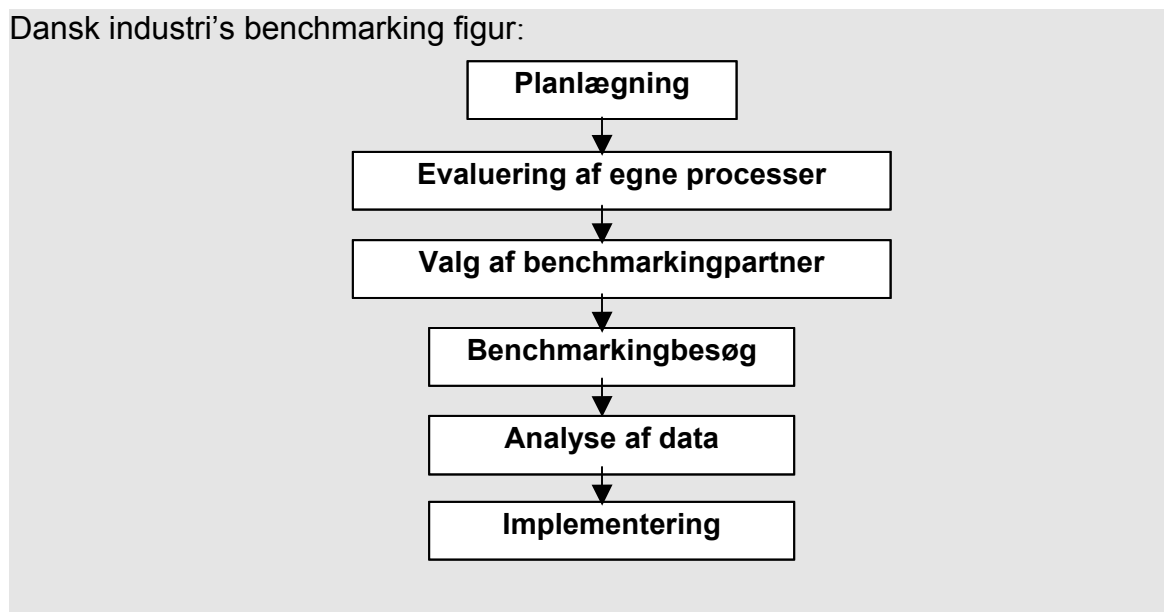
Figur 2. Vazari's metode



Hos Vaziri ses benchmarking som en komponent i det overordnede TQM program. Han ser ligeledes benchmarking som en måde at lede procesinnovation. Glen Peters ser ligesom bl.a. John Bullivant benchmarking som en proces til ledelse af ændringer. Men hans fokus er som tidligere nævnt på selve det at handle (take action).

Den sidste model, som vi kort vil beskrive, er den model Dansk Industri anvender. Modellen adskiller sig lidt fra de tidligere ved at være mere praksisorienteret. Den indeholder blandt andet et punkt som hedder benchmarkingbesøg, og som omfatter udviklingen af en spørgeguide og besøg hos benchmarkingpartner.

Dansk industri's benchmarking figur:



Ligesom i mange af de tidligere modeller, er det vigtigste element i denne model evaluering af egne processer. Selvom de forskellige forfattere placerer deres fokus forskelligt, er konklusionen den, at benchmarking er meget udbredt indenfor et hav af forskellige områder. I det næste afsnit beskrives kort nogle af de områder, hvor det anvendes, men hovedvægten vil være på benchmarking i den offentlige sektor.

4. Anvendelse af benchmarking

Benchmarking anvendes i fremstillingssektoren, i den finansielle sektor, i innovation, i procesinnovation, i forbindelse med human ressourcer, i benchmarking af den personlige funktion eller kundetilfredshed, men det anvendes også i marketing, i uddannelse og i forskning & udvikling.

Benchmarking og kvalitetssystemer anvendes endvidere i den offentlige sektor.

Firmaer i UK og Japan konkluderer¹¹, at benchmarking er det mest magtfulde værktøj til vurdering af industriel konkurrencedygtighed og for påskønnelsen af ændringsprocessen i virksomheder som stiler efter verdensklassepræstation. Men det, som måske er mest vigtigt, når vi ser på den offentlige sektors organisationer, er den stigende opmærksomhed på værdien og metoderne til delt information.

“The motivation for benchmarking in the public sector is not simply for competitive advantage, rather it is to support two traditional but important values, pride in your work and that of your organisation and vocation commitment of providing a public service and providing it well”¹²

Benchmarking har en speciel værdi i den offentlige service. Ikke kun er den Offentlige sektor underkastet et stigende pres til at ligne virksomheder, dvs. at være konkurrencedygtige og at demon-

¹¹ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

strere "value for money", men servicen er forventet at demonstrere lighed og retsindighed alt sammen under kritisk overvågning fra de ansatte, brugerne, revisorer, pressen, politikerne og det offentlige (public account committees).

Det følgende eksempel viser hvilke kriterier, der anvendes til vurdering af offentlig service på hospitalerne i Wales:¹²

- A. Effectiveness - for individual patients and the population as a whole;
- B. Equity - equal treatment for equal need;
- C. Access to service - in terms of location and waiting times, and so on;
- D. Relevance to need - for the whole community and appropriate care for the individual;
- E. Social acceptability - including responsiveness to views of patients and their families and friends;
- F. Efficiency - value for money.

De lokale myndigheder ser benchmarking som et magtfuldt værktøj til kvalitetsledelse og et redskab, som kan hjælpe både den og resten af den offentlige sektor til at møde kravene til en nyt regime. Under det nye regime vil der blive forventet af myndighederne, at de bekræfter helheden og overensstemmelsen af informationen, der fortæller om deres præstation. Samtidig forventes det, at de rapporterer offentligt, om de har nået "best value", og om dette bidrager til planer om "hjælpeaktion" (remedial action).

I Storbritannien har benchmarking været anvendt af de lokale myndigheder i flere år, før det er blevet et buzzword i den private sektor. Siden tidligt i 1980'erne har revisionskommissionen anvendt det til at identificere gode og dårlige præstationer mellem de lokale myndigheder. Kommissionen har så på baggrund af disse sammenligninger udviklet måder til udbredelsen af "god praksis" via publicering af råd og eksempler for gode processer og praksis. Alt dette har været en del af kommissionens fortolkning af lovmæssige beføjelser til at fremme "value for money" hos de lokale myndigheder. Det, der er unikt ved denne sektors benchmarking, er, at den er obligatorisk, og at hele metoden er forbundet med den eksterne revisionsproces, samt at benchmarkingresultaterne ofte er offentliggjort, samt at gode og dårlige præstationer er synliggjort i tabeller. Dette er forskelligt fra den private sektors benchmarking, hvor det hele er en intern affære. Artiklen hævder, at det er fordi, at det er obligatorisk, at det har været en succes. Begrænsninger ved denne obligatoriske benchmarking er, at den kun har været udført på et indledende stadie. Brug af sammenlignelige statistikker til at se, hvor præstationen kunne forbedres, er informativt, men det er nytteløst, hvis der ikke gøres noget for at forbedre præstationen. Det, som virkelig betyder noget, er den organisatoriske indlæring, som burde være resultatet af benchmarkers udveksling af ideer omkring processer, og de heraf medfølgende forbedringer. Et andet problem er, at bindingerne til den eksterne revision har betydet, at de lokale myndigheder kun opfatter benchmarking i regnskabsmæssig forstand og ikke ser muligheder for at anvende benchmarking som et ledelsesværktøj til kontinuerlige forbedringer.¹³

¹² "Benchmarking for continuous improvement in the public sector" by John Bullivant (1994).

¹³ "Benchmarking for best value in local authorities" by Mary Bowerman and Amande Stephens (1997).

Generelt er benchmarking vigtig i følgende områder:

- Virksomhedsplanlægning
- Policy deployment
- Løsning af virksomhedsproblemer
- Løbende forbedring
- Kundetilfredshed
- Kulturelle ændringer
- Kontinuert indlæring
- Managing by fact
- Strategisk udvikling

En integreret metode til måling finder vi ifølge Kaplan and Norton (1993) i "The balanced scoreboard", som dækker alle områder, der er kritiske for en virksomheds konkurrence. "The balanced scoreboard" finder man også i Økonomistyrelsens anbefaling til mål og resultatstyring. "The balanced scoreboard" indeholder 4 vigtige elementer, som ifølge Kaplan and Norton (1993) består af:

Kundeperspektivet (data inkl. service, kvalitet, ansvarlighed og omkostninger)

Procesevner (kvalitet, hastighed, levering, omkostninger og synergi)

Fokus på innovation (learning organisation - forbedring - people)

Det finansielle perspektiv med fokus på interessenterne.

4.1 Eksempler fra private danske virksomheder

Dansk Industri og de 9 virksomheder, som har deltaget i benchmarking projektet, har en lidt anden opfattelse af benchmarking end den, man typisk finder i litteraturen.¹⁴ De mener, at det grundlæggende og det virkelig nye element i benchmarking er ideen om, at man kan lære meget af virksomheder inden for helt andre bran-

cher end ens egen, så længe man fokuserer på en afgrænset proces.

Denne opfattelse udmønter sig blandt andet i, at Danfoss i forbindelse med benchmarking af uddannelsesafdelingen (med fokus på de administrative processer og arbejdsopgaver i forbindelse med tilmelding, planlægning og afvikling af kurser) ser "tilmelding af kurser" som "booking af reservationer" og dermed SAS som en mulig benchmarkingpartner. De sammenligner sig altså med et rejsebureau mht. lige denne funktion, fordi de er de bedste til booking. Deres målepunkter for denne del af uddannelsesafdelingen er:

- Personalefunktionens uddannelsesadministration er troværdig
- Imødekommenhed
- Hurtige tilbagemeldinger
- At få hurtig kontakt til afdelingen
- Rettidig information omkring kurser
- Informativt materiale
- Rådgivning/ det at kunne få yderligere oplysninger om kurser
- At vide hvem der laver hvad
- Personlig dataregistrering

For at evaluere deres egne processer laver de en undersøgelse, hvor hver af de ansatte i uddannelsesadministrationen vurderer de enkelte målepunkter (med værdi 1-5), dette sammenlignes så med kundernes (de yderligere ansattes) holdning til de samme punkter.

Først efter denne selvevaluering går de i gang med at sammenligne sig med andre.

Et tilsvarende eksempel finder vi hos Grundfos, som også lavede benchmarking af deres uddannelse. Deres fokus var på uddannelse af specialiserede medarbejdere som ingeniører og teknikerne.

Fire kritiske succes faktorer:

1. At medarbejderne kommer på det rigtige kursus
2. At kursusudbudet har høj kvalitet
3. At medarbejderne føler sig fagligt ajour
4. At kursusudbudet er baseret på Grundfos aktuelle behov

Under beskrivelsen af egne processer vidste det sig, at 300 kurser var forældede, så allerede inden sammenligning med andre fandt de frem til mulige forbedringer ved at fjerne disse kurser, så overskueligheden over kurser bliver større. Samtidig iværksatte de ændringer, som betyder, at det fremover sikres, at der ikke forekommer forældede kurser.

De målepunkter som Grundfos opstillede for de 4 kritiske succes faktorer var:

- Kursusevaluering
- Tid brugt til kursusaktivitet pr. medarbejder pr. år
- Fratrædelsesinterview
- Spørgeskemaundersøgelse
- Realiseret/budgettet

På baggrund af dette opstillede Grundfos følgende kriterier til mulig benchmarking partner:

- a) gode til uddannelse/træning/opsamling af viden
- b) af/for specialister
- c) med forskellige specialer
- d) indenfor foranderlige og dynamiske felter/områder

Dette medførte en bruttoliste på 15 mulige benchmarking partnere, som indbefattede:

- Små og store private virksomheder fra forskellige brancher
- Offentlige institutioner
- Uddannelsesinstitutioner

Det næste problem er så at finde ud af, hvem der så er den allerbedste af de 15 udvalgte. Et præstationsmål kunne være: Effektmåling af uddannelse - men dette forudsætter ifølge Grundfos, at benchmarking partneren har disse målinger, hvilket desværre ikke altid er tilfældet.

Konklusion på dette må være, at det er vigtigt at kende sig selv, før man sammenligner sig med andre. I forbindelse med sammenligningen er det vigtigt ikke at være snæversynet, den bedste til

¹⁴ "Benchmarking i praksis, Model og eksempler fra danske virksomheder", Dansk Industri (1997).

den pågældende funktion findes ikke nødvendigvis i samme branche som en selv, nytænkning er her et vigtigt nøgleord.

4.2 Benchmarking anvendt i Forskning og Udvikling

Et eksempel på benchmarking af forskning og udvikling, der er fundet i den udenlandske litteratur, bruger følgende indikatorer til etablering af sammenligninger:¹⁵

- Løbende forbrug
- F&U brugt per ansat
- F&U som procent af salg
- F&U som procent af indtjeningen
- Procent af salg per ansat.

I dette projekt undersøges, om der er en stærk korrelation mellem F&U og præstationen for virksomhederne på markedet.

Den praksis, som er fundet i Japan, som adskiller sig fra den i USA og Europa og som ofte fører til bedre forskning og udvikling, inkluderer:

- F&U aktivitet baseret på team samarbejde og udvikling af synergier og niveauer af kreativitet i samme retning.
- Opmuntringen til eksperimenter og kreativitet – Fejl opfattes som muligheder for forbedringer.
- Ledelsen af innovation er gennem gradvis stigende forbedringer, optimering af indlæring, forhøje/forøge evne and effektiv udnyttelse af tilgængelige ressourcer.
- Fokus på processerne, vedvarende evner og konsistens mht. til at levere succesfulde produkter og services.
- Fokusering på kunderne gennem vedvarende høje konkurrencestandarder, kvalitet, ansvarlighed og markedsfremtræden.

4.3 Danske eksempler på benchmarking anvendt i Forskning og Udvikling

Overvejelser omkring benchmarking foregår for tiden på en lang række danske sektorforskningsinstitutter, og sammenslutningen af direktører i sektorforskningen (SEDIRK) indsamler for tiden oplysninger herom.

I dette notat skal sektorforskningens overvejelser kun beskrives ved korte eksempler fra Danmarks og Grønlands Geologiske Undersøgelser (GEUS) samt fra Statens Veterinære Serumlaboratorium (SVS).

Følgende afsnit er hentet fra SVS beskrivelse som eksempel på intern benchmarking:

SVS har arbejdet meget bevidst med at sikre kvaliteten af forskningen. Der er i øjeblikket indført en intern kvalitetssikringsprocedure af de enkelte projekter, hvor en afdelingsleder og en medarbejder fra en afdeling evaluerer et eller flere projekter fra andre afdelinger. Evalueringen foretages på baggrund af de opstillede mål for det enkelte projekt. Det første udkast til evaluering sendes til projektgruppen, så der er mulighed for dialog mellem projektgruppe og evaluatorene. Efter dialogen samles alle evalueringerne, og de er tilgængelige internt på institutionen (der er planer om at lægge dem på intranettet). Konsekvensen af denne interne evaluering har været en større åbenhed omkring de enkelte projekter, en større viden på tværs af afdelinger om projekter, og på læn-

¹⁵ "Practical Benchmarking, The Complete Guide" by Mohamed Zairi & Poul Leonard (1994).

gere sigt bliver der mulighed for mere ensartethed i vurderingen. Den interne evaluering er et supplement til den evaluering, der hele tiden har fundet sted, både i forhold til de enkelte medarbejdere, i forbindelse med bedømmelse i relation til stillingsstrukturen og for det enkelte projekt/den enkelte artikel med feedback fra afdelingsleder og direktør (der godkender alle artikler/bøger)¹⁶.

Den form for evaluering eller benchmarking, som indgår i ovennævnte eksempel, knytter sig til enkelt-projekter.

De spørgsmål som det ville være relevant at stille til dette "interne evalueringssystem" er bl.a.:

- Hvordan sikrer man kvalitet - hvilke processer/funktioner indgår?
- Hvilke funktioner prioriterer man højest? Hvad er det vigtigt at blive bedre til?
- Hvilke kriterier anvendes til evaluering?
- Hvad er målene for de enkelte projekter?
- Hvad er situationen i dag, og hvor langt er man fra målene?
- Er der andre projektgrupper i virksomheden SVS, som har lignende funktioner? Og hvordan fungerer de?
- Kan man gøre tingene anderledes, bedre og hvordan? Har man indenfor projektet de midler, der skal til for at realisere de opstillede mål? Eller med andre ord, hvor langt kan man nå med de midler, man har? Og hvad er vigtigst at starte med?

Næste skridt kunne så være ekstern benchmarking - hvad gør de "bedste" på området? Problemet er selvfølgelig først at finde ud af: hvem er de "bedste"?

GEUS

Det ser ud som om, at GEUS allerede har gennemgået de første stadier i en typisk benchmarking proces. GEUS opfylder kravene til ISO9001, som typisk indeholder en meget detaljeret beskrivelse af egne processer (projekter/funktioner). I forbindelse med ISO9001 proceduren er der krav om, at man skal definere nogle kvalitetsstandarder. Disse kvalitetsstandarder kan både være internt og eksternt bestemte. Hvis der findes internationale standarder følger man typisk dem, men det er ikke et "must". Hvis disse standarder er udtryk for det niveau de "bedste" har, ville den eksterne benchmarking være at finde ud af, hvordan de "bedste" når den standard. En undersøgelse af deres metoder kunne give inspiration til nytænkning, som i sidste instans gerne skulle resultere i, at GEUS befinder sig blandt de "bedste". Benchmarking er ikke en del af ISO9001, men det er en naturlig forlængelse af den proces, man allerede er i gang med. Ved intern benchmarking er det primære formål at sammenligne sig med den bedste praksis/standard i virksomheden på det pågældende område. GEUS er et typisk eksempel på en virksomhed, som handler i overensstemmelse med sine ressourcer, hvilket også er vigtigt, når man taler benchmarking.

Det efterfølgende citat er hentet fra rapporten til illustration af deres ressourcebevidsthed:

"På det strategiske niveau fastlægger direktionen GEUS' overordnede kvalitetsmålsætninger og strategien for, hvorledes disse målsætninger opnås. I forlængelse heraf er det direktionens ansvar at sikre, at kvalitetsmålsætninger står i et rimeligt forhold til instituttets ressourcer, herunder organisatorisk formåen. Ligeledes definerer direktionen kvalitetsrelaterede funktioner samt de hertil

¹⁶ Udkast til rapport om sektorforskningsinstitutioner af Kamma Langberg og Kristian Kindtler, Analyseinstitut for Forskning. Rapporten forventes at udkomme medio november 1998.

knyttede ansvar og kompetencer på de forskellige ledelsesniveauer i kvalitetsstyringen. Endelig gennemfører direktionen sammen med kvalitetsledelsen en periodisk intern auditering af, at sidstnævntes taktiske funktioner svarer til det planlagte.”

Om GEUS på nuværende tidspunkt anvender benchmarking fremgår ikke af rapporten, men det ville være nemt for dem at gennemføre benchmarking, da de har beskrivelser af egne processer og er vant til at sammenligne processerne med en standard. Det nye ville være fokus på prioritering af, hvad der er vigtigst, samt at finde ud af hvem der er “bedst” på dette område, og hvordan den pågældende virksomhed/afdeling er blevet det.

4.4 Benchmarking anvendt i uddannelsessektoren

I dette afsnit om anvendelse af benchmarking i uddannelsessektoren gives først eksempler hentet fra den udenlandske litteratur med eksempler på anvendelse i hhv. Storbritannien og USA. Efter dette beskrives det danske undervisningsministeriums kvalitets og benchmarking projekt.

Benchmarking indenfor akademiske institutioner i Storbritannien ser på forskellige kritiske processer. Disse kerneprocesser repræsenterer de overordnede styrker for hver akademisk institution, institutionernes troværdighed og evne til at konkurrere for at tiltrække de studerende til kurser og succesfuld licitation for forskningsmidler.

Universiteter i Storbritannien er rangordnet efter et mål for deres relative styrke. Rangordningen af universiteterne sker på baggrund af 7 hovedfokuseringer. Disse fokuseringer er på undervisnings-evner mere end de er på forskningsstyrker. Rangordningen opdateres engang hvert akademisk år for at reflektere konsistens og vise trends af forbedringer. Rangordningen er baseret på tilmeldingsgrad, ansættelsesniveau og afdelingsbudgetter. På denne måde vil rangordningen afspejle rygter og efterspørgsel efter universiteterne, ligesom rangordenen vil være afhængig af ressourcer, som er tilgængelig på de vigtigste områder.

Georgia Institute of Technology (GIT) ¹⁷

GIT implementerer løbende, på en meget systematisk måde, kvalitetsforbedringsprogrammer.

1. De har en **vision** om at blive det fornemste teknologiske universitet i nationen.
2. Deres **mål** er at målsætte ikke kun det, som var fortalt, men, som det var fortalt, fra perspektivet af deres kunders behov.
3. **Processerne** er etableret med fokus på:
 - Undersøgelse af betydningen af alternative evalueringer og karakterprocedurer;
 - timing af materiale introduktion; og
 - alternative leveringsmetoder.
4. **Målekriterierne** er skabt til at overvåge deres effektivitet i at møde kundernes forventninger.
5. De **udvikler** data indsamlingsstrategier.

¹⁷ "Practical Benchmarking: The complete guide" by Mohamed Zairi & Poul Leonard (1994).

5. Resultatstyring og Kriterier

5.1 Kvalitet og resultatstyring i den Offentlige sektor i Danmark

Som nævnt i indledningen har Undervisningsministeriet udarbejdet et grundlag for en fælles metode til at øge sammenligneligheden i kvalitetsvurderinger mellem uddannelsesområderne.¹⁸

Ministeriets metode bygger på to hovedtyper af mål for resultatvurderingen:

- Deciderede målsætninger for uddannelsessektorens præstationer/resultat; og
- Forudsætninger/rammevilkår

Målsætningerne er et udtryk for nogle konkrete, formulerede og alment accepterede overordnede mål for uddannelsessystemet. Målsætningerne har tilknyttet én eller flere indikatorer, som hver har ét eller flere kriterier for, om kvaliteten er god. Indikatorernes værdi bestemmes ud fra et verifikationsgrundlag. Graden af målopfyldelse kan dermed vurderes ud fra indikatorerne og de tilhørende kriterier.

Målsætning angiver hensigt og intention: Hvor vil man hen? Målsætningen kan udtrykkes ved en retning, en kurs for bevægelse eller ved noget, der kan nås. Et eksempel på en målsætning er, at ressourcerne skal udnyttes effektivt. Det overordnede mål er, at der skal sikres bedst mulig kvalitet for de ressourcer, der er til rådighed.

Indikatorer viser noget om målopfyldelsen. Nogle måltyper kan måles direkte, andre må måles indirekte. Typisk skal der flere indikatorer til at give en fyldigt indsigt i graden af målopfyldelse. En indikator for "effektiv ressourceudnyttelse" kunne være hhv. fuldførelses- og frafaldsprocenter. Disse kan måles og således vurderes det, om de ligger på et tilfredsstillende niveau.

Kriterier er angivelsen af hvilket indikatorniveau, som angiver det acceptable eller ønskelige. Kriterier er en fastsættelse af det niveau på indikatoren, som præciserer, hvornår målopfyldelsen er nået. Det er karakteristisk, at indikatoren ligger fast, mens kriteriet kan variere over tid og sted. Et kriterium for fuldførelsen kunne være mindst 75 %.

Verifikationsgrundlag. For at kunne beskrive indikatorerne, må der indsamles data/udsagn o.l., som kan benyttes til at fastlægge, hvor på skalaen kvaliteten befinder sig. Det er vigtigt, at disse data/udsagn defineres og indsamles.

Som det fremgår svarer hele denne proces til det, man gør i forbindelse med benchmarking. Med udgangspunkt i benchmarking er det nye at sammenholde disse indikatorer og kriterier med internationale standarder eller "best practices" og derigennem sikre, at niveauet afspejler den ønskede kvalitet.

Forudsætninger/rammevilkår indeholder ikke noget budskab om kvalitet i sig selv, men beskriver om grundlaget for en realisering af kvalitetsmålene er tilstede. Til forudsætninger/rammevilkår er der også tilknyttet indikatorer, der typisk er tilstandsbeskrivende fx som nøgletal, men disse kan også være af kvalitativ karakter, fx normer og værdier hos ledelse og lærere. Til disse indikatorer vil der ikke nødvendigvis altid være en egentlig kvalitetsmålsætning med et kriterium for den gode

¹⁸ "Kvalitet der kan ses": Undervisningsministeriet 1998.

kvalitet, men de giver en beskrivelse af forudsætningernes tilstand.

Fagligt og pædagogisk kvalificerede lærere og ledere er et eksempel på en forudsætning/rammevilkår. Det er ikke nødvendigvis ensbetydende med pædagogiske kvalifikationer, men det er et rammevilkår, som giver et godt udgangspunkt for kvalitet i undervisningen. Forudsætningen/rammevilkåret kan tilknytte indikatorer om grunduddannelse og pædagogiske kvalifikationer eller omfanget af efteruddannelse. Forudsætningen/rammevilkåret skal beskrive, hvordan udgangspunktet for den overordnede målsætning om en høj kvalitet ser ud.¹⁹

Det centrale i kvalitetssystematikken er definitionen af målsætninger og forudsætninger/rammevilkår.

De overordnede målsætninger er:¹⁹

- A. Systemet skal sikre uddannelses tilbud til alle
- B. Der skal være sammenhæng i uddannelsessystemet
- C. Ressourcerne skal udnyttes effektivt
- D. Højt fagligt niveau i undervisningen
- E. Befolkningen skal have et højt uddannelsesniveau.
- F. Uddannelsessystemet skal bidrage til samfundets udvikling og økonomiske vækst samt være tilstrækkeligt fleksibelt til at modsvare arbejdsmarkedets behov.
- G. Motivation for fortsat uddannelse.

De overordnede forudsætninger/rammevilkår er:¹⁹

- H. Fagligt og pædagogiske kvalificerede lærere og ledere
- I. Kvalificerede og motiverede studerende
- J. Udviklingsorientering
- K. Økonomiske og fysiske rammer
- L. Støttefunktioner.

For den enkelte afdeling eller det enkelte uddannelsessted gælder det, at man ud over disse generelle målsætninger også selv skal formulere nogle målsætninger, som er gældende for netop den pågældende institution, samt tilsvarende finde ud af hvilket udgangspunkt man har. Hvor er vi? Og hvor skal vi hen? Det er et led i benchmarking processen at stille og besvare disse spørgsmål. Kvalitet i uddannelse og undervisning drejer sig således om både processer og resultater, som i sit yderste led skal identificeres på individniveau i form af kvaliteten i de læreprocesser, gennem hvilke den enkelte i samspil med sine omgivelser opbygger relevante faglige og personlige kvalifikationer.²⁰

Kvaliteten i en læreproces er afhængig af en lang række forhold, som i større eller mindre grad kan måles præcist og objektivt. Forhold som de givne rammer for en uddannelse i form af en plan for undervisningen, undervisningsmaterialet og de fysiske rammer er af forholdsvis objektiv karakter, mens kvaliteten i den indsats, der ydes af de "lærende" selv, af elever og studerende - som både proces og resultat afhænger meget af - er vanskelig at vurdere objektivt. Selvom det er tegn på kvalitet, at lærerne, uddannelsesinstitutionerne etc. befordrer og motiverer denne indsats, kan man ikke blot betragte elever og studerende som "brugere" og lærerne som leverandører af service-

¹⁹ "Kvalitet der kan ses": Undervisningsministeriet 1998.

ydelse. Skolen og uddannelsesinstitutionerne skal ikke "blot levere den efterspurgte vare" og skabe "kundetilfredshed" her og nu. Kvaliteten kan derfor ikke blot måles via brugerundersøgelser, lige så lidt som det kan ske alene gennem optælling af antal beståede eksaminer pr. år. Brugerundersøgelser og opgørelser over antal færdiguddannede kan være med til at give et billede af kvaliteten, men disse mål skal indgå i kombination med andre datatyper.²⁰

I princippet er det nødvendigt at kombinere oplysninger, der vedrører både forudsætninger, processer og resultater på individ-, gruppe-, institutions-, uddannelses- og systemniveau, hvilket i praksis på grund af kompleksitet og ressourceforbrug naturligvis kun kan lade sig gøre i stærkt begrænset omfang. Kvalitetsbeskrivelser må derfor bygge på et begrænset antal indikatorer udvalgt således, at der er balance mellem forskellige kategorier. Det er i denne sammenhæng vigtigt at pege på den problemstilling, der ligger i erfaringen af, at "man får det, man måler". Denne effekt opstår allerede på grund af det opmærksomhedsskabende i selve målingen. Fokuseringen udgør i sig selv et incitament, som på sigt kan være svært at holde adskilt fra den normale målstyring.²⁰

På grund af problemstillingen med, at "man får, hvad man måler", er prioriteringen og udvælgelsen af målsætninger så afgørende en del af benchmarking, fordi det herigennem sikres, at man sætter mest fokus på det, som har størst betydning, og dermed ikke bruger en masse kræfter på et problem, som ikke har den store betydning for at nå til det mål, man har sat sig.

Et overordnet kvalitetssystem indebærer en risiko for, at de mere bløde og integrerede kvalitetsovervågningsmetoder, der i dag findes i uddannelsessystemet, fx i institutionsmiljøerne og i forbindelse med små og pr. tradition etablerede uddannelser, overses og evt. erstattes med mere formaliserede systemer. Det negative herved hænger sammen med, at erstatning af intuitive metoder med mere formaliserede og bureaukratiske metoder vil kunne være mindre effektive set i forhold til et udviklingsperspektiv, fordi det kan skubbe engagementet til side, og u hensigtsmæssigt, fordi det kan være mere omkostningskrævende og fortrænge værdifulde aktiviteter. Begrænsningen i at basere sig på sådanne mere intuitive systemer ligger naturligvis i, at der kan være mangel på overordnet målsætning, ekstern synlighed og udgangspunkter for intervention, hvilket betyder, at denne form ikke er så hensigtsmæssig, når det drejer sig om tilvejebringelse af dokumentation for større dynamik og navnlig ikke er hensigtsmæssig, når det drejer sig om vurderinger af helhedsprægede forandringer. "Kvalitet der kan ses" - projektet skal tjene som et supplerende kvalitetsevalueringssystem, der kan give inspiration til, at udviklingen i uddannelsessektoren følger en ensartet målsætning og tankegang.²⁰

Med de utallige tilgange, der kan vælges til vurdering af kvalitet, bliver det åbenbart, at kvalitetsbegrebet i sig selv er flertydigt - ikke mindst fordi uddannelse både drejer sig om udmøntning af politik, beslutninger om aktiviteter, service og opgaveløsning i henhold til lovgivning, regler og aftaler, og om interessenteres rettigheder, muligheder og handlinger. Produkt og proces er ofte vanskelige at adskille, og selve undervisningsforløbene er ligesom interessenterne forskellige og ikke mulige at standardisere. Da undervisning er individorienteret, er situationstilpasning nødvendig, samtidig med at lærerindsatsen må tage hensyn til praktiske forhold og politiske prioriteringer.²¹

²⁰ "Kvalitet der kan ses": Undervisningsministeriet 1998.

Det er vigtigt i vurderinger af kvalitet af undervisning og uddannelse at medtænke og integrere de processer eller mekanismer, der kan benyttes i de enkelte uddannelser til at skabe, sikre og videreudvikle kvaliteten. En sådan gennemgående "kvalitetstankegang" er en forudsætning for en stadig udvikling af kvalitet.

5.2. Kriterier for de videregående uddannelser

Målsætninger:²¹

A. Flest muligt kvalificerede ansøgere optages på den ønskede uddannelse.

- Antal optagne på ønsket uddannelse i forhold til antal kvalificerede ansøgere til uddannelsen

B. Ressourcerne skal udnyttes effektivt

- Fuldførelses- og frafaldsprocenter
- Udgifter per studenterårsværk
- Studenterårsværk per lærerårsværk
- Antal afholdte timer i % af de planlagte
- Studietider, fordelingen af fuldførte studerende på samlet forløbstid i forhold til normeret studietid.

C. Højt fagligt niveau i undervisningen og forskningen

- Evaluering af undervisningen
- Internationale sammenligninger af det faglige niveau
- Resultat af eksamen og prøver, herunder censorvurderinger
- Fremmødeprocent
- Uddannelserne har definerede og prioriterede kerneområder/Uddannelsens faglige profil
- Antal af timer, der varetages af VIP-forskere
- Klar og præcis organisering i studieprogrammer
- Forskermiljøets størrelse
- Videnskabelig produktion
- Formidling af resultater
- Evalueringer og Peer reviews
- Omfang og betydning af eksterne forskningsbevillinger
- Fordeling mellem frie forskningsmidler og programmidler

D. De færdiguddannede skal have gode faglige og personlige kvalifikationer

- Resultat af prøver
- Evalueringer af uddannelser
- Andel af studenter, der deltager i demokratiske processer
- Grad af studerendes samarbejdsevner og omstillingsparathed selvtillid/selvopfattelse

²¹ "Kvalitet der kan ses": Undervisningsministeriet 1998.

E. Uddannelsen fører til relevant beskæftigelse/videreuddannelse og modsvarer arbejdsmarkeds behov

- Beskæftigelsesstatistik
- Aftagerundersøgelser

F. Systemet skal give motivation for fortsat uddannelse

- Efteruddannelsesprocent

Forudsætninger/rammevilkår

G. Fagligt og pædagogisk kvalificerede lærere og ledere

- Faglige og pædagogiske krav og andelen af lærere, der opfylder disse
- Grunduddannelse og pædagogiske kvalifikationer (andre uddannelser)
- Studerendes vurderinger af undervisere
- Efteruddannelse

H. Kvalificerede og motiverede studerende

- Fuldførelsesprocent
- De studerendes arbejdsindsats
- Lærervurderinger af de studerendes forudsætninger

I. Udviklingsorientering

- Handlingsplaner, der evalueres og opdateres, for institutions- og personaleudvikling (rekruttering, fastholdelse, udvikling og afvikling)
- Ressourceforbrug til forskning og udviklingsarbejde
- Procedurer for kvalitetssikring og -udviklende mekanismer
- Omfang af internationalisering

J. Økonomiske og fysiske rammer

- Ressourceforbrug til uddannelse
- Funktionsfordelt ressourceforbrug
- Bygningsareal per studenterårsværk
- Tidssvarende biblioteksfaciliteter
- Antal studerende per tidssvarende computer
- Ressourceforbrug til renovering og nybygning
- Tilgængelighed for personer med handicap

K. Støttefunktioner

- Ressourcer til studie- og erhvervsvejledning
- Andel af studerende med uddannelsesstøtte

I kriterierne for de videregående uddannelser indgår enkelte kriterier, som relaterer sig til forskningsmiljøet. Det vigtige i forbindelse med en benchmarking-tilgang er at prioritere resultatstyringskriterier samt at beskrive, hvor man nu er f.eks. ved videnskabsproduktion eller fuldførelsesprocent, og hvilket niveau man ønsker for de pågældende indikatorer? - Hvilket niveau har de, som

man sammenligner sig med? Og hvilket niveau har de, som man burde sammenligne sig/sin institution med? - Dernæst hvordan man når dette niveau udfra de forudsætninger og rammer, som man har? Det væsentligste i processen er ikke om institutionen på det enkelte kriterium når det fastsætte niveau, det væsentligste er derimod hele forståelsen af vigtigheden af kvaliteten samt det totale billede af, hvor man står, og hvor man vil/skal hen.

Som en udløber af benchmarking debatten inden for uddannelsesområdet, er der opstået et særligt begreb, med betegnelsen "pædagogisk benchmarking", som adskiller sig fra "økonomisk benchmarking" ved at lægge vægt på en sammenligning af de bløde værdier i eksempelvis skolens kerneydelse: undervisning. I dette eksempel med undervisning er det lærernes forløb i undervisningen, som skal sammenlignes. Essensen er her at fokusere på den undervisning, der fungerer godt i stedet for at fordybe sig i den undervisning, der virker dårligt. Det handler altså om at synliggøre de gode forløb, så andre kan lære af dem.

Pædagogisk benchmarking kan som metode blive et væsentligt værktøj i virkeliggørelsen af den lærende organisation. Udgangspunktet er en analyse af den nuværende situation og et ønske om at realisere en fælles vision. Ved at analysere og reflektere over egen praksis og sammenligne med hvad andre gør, skabes der mulighed for at handle på nye måder. Pædagogisk benchmarking kommer derved til at synliggøre koblingen mellem kvalitetsudviklingen, evalueringen, lærerprofessionalismen og ledelsen. Hvis man skal overføre tankegangen fra pædagogisk benchmarking med fokus på de bløde værdier til universiteterne og andre forskningsinstitutioner, må det være et spørgsmål om at fremhæve den gode forskning ved at synliggøre metoder mv., så andre kan lære af dette.²²

6. Overvejelser over kriterier til kontrakter for universiteter

Målsætning i flg. Forskningsministeren: De danske institutioners kvalitet og gennemslagskraft nationalt såvel som internationalt skal øges.

Målsætningen er formuleret dynamisk forskellig fra stationær. Der er tillige i ministerens udmeldinger et ønske om en anspore til omstillingsparathed, men dette aspekt må opfattes som et middel til ovenstående målsætning, ikke som et mål i sig selv. Udtalelser fra ministerens side om stærke ledere må ligeledes ses som et led i ovenstående målsætning.

Kontrakter indgås mellem parter dvs. institutionerne og Ministeriet, hvilket vil sige, at parterne skal være enige om, hvad der skal indgå i aftalen; men uanset dette kan udformningen af kontrakter ses som et udslag af forskningspolitik.

Forskningspolitik inddrager i disse år i stigende grad evalueringer. Nedenstående er givet en form for kortfattet oversigt over, i hvilket omfang erfaringer fra evaluering af forskning kan indgå i overvejelserne om udformningen af kontrakterne. Idet det ud fra en antagelse om "common performance orientation" må forventes, at de grundlæggende kriterier skal være gennemgående for alle typer af kontrakter med uddannelses-/forskningsinstitutioner, skal det allerede indledningsvis anføres, at kun en del kriterier kan være generelle for alle institutioner uanset faglig sammensætning, men en væsentlig del vil være fagspecifikke og institutionsspecifikke. Det gennemgående træk i udtalelser fra europæiske administratorer/forskere med erfaring i kontrakter

²² "Nyhedsmagasinet Danske Kommuner: Lærerne skal dele gode erfaringer" af Mai-Britt Herløv Petersen.

fra europæiske administratorer/forskere med erfaring i kontrakter etc. inden for offentlig forskning er, at der advares mod at bruge de sædvanlige detaljerede evaluerings-kriterier for håndfast. Det afgørende spørgsmål bliver her, hvor langt ned i institutionens opbygning/sammensætning man fra Ministeriets side agter at gå ved udformningen af kontrakterne, idet institutionernes interne fleksibilitet vel skal opretholdes.

Baggrundskarakteristik af institutionen, der er genstand for kontrakten, kan/bør vel føre til specielle kontrakter for de enkelte institutioner eller alternativt, at kontrakterne indeholder elementer, der tager højde for institutionernes bredde, sammensætning og fysiske rammer, herunder den geografiske placering. Nedenstående angår derfor alene det generelle, som kan indgå i overvejelserne for alle.

Institutionens karakter dvs. dens størrelse forstået ved forskerstaben og antallet af forskningsfelter, som institutionen dækker, institutionens bredde defineret ved antal fagområder, som institutionen dækker, antal uddannelser, antal studenter/lærere, der "belaster" institutionen. Studenterbelastningen er vel pr. definition et spørgsmål for Undervisningsministeriet, men belastningen for den enkelte ansatte varierer fra institution til institution, hvilket giver store forskelle i reel forskningstid. Sammenhængen mellem forskning og undervisning gør det vanskeligt at adskille ressourcetildelingen, såfremt man ikke snævert sigter på at indgå kontrakt om de ressourcer, der skal anvendes til forskning. Alternativt kan man forestille sig, at der indgås kontrakter mellem institutionerne og de to ministerier på samme tid. Sidstnævnte situation vil matche den i Danmark forekommende traditionelle sammenkædning mellem universiteternes og universitetslærernes forskning og deres undervisning; en sammenligning som bl.a. afspejles i de ovenfor nævnte kriterier for de videregående uddannelser (p. 20).

Forskningens omfang, niveau/ kvalitet og relevans/konsekvens for omgivelserne er centrale elementer. Der er dog europæisk enighed om, at der er store vanskeligheder forbundet med at definere relevansen af forskellige forskningsområder, nogle mere end af andre. Endvidere skal inddrages forskningsformidling som led mellem institutionen og dens omverden. På alle disse felter kan der jævnfør målsætningen inddrages overvejelser mht., om der kan ske ændringer i form af forbedringer!

6.1 Forskningsprojekter

Forskningens omfang kan måles ved antallet af igangværende projekter, og her kan der sættes en målsætning om at tilstræbe deltagelse i et øget omfang i international forskning. Det skal her bemærkes, at forskning i samarbejde med det lokale/regionale eller nationale erhvervsliv ligeledes kan være en målsætning, der kan opstilles for den enkelte institution.

Traditionelt kan der måles på:

- Hvor mange projekter ledes/koordineres fra institutionen?
- Hvor mange projekter deltager institutionen i?
- Hvor mange af disse er internationale? Erhvervsrelaterede eller anden målsætning?
- Hvor mange bevillinger har institutionen fået hjem?
- Ekstern finansiering kan opfattes som et succeskriterium i sig selv, men er øget ekstern finansiering i sig selv et mål? Der bruges megen tid på ansøgninger herom.

6.2 Produktivitet

- Produktion på de enkelte fag i form af ph.d.'er kan aggregeret bruges til samlet mål for institutions output (hvorimod produktion af kandidater er mere et spørgsmål til Undervisningsministeriet).

Hvis man alene sigter på en øgning af kvantitative mål, kan der opstilles målsætninger mht., at antallet af ph.d.'er skal søges øget. En anden mere kvalitativ diskussion er, i hvilket omfang man kan inddrage overvejelser om, hvor de færdiguddannede går hen? I hvilket omfang går de til erhvervslivet, til offentlig administration, til internationale organisationer? Det skal bemærkes, at der inden for OECD i disse år er særlig opmærksomhed rettet omkring "flow of human capital", defineret som hvor de uddannede går hen, hvilken mobilitet er der blandt akademikere? Til overvejelser herom kunne der sondres mellem kandidater fra forskellige uddannelses-institutioner. Projektet er på europæisk plan endnu ikke gennemført, men de første papirer, som er oplæg til projektet, fremlægges ved et OECD-NESTI møde juni 1998.

Et stort afsæt af kandidater til væsentlige poster/væsentlige sider af det danske og internationale samfund kan inddrages i overvejelser om delmål, men der kan være et problem med tidsperspektivet i afsætningen.

- Produktion i form af færdiggjorte forskningsprojekter, publikationer, artikler etc. kan betragtes som en anden del af produktionsmålet, men det kan også ses som en del af formidling, jf. nedenfor. Den viden, der som følge af forskningsaktivitet kumuleres, men som ikke umiddelbart giver sig udtryk i publikationer, kan derimod ikke måles på kort sigt.

6.3 Formidling

- Antal bøger, monografier, kapitler til readers, artikler, referencer eller ej.
- Udgivelser på internationale forlag versus udgivelser på eget forlag.

Metodisk kan bibliometriske mål anvendes, men der må advares mod brugen af forenkede mål baseret på de store oversigtsværkers opgørelser bl.a. på grund af kulturforskelle, fagforskelle, traditioner etc. ligesom problemet med selvreferencer skal bemærkes.

- Anden form for formidling end bøger, f.eks. foredrag, avisindlæg, workshopbidrag etc.

Bibliometriske mål kan i princippet anvendes, men kan ikke umiddelbart anvendes, da oplysninger herom ikke eksisterer i databaser p.t.

6.4 Samarbejde med omverdenen

- Kontakter med andre danske institutioner i form af samarbejdsaftaler bl.a. om studenterudveksling.
- Kontakter med andre universiteter i form af fælles projekter.
- Kontakter med udenlandske universiteter om projekter/uddannelsesprogrammer.
- Omfang af EU-støttede aktiviteter.

6.5 Rådgivning på forskellige niveauer

- Hvor mange institutter ved den enkelte institution er involveret i ekstern rådgivning til hvem (Ministerier, internationale organer, nationale instanser eller andre)? Hvor ofte?
- Deltagelse i offentlig debat er en anden form for dialog med omgivelserne.

6.6 Deltagelse i bedømmelsesudvalg, internt/eksternt

- Typer af bedømmelsesudvalg, grad af international orientering.
- Referering af artikler.

6.7 Tidsskriftsredaktionsarbejde, nationalt/internationalt

Der skal her skelnes mellem arbejde for tidsskrifter m./u. referering.

6.8 Afholdelse af konferencer/deltagelsesomfang

- Konferencer som institutionen leverer.
- Gæsteforelæsninger, som en institution afholder eller leverer foredragsholdere til.
- Deltagelse i konferencer, indenlandsk/udenlandsk/discussant/moderator.
- Paper-/poster-præsentation ved konferencer.

6.9 Effektivitet (produktion pr. given enhed)

- Effektivitet, antal årsværk pr. publikation/antal publikationer pr. årsværk.
- Effektivitet, antal kandidater pr. årsværk.
- Effektivitet, antal ph.d.'er/doktorer pr. forskerårsværk.
- Publiceringer pr. forsker.

Det skal bemærkes, at der vil være store forskelle fagmiljøer imellem. Sammenvejede mål kan overvejes.

6.10 Innovationsevne

- Patenter (kun relevant for visse områder); det skal bemærkes, at danske forskere har ringe tradition for patentering.
- Omstillingsevne, tilpasning til politiske signaler/forskningsprogrammer.

I et væsentligt omfang kan kriterier fra evaluering bruges som udgangspunkt for overvejelser om indholdet af kontrakter ud fra målsætningen om, at det der indgås kontrakt om, vel er det, der skal evalueres efter.

Reaktionen fra miljøerne vil så vidt vides være, at der sker en del, som måske nok kan evalueres, men som ikke kan kvantificeres. Men selv ved forhold, der kan kvantificeres, vil der stadig være forhold, som er vanskelige at evaluere reliabelt.

Europæiske erfaringer

Anvendelse af erfaringer fra udlandet er inddraget i ovenstående, som lægger vægten på målelige kriterier. Derudover skal bemærkes, at ved EU DG-XII's netværk-konference i London d. 8. juli 1998 gav Stefan Kuhlmann en oversigt over "advanced methods of socio-economic evaluation". Ved denne gennemgang beskrev han udviklingen i evalueringer på europæisk plan som gående fra kvantitativ til kvalitativ, fra ex post til ex ante, dvs. fra faktiske præstationer til forventede. Denne

bevægelse passer i sig selv med en overgang til et ønske om at lave kontrakter.

Derudover diskuteres generelt overvejelser mht. om "stærke miljøer", defineret som dem, der kan score højt på ovenstående sæt af kriterier, skal styrkes yderligere. Det diskuteres, om det er hensigtsmæssigt alene at udtrykke tillid til "stærke miljøer", og en form for opgør med "svage". Det fremføres, at man måske forskningspolitisk skulle satse på at styrke nationalt svage områder ved tilførelse af eksternt ekspertise evt. i form af forskningsprofessorater til opbygning af miljøer, hvilket bl.a. er sket i Irland.

En karakteristik af stærke/svage miljøer kan indgå i en form for selvpræsentationen, men denne dimension er vanskelig at få ind i en egentlig måling, der kan gøres til genstand for en kontrakt, med mindre den måles ved omfanget af internationale kontakter/projekter eller ved en sammenvejning af en række af de ovennævnte mål.

I Holland inddrages et væsentligt omfang af selvevaluering fra universiteternes side i de evalueringer, som hollandske universiteter gennemfører. Selvevaluering indgår også i det europæiske rektorkollegiums evalueringsprogram.

Overvejelser om validitets- og reliabilitetsproblemer fremgår bl.a. af OECD-rapporten *The Evaluation of Scientific Research: Selected Experiences*, OECD/GD (97)194. Til vedlagte er der trukket på denne rapport, og der er derudover trukket på erfaring/informationer fra NESTI samt EU/DG-XII-evalueringsudvalg.

Der advares i OECD-rapporten mod, at evalueringskriterier bliver styrende, ligesom der gøres opmærksom på, at evalueringer/opstillinger af målsætninger bør ske i fuld bevidsthed om, at forskningen har en række afledte funktioner på en given institution og i det omgivende samfund. Netop disse advarsler fører til, at der skal indgå andre kriterier end de sædvanlige, der indgår i forskningsevalueringer.

Jf. OECD-rapporten er der to basale komplementære måder at evaluere på: anvendelse af kvantitative indikatorer eller anvendelse af peer reviews. I Frankrig bruges i vid udstrækning peer reviews i sammenhæng med universiteternes egne udmeldinger med planer for institutter/laboratorier.

Der foregår for tiden en debat om det hensigtsmæssige i at forlade sig på peer review, idet der sker en selektion af bedømmere ved valget af peers, og dermed en implicit og ofte ubevidst selektion af bedømmelseskriterier, såfremt disse ikke er fastlagt på forhånd.

Materiale fra Australien hentet via nettet fortæller, at der til ratings er anvendt 89 forskellige ratingdimensioner. Kun en mindre del af disse er relevante for overvejelser om kontrakter, idet en stor del vedrører materielle forhold, der ikke anses for relevante.

Materiale vedr. britisk praksis fortæller, at man i UK siden 1992 har haft et kontor, HEFCE, som fordelte midler til uddannelse og forskning på 136 engelske universiteter og colleges. Udgangspunktet for fordelingen har været "research assessments". I materialet fra UK problematiseres fremgangsmåden med anvendelse af de mange kvantitative mål som basis for tildeling af midler,

og der anbefales i stedet øget brug af peer reviews (RAE 2/97 p. 4). Advarslerne lyder derefter, at der skal gå mindst fire år mellem evalueringerne, for ellers bliver de for styrende, for dominerende.

Erfaringer med det engelske system, der bygger på en kombination af peer reviews med kvantitative indikatorer, er p.t. under evaluering. En samlet rapport vedr. de britiske evalueringer vil foreligge senere på året.

Sammenfatning

Det er fortsat et spørgsmål om, i hvilket omfang de ovenfor fremlagte kriterier kan anvendes som udgangspunkt for de kontrakter, der skal indgås med danske universiteter. En af Analyseinstitut for Forskning organiseret konference i september 1998 stillede bl.a. dette spørgsmål til et europæisk panel. Svarene var ikke entydigt positive, - ej heller entydigt negative. De ovenfor beskrevne forskelle i europæisk praksis afspejledes i holdningerne. Sammenfattende var anbefalingerne, at de udvalgte kriterier i givet fald skulle være relevante og accepterede af de involverede parter, for kun derigennem kan der opnås validitet og tilslutning til arbejdet med at styrke danske universiteter.

7. Metodeovervejelser

I forbindelse med arbejdet med at koble benchmarkingmodeller til udviklingskontrakter er der flere forhold, der skal overvejes. Blandt disse er overvejelser mht. valg af enhed/genstand for måling/genstand for kontrakt. Hvilken enhed skal evalueres/skal der indgås kontrakt med? Universitetet eller dele af dette? I følge forskningsministerens udspil, er der tale om kontrakter med universiteterne, der skal have stærke ledere, der kan klare styringen og kontrollen nedefter i systemet. Enkeltpersoner, institutter, fakulteter kan være enheder, der måles på, men disse er ikke genstand for kontrakt i den for tiden diskuterende version.

Derudover er der overvejelser over, hvordan man finder "den bedste" at sammenligne sig med. Grundlæggende drejer dette problem sig om præstationsmåling. Andre ord for det samme er effi- cients- eller effektmåling.

Den såkaldte DEA-metode er grundlæggende en metode til tildeling af effi- cientsgrader til alle de organisationer/institutioner, der indgår i analysen, ud fra en sammenligning mellem disse²³. Denne metode giver mulighed for rangordning af organisationer/institutioner efter effi- cientsgraden. Undersøgelsen viser ligeledes om den enkelte organisation er effi- cient i forhold til referencegruppen eller med i forhold til "best performance"/de dygtigste organisationer/institutioner.

DEA-metoden kan altså anvendes som evalueringsværktøj i forbindelse med benchmarking. Eva- lueringsmålene kan have både kvalitative og kvantitative karakter, og der er ligeledes tale om et værktøj, som ikke kun anvender et mål, men hvor mange parametre kan indgå i analysen og den "bedste" findes på baggrund af en "optimal" vægning af disse.

Effektivitetsanalyserne kan have til formål at vurdere i hvor høj grad, ressourceanvendelsen fører til den ønskede målsætning. Finansministeriet har i 1992 anvendt DEA-metoden til at sammenligne

²³ "Produktivitetsanalyser", Flemming Christensen, Peter Fristrup og Jens Leth Hougaard, Jurist- og Øko- nomforbundets Forlag (1991).

produktiviteten på danske seminarier.²⁴ Denne undersøgelse er begrænset til udelukkende at omfatte økonomiske variabler som driftsudgifter; der indgår ingen vurdering af kvalitet, omdømme eller tilfredshed med seminaret som helhed. Det gode ved undersøgelsen er, at den åbner op for diskussionen på de enkelte seminarier om at sammenligne sig med dem, der klarer sig bedst på et specifikt område, her "minimering" af driftsudgifter. Finansministeriet anvender undersøgelsen som indikator for besparelspotentiale hos de virksomheder/institutioner, som ikke er efficient eller med danske ord, som endnu ikke ligger på linje med de bedste.

8. Problemer med benchmarking i forhold til udviklingskontrakter

En forkert anvendelse af benchmarking kan helt klart være et problem. En måde at anvende benchmarking forkert er, hvis der anvendes en masse ressourcer på emner, som slet ikke er væsentlige for virksomheden eller institutionen.

Et andet problem er, at man får, hvad man måler, hvilket uhensigtsmæssigt kan trænge andre vigtige elementer i baggrunden. Hvis vi igen tager universiteterne som eksempel har disse 3 hovedopgaver: Uddannelsesopgaven, forskningsopgaven og formidlingsopgaven. Hvis der kun er fokus på evaluering af undervisning vil dette få negative konsekvenser for forskningsopgaven, som så vil blive nedprioriteret, fordi man ikke opnår flere midler til universitetet eller opnår bedre "resultat i evalueringen" ved at have god forskning.

Hvis der derimod alene er fokus på forskning og formidling, så kan dette medføre en nedprioritering af undervisningsopgaven.

Det kan også være problematisk at sammenligne sig med andre, før man kender "sig selv", da situationen for den enkelte institution måske viser sig at være anderledes end umiddelbart antaget, når man kommer til bunds med granskning af "sin egen institution".

Evalueringer giver altid anledning til omfattende diskussioner. Er det egentlig kvalitet, der måles? Og er det de vigtigste kvaliteter, der er målt? Er der eksempelvis taget tilstrækkeligt hensyn til, at universiteterne er forskellige? At placering og udvikling i regional og faglig kontekst har været og fortsat er forskellig? Har evalueringspanelernes medlemmer, såfremt sådanne anvendes, de rette kvalifikationer? Kan man i det hele taget måle kvalitet? De mulige indvendinger er utallige.

For måling og evaluering generelt gælder det generelt, at det er vanskeligere at måle kvalitative elementer som eksempelvis tilfredshed og kvalitet, end det er at måle kvalitative elementer, men omvendt vil det i mange specielt offentlige virksomheder give det bedste helhedsbillede, hvis analysen indeholder andet end kvantitative målinger.

Forskelligheder er umiddelbart vigtige at lægge mærke til, når der foretages sammenligninger med forskelligartede enheder, så man ikke sammenligner størrelser, der pr. natur eller pr. kultur ikke er sammenlignelige, populært sagt skal man lade være med at sammenligne pærer med appelsiner. Problemer er specielt udbredt inden for sammenligning af publikationer de enkelte universitetsafdelinger imellem. Og dermed er der problemer ved sammenligninger universiteter imellem, såfremt disse ikke har den samme faglige sammensætning. Der er konkret ikke de samme traditioner for at publicere i tidsskrifter på de humanistiske fakulteter, som der er på de samfundsvidenskabelige fakulteter. Og der er andre forskelle, når man sammenligner med

²⁴ "Produktivitetsanalyse af Seminarieområdet", Finansministeriet og Undervisningsministeriet (marts 1992).

fakulteter. Og der er andre forskelle, når man sammenligner med traditionen for publicering inden for medicin.

Et andet problem ved evalueringer er, at de er tidskrævende. Den tid, der anvendes i forbindelse med evalueringsprocessen, skal tages fra den tid, der kan anvendes til virksomheden/institutionens andre opgaver.

Et meget centralt problem er den modstand, man mange steder finder til måling og evaluering af forskning og videnskabeligt arbejde generelt. En generel holdning man er oppe imod ved indførelsen af udviklingskontrakter og deraf følgende evalueringer eller andre tilsvarende målinger er opfattelsen af, at "målinger/evaluering vil begrænse den enkeltes frihed".

Endnu en hurdle er, at mange har dårlige erfaringer med evaluering, og det er altid meget sværere for en organisation at motivere medlemmerne til at deltage og lære fra en evalueringsproces, når deres erfaringer er negative. Der kan være flere forklaringer på disse dårlige erfaringer; et element er, at de tidligere evalueringer har været for simple og dermed ikke opfattes som retvisende. Ved at bygge på valide multidimensionelle evalueringer kan man måske fremover vende denne del af de negative erfaringer til positive erfaringer og dermed opnå et bedre resultat. Et andet element er, at "evalueringerne har været unfair". Ved at gennemføre nogle evalueringer, som miljøerne på forhånd har accepteret nødvendigheden af at gennemføre, kan forskerne måske ændre opfattelser i retning af, at "evalueringerne er fair", nyttige og måske endda nødvendige.

Det sidste centrale element, som skal behandles i dette notat, er "belønning/straf" problematikken. Hvis der ikke er knyttet nogle konsekvenser (økonomisk belønning eller negative sanktioner) til det resultat, man opnår, vil dette i sig selv virke som et manglende incitament til at gøre en øget indsats. For at sikre forståelse for evalueringens formål og for at motivere miljøerne til at gøre sig endnu bedre end de allerede er, må det anses for nødvendigt at knytte en vis del af ressourcerne til resultaterne af evalueringen. Erfaringer fra udlandet, eksempelvis fra Irland, viser dog, at negative evalueringer ikke nødvendigvis fører til negative sanktioner, idet målsætningen forskningspolitisk i sidste ende er politikernes. Negative evalueringer kan ud fra en politisk prioritering føre til en øget investering, såfremt den pågældende institution ønskes optimeret. Det afgørende for de politiske overvejelser i en sådan situation må antagelig være, om der ved institutionen er tilstrækkeligt at bygge videre på. Har institutionen vist vilje til at indgå i overvejelser om, hvordan institutionen kan komme videre og blive blandt de bedste af sin art? På denne måde kan man argumentere for, at ringen i dette notat er sluttet. Benchmarking, uanset at det ikke er en epokegørende ny metode, pålægger den enkelte institution at overveje sin egen situation og overveje med hvem, det er rimeligt at sammenligne institutionen. Overvejelserne kan gøres til genstand for interne drøftelser med deraf følgende opmærksomhed og eventuel selverkendelse. Sammenligningerne med andre kan gøre institutionen/universitetet i stand til at definere sine mål, som kan udtrykkes i en udviklingskontrakt, såfremt målene accepteres som relevante og gyldige mål for institutionen af kontraktpartneren, der i dette tilfælde er Forskningsministeriet eller evt. både Forskningsministeriet og Undervisningsministeriet. Elementerne i den indgåede kontrakt skal gøres til genstand for almindelig åben debat, for uden denne offentlighed om målsætningen kan det ikke forventes, at institutionerne kan leve op til målsætningen. Stærke ledere har også brug for, at andre, og her ikke mindst de ved institutionen ansatte, har kendskab til målene. Ideelt set skal målene opstilles som led i en intern dialog, førend målene forhandles med den eksterne kontraktpartner in casu ministeriet.

9. Sammenfatning vedr. benchmarking i forhold til udviklingskontrakter

Benchmarking er det nye ord for at sammenligne sig med andre med det formål selv at blive bedre. Det må konkluderes, at benchmarking er et godt redskab til motivering og forbedring af specifikke funktioner samt en hjælp til nytænkning.

Der findes mange benchmarking modeller, men fælles for dem alle er en planlægnings fase, en analysefase (som indeholder forståelse og evaluering af egne processer samt sammenligning med andre) og endelig en implementeringsfase.

Bevidstheden om, at man får, hvad man måler, er vigtig, så man ikke anvender benchmarking forkert og bruger en masse ressourcer på at forbedre noget, som ikke er væsentligt.

DEA-metoden er et værktøj, som er velegnet for offentlige virksomheder til at finde den eller de "bedste" på et givet område. Det er altså en måde at danne en standard, som man kan sammenligne sig med.

Benchmarking er et redskab til hjælp til mål og resultatstyring, men det er vigtigt at fokusere på alle virksomheden/institutionens opgaver, og det er vigtigt, at en evalueringen, som efterfølger en eventuel udviklingskontrakt, er multidimensionel og fair over for det grundlag, kontrakten er indgået på. Det er derfor vigtigt, at udviklingskontrakter og deraf følgende evalueringskriterier opstilles i samarbejde med dem, der senere skal evalueres efter disse kriterier, for dermed samtidig at reducere den modstand, som allerede findes til måling og evaluering af forskningsinstitutioner.

Evalueringer kan bruges til at sikre og fremme kvalitet, men forventning om evaluering kan også være årsag til dyrkelse af evalueringskriterier, som ikke nødvendigvis får det bedste frem i institutionerne. Der er i denne sammenhæng adskillige advarsler fra europæisk hold om, at evalueringer eller såkaldte assessments ikke misbruges administrativt som automatisk bevillingsgrundlag eller som grundlag for automatisk reallokering af bevillinger.

10. Forskningsministeriets udspil til målsætninger

Forskningsministeriet vil i følge meddelelse til universiteterne i forbindelse med udviklingskontrakterne med universiteterne lægge vægt på følgende målsætninger:

- Styrket samarbejde med det omliggende samfund
- Bedre og hurtigere udnyttelse af forskningsresultater
- Styrkelse af internationalisering
- Fornuftig arbejdsdeling i det samlede forsknings- og uddannelsessystem
- Koncentration og fokusering på styrkepositioner
- Effektiv ledelse og administration
- Klar løn- og personalepolitik med henblik på rekruttering og fastholdelse af de bedste forskere og lærere
- Rationel bygningsforvaltning med henblik på bygningsmassens anvendelse og vedligeholdelse

Kun en del af disse målsætninger har tidligere været gjort til genstand for måling.

Benchmarking-metoden kan anvendes i forbindelse med målsætninger, som ovenfor angivet, idet en gennemførelse af ovenstående målsætninger kræver overvejelser internt i den enkelte instituti-

on, om hvordan institutionen er placeret på den enkelte dimension sammenholdt med andre, eller sagt på anden måde: hvad er "best practise" inden for de enkelte mål. Hvor ligger man mindre godt end andre tilsvarende institutioner? Hvem skal man sammenligne sig med? Hvor kan man hente inspiration? Hvad er det muligt at forbedre? Sådanne overvejelser er typiske overvejelser, der kan foregå inden for benchmarking traditionen.

Den anden del af processen er derimod den del, der jf. ministeriets udmeldinger til institutionerne relaterer sig til, hvordan disse foranstaltninger skal måles, og derefter til spørgsmålet om hvilke målingsinstrumenter, der skal etableres for at måle effekten af udviklingskontrakterne? Kan målinger foregå på anden vis end ved evalueringer af mere kvalitativ karakter?

Afsluttende anbefaling: at der gennemføres en undersøgelse med inddragelse af universiteterne af, hvordan disse målsætninger kan udmøntes i målbare kriterier.

Litteraturliste

"Practical Benchmarking: The complete guide", Mohamed Zairi & Poul Leonard (1994)

"Benchmarking for continuous improvement in the public sector", John Bullivant

"Benchmarking customer service", Glen Peters (1994)

"Benchmarking", Peter Scott, Management Accounting, vol. 74., nr. 7., p. 50

"Benchmarking - performance gap - real or a problem of definition?", Ian Malcolm, Management Accounting, 1997, vol. 75, nr. 4, p. 38-39

"The Value of Benchmarking management reporting", Ian Malcolm, Management Accounting, 1996, vol. 74, nr. 6, p. 18-20

"Benchmarking for best value in local authorities", Mary Bowerman and Amanda Stephens, Management Accounting, 1997, vol. 75, nr.10, p. 76-77

"Benchmarking i praksis, Model og eksempler fra danske virksomheder", Dansk Industri, 1997

"Kvalitet der kan ses": Undervisningsministeriet, 1998

"Produktivitetsanalyser", Flemming Christensen, Peter Fristrup og Jens Leth Hougaard, Jurist- og Økonomforbundets Forlag, 1991

"Produktivitetsanalyse af Seminarieområdet", Finansministeriet og Undervisningsministeriet, marts 1992

"Lærerne skal dele gode erfaringer", Mai-Britt Herløv Petersen, Nyhedsmagasinet Danske Kommuner

"Midt i en evalueringstid", Henrik Toft Jensen, RUC NYT - information og debat fra Roskilde universitetscenter nr. 1, 31. August- 21 september 1998

"Guidelines for Self-Evaluation and Preparation for Review Team Visits", CRE Institutional Evaluation programme, Geneva, July 1998

"Institutional Evaluation as a Tool for Change", CRE Institutional Evaluation programme, Mai 1997

"Institutional Evaluation and Management Quality", CRE Institutional Evaluation programme: Background, Goals and Procedures, Geneva, July 1998

“The Value of Benchmarking management reporting”, Ian Malcolm Management Accounting, London, 1996, vol. 74, nr. 6, p. 18

“Benchmarking the performance of human resource management”, Long Range Planning, 1994, vol. 27, nr. 6, p. 43

“From knowledge to action: The impact of benchmarking on organisation performance”, Long Range Planning, 1997, vol 30, nr. 3, p. 427

“Vejledning om resultatkontrakter”, Økonomistyrelsen, juni 1996

“Er vi på rette spor?” - Fremtidsperspektiver for den statslige økonomistyring, Økonomistyrelsen, Artikelsamling 1998

“Mål og målinger” Økonomistyrelsen, maj 1998

“Sektorforskningens rolle og rammebetingelser”, Udkast til rapport om sektorforskningsinstitutioner af Kamma Langberg og Kristian Kindtler, Analyseinstitut for Forskning. Rapporten forventes at udkomme medio november 1998.

“Consensus conference on the theory and practice of research assessment - October 1996”, European Science Foundation, European Science Policy Briefing, nr. 3, June 1998