

De ufattelige tanker

- en perspektivering af området for kunstig intelligens

Speciale ved institut for filosofi, SDU Odense

Udarbejdet af:

Torsten Bo Hansen

Vejleder: David Favrholt

Indholdsfortegnelse

1	HOMO SAPIENS VERSION 2.0	4
2	GRUNDANTAGELSER I TEORIER OM KUNSTIG INTELLIGENS.	8
2.1	KAN COMPUTERE TÆNKE?	9
2.2	DET KLASSISKE PARADIGME	10
2.2.1	<i>Direkte repræsentation</i>	11
2.2.2	<i>Mekanisk fornuft – et paradoks</i>	13
2.3	DET KONNEKTIONISTISKE PARADIGME	17
2.3.1	<i>Overvåget indlæring</i>	19
2.3.2	<i>Hopfield-netværket</i>	19
2.3.3	<i>Ikke-overvåget indlæring</i>	21
2.4	PARADIGMERNES PLAUSIBILITET	22
2.5	BEVIDSTHEDSOPFATTELSER	24
3	DET KLASSISKE PSYKOLOGISKE INTELLIGENSBEGREB	27
3.1	IK, G-FAKTOR OG FLERFAKTOR	27
4	MULTIPLE INTELLIGENCE	32
4.1	INTELLIGENSBEGREBETS TERMINOLOGI OG ANALYSERAMME.....	34
4.2	DE MANGE INTELLIGENSER.....	37
4.3	DET BIOPSYKOLOGISKE POTENTIALE	39
4.3.1	<i>Subintelligenser</i>	40
4.4	BEVIDSTHEDSOPFATTELSEN	43
5	ENHEDSDUALISME	52
5.1	TEORIEN OM IMPLEMENTATION	52
5.2	ORGANISATIONEL INVARIANS.....	54
5.3	FÆNOMENALE OG FYSISKE EGENSKABER FOR BEVIDSTHED	55
5.4	FADING QUALIA	58
5.5	INFORMATIONSTILSTANDE.....	61
6	IMPLEMENTATION	64

7	DEN NATURVIDENSKABELIGE BEGRÆNSNING	70
7.1	WHAT IS IT LIKE?	70
7.2	SUBJEKT OG OBJEKT VIRKELIGHED	72
8	VERSION 3.0?	79
9	SUMMARY OF THE INCONCEIVABLE THOUGHTS.....	85
10	APPENDIKS	89
10.1	BREVVEKSLING MED HOWARD GARDNER	89
11	LITTERATURLISTE	94
11.1	BØGER.....	94
11.2	ARTIKLER.....	96

1 Homo Sapiens version 2.0

We know what we are, but not what we may become.

- William Shakespeare.

I bogen *The Age of Spiritual Machines* af Ray Kurzweil indleder dette citat af Shakespeare beskrivelsen af et fremtidsscenario år 2099. I fremtidsscenariet er der ikke så få ting, der tager sig radikalt anderledes ud fra den hverdag og virkelighed, som vi er bekendt med i dag. Som et resultat af anvendelsen af implantater, der optimerer menneskets erkendelse og maskinernes evne til at lære menneskelig intelligens, vil maskine og menneske i fremtiden erkendelsesmæssigt og intelligensmæssigt nærme sig hinanden. Skellet mellem menneske- og maskinintelligens vil derfor i sidste instans blive udvisket. Maskiners intelligens vil i fremtiden være baseret på modeller af menneskets intelligens, men fysisk vil deres intelligens være inkarneret i elektroniske og fotoniske kredsløb, som er direkte ækvivalenter til de menneskeligt kulstofbaserede celleprocesser. Maskinintelligensen vil ifølge Kurzweil, som et resultat af den teknologiske udvikling, overgå den menneskelige intelligens ikke bare i hastighed, men også i omfang! Bevidsthed vil i fremtiden ikke være bundet til nogen bestemt hardware, men i stedet eksistere som software, der kan 'køre' på forskellig hardware¹.

Mange vil tro, at der her er tale om science fiction, men *The Age of Spiritual Machines* er tværtimod ment som et seriøst bud på fremtidens samfund. Kurzweils hypotese er den, at den teknologiske udviklings fortsatte vækst uvilkårligt vil medføre en virtuel intelligensform, der med tiden vil overgå menneskets i enhver henseende. I bund og grund betragter Ray Kurzweil kunstig derfor intelligens som et spørgsmål om at have den fornødne computerkraft til rådighed. Men er intelligens i grunden alene et spørgsmål computerkraft, som Kurzweil antyder?

¹ Kurzweil 1999, s. 234.

Jeg levner meget plads til denne omtale af Ray Kurzweils seneste bog, idet han på mange områder nærmest udtrykker en ukritisk hengivenhed til vores fremtidige teknologiske for- måen på området for kunstig intelligens. Hengivenheden deles af en lang række forskere indenfor samme disciplin. I øjeblikket eksperimenterer den engelske professor i kyberne- tik, Kevin Warwick, med implantater og overvågning af nerveimpulser. Et af hans forsøg går ud på under smerte at tappe nervesignaler fra ham selv og derefter påføre hans kone disse. Lykkes dette, mener Warwick at de begge oplever nøjagtig den samme smerte!² Såfremt Kurzweils og Warwicks antagelser er korrekte, synes mange filosofiske problemer at have fundet deres løsning. Filosofihistorien taget i betragtning, er der derfor grund til her at indtage en kritisk position.

Det indledende citat af Shakespeare antyder, at vi ved , hvad vi er, men ikke hvad vi kunne blive³. Men ved vi nu også hvad vi er, eller har vi blot en række antagelser om vores egen menneskelige, og dermed bevidsthedens natur?

I et forsøg på at simulere eller syntetisere den menneskelige bevidsthed, må man nødven- digvis have et grundlag at gøre dette ud fra i form af en bevidsthedsopfattelse. De forskel- lige kognitionsteoriens grundlag - bevidsthedsopfattelsen eller intelligensopfattelsen - er set med filosofiens øjne en nærmere undersøgelse værd - ikke mindst for at afdække, hvordan de forskellige teorier definerer deres eget udgangspunkt - bevidstheden. En sådan undersø- gelse er desuden særligt interessant til at belyse, hvordan teorierne forholder sig til det filo- sofiske spørgsmål om forholdet mellem sjæl og legeme indenfor bevidsthedsfilosofien.

En ikke ualmindelig anskuelse indenfor visse teorier om kunstig intelligens er at betragte mennesket analogt med en computer, hvor hjernen antages at svare til en computers hard- ware og bevidstheden anskues som værende ækvivalent med en computers software⁴. Fo- restillinger som disse medfører, at grundlæggende filosofiske problematikker, som for ek- sempel sjæl-legeme problematikken, får om ikke endnu større, så i det mindste fornyet relevans. Filosofien synes altså på en frugtbar måde kunne indgå i et parforhold med kog- nitionsforskningen.

² Fyens Stiftstidende den 16/12-2001, Fokus s.4.

³ Oversættelse ved Ingelise Gullack.

⁴ Ray Kurzweils fremtidsvision beror således på betragtninger af denne slags.

Selve begrebet 'kunstig intelligens' dækker overordnet over en form for bevidsthed, nemlig intelligens, der søges genskabt i et kunstigt, det vil sige unaturligt miljø, - et miljø uden for mennesket. 'Kunstig' skal altså ikke forstås som værende falsk, men derimod blot som en genskabelse af intelligensen i et andet habitat end naturens egen⁵.

Kognitionsforskningen er i sit forsøg på en udtømmende beskrivelse af menneskets bevidsthedsnatur en hybrid sammensat af forskellige videnskabelige discipliner, der beskæftiger sig med henholdsvis psykologi, kunstig intelligens, lingvistik, logik, neurovidenskab og videnskabs- og sprogfilosofi⁶. At tage et filosofisk udgangspunkt synes altså ikke at være i modstrid med kognitionsforskningens grundlag. Indenfor kognitionsforskningen er filosofien dog kun partielt repræsenteret i form af disciplinerne videnskabs- og sprogfilosofi.

Det er min overbevisning at filosofien har mere at skulle have sagt i denne henseende og at filosofien kan bibringe med yderligere relevante betragtninger, der kan berige diskussionen om kunstig intelligens.

Som udgangspunkt for mine undersøgelser vil jeg, noget firkantet om man vil, gribe fat i formuleringen 'kunstig intelligens' for at se, hvad begrebet intelligens egentlig dækker over i relation til menneskets bevidsthedsnatur. Nu er kognitionsforskningen ikke det eneste område der beskæftiger sig med den menneskelige bevidsthed, endsige intelligensbegrebet. For således at nuancere denne undersøgelse yderligere, har jeg valgt at inddrage det psykologiske intelligensbegreb repræsenteret ved Howard Gardners teori om mange intelligenser.

Disse forhold har samlet givet anledning til følgende problemformulering:

Gennem en analyse af intelligensbegrebet indenfor Howard Gardners teori om Mange Intelligenser (Multiple Intelligences) vil jeg undersøge, hvorvidt dette psykologiske intelligensbegreb kan medføre nye perspektiver i diskussionen omkring kunstig intelligens.

Jeg har netop valgt at inddrage Howard Gardners intelligensteori ud fra den iagttagelse, at kognitionsforskningen og psykologien i grunden beskæftiger sig med det samme emneom-

⁵ Bernsen & Ulbæk 1993, s.30.

⁶ Bernsen & Ulbæk 1993, s.9-10.

råde – bevidstheden; blot har de et forskelligt sigte. I og med at begge discipliner omhandler bevidstheden i en eller anden forstand, er det mit håb, at den ene disciplin kan tjene til at belyse den anden.

En begrebslig afklaring synes allerede nu på sin plads. Med bevidsthed henvises til det forhold, der med John Searle karakteriseres som et indre, første-personlig, kvalitativt fænomen⁷ eller med Erich Klawonn som den essentielle subjektivitet ved oplevelser eller erfaringer⁸. Lad mig her understrege, at jeg som grundlag for opgavens betragtninger tager afsæt i samme karakteristik af dette fænomen; deraf følger ikke, at jeg skulle dele Searle eller Klawonns deraf senere følgende antagelser.

For at imødekomme problemformuleringens sigte introduceres indledningsvis en overordnet behandling af området for kunstig intelligens og områdets betragtninger omkring bevidsthed. Disse betragtninger knytter sig til det klassiske og konnektionistiske paradigme. Efterfølgende er det relevant at se på Howard Gardners intelligensbegreb og hvilke betragtninger han anlægger i forholdet omkring bevidsthed. Gardners brug af begrebet om computationelle egenskaber i relation til hjernen, synes af en sådan vigtighed for hans teori, at dette begrebs bevidsthedsmæssige implikationer efterfølgende søges anskueliggjort gennem David Chalmers brug af dette begreb indenfor hans teori, kaldet *Enhedsdualisme*. Chalmers synspunkt kan af hensyn til opgavens karakter ikke stå alene, hvorfor en diskussion med udgangspunkt i John Searles *biologiske naturalisme* efterfølgende vil finde sted. Idet opgaven indledningsvis starter med en række overordnede betragtninger omkring bevidsthed, for herefter at bevæge sig mere i dybden, anser jeg det for hensigtsmæssigt endnu engang at løfte perspektivet. En form for metaperspektiv på det grundlæggende forhold mellem subjekt og objekt, samt en kritisk stillingtagen til vor egen position i bedømmelsen af dette forhold, rejses med Thomas Nagels objektive fænomenologi, hvorfor denne position sluttelig vil blive inddraget.

I og med at undersøgelsen har til formål at nuancere og perspektivere aspekter ved kunstig intelligens vil fremstillingen af de valgte teorier primært være problematiserende hen imod at afdække bevidsthedens relation hertil.

⁷ Searle 1997, s.5.

⁸ Klawonn 1991, s.128.

2 Grundantagelser i teorier om kunstig intelligens.

Nærværende kapitel har til formål at belyse bevidsthedsopfattelsen, der elaboreres med indenfor teorier om kunstig intelligens. Af selv samme grund har jeg valgt ikke at gå i detaljer med teoriernes tekniske struktur, men udelukkende at beskæftige mig med aspekter, der kan belyse de anvendte bevidsthedsbegreber og problemer herved.

Inden vi differentierer området for kunstig intelligens er det hensigtsmæssigt at se på, hvad man egentlig forstår ved kunstig intelligens. Kunstig intelligens er en samlebetegnelse, der dækker over:

”...de samlede bestræbelser i teori og praksis på at få en computer til at overtage eller assistere enhver menneskelig intellektuel funktion.”⁹

Kunstig intelligens, hvad enten den er syntetiseret eller ’blot’ simuleret, er således et forsøg på at tildele en computer menneskelige egenskaber. Det forhold at computere ikke besidder en krop, ser man indenfor kunstig intelligens ikke som en hindring for at kunne syntetisere eller simulere den menneskelige bevidsthed¹⁰. Som udgangspunkt kan man derfor konstatere at grundlaget for teorier om kunstig intelligens beror på en skelnen mellem bevidsthed og krop, hvor bevidstheden betragtes som en autonom størrelse, der således kan isoleres fra kroppen og studeres for sig selv¹¹.

Et udgangspunkt for en beskrivelse af erkendelsen er, at der ligger nogle grundantagelser til skue. Grundantagelserne fungerer som en slags paradigme, inden for hvilke man søger at beskrive fænomenet kognition. Teorier om kunstig intelligens kan deles op i to sådanne paradigmer: det klassiske- og det konnektionistiske paradigme¹² eller om man vil, repræ-

⁹ Kirkeby 1988, s.24.

¹⁰ Ole Fogh Kirkeby ser kroppens relation til bevidstheden som essentiel for vort bevidsthedsbegreb. Kirkeby 1988.

¹¹ Kirkeby 1988, s.24. Nyere forskning indenfor kunstig intelligens peger dog i højere grad hen imod en inddragelse af kontekst. I denne sammenhæng er robotens ’krop’ uundværlig i og med at sensorer transmitterer data, hvorved robotten interagerer med sine omgivelser og herigennem antages at udvikle mentale egenskaber jf. Juyang Weng m.fl.: *Autonomous Mental Development by Robots and Animals*.

¹² Bernsen & Ulbæk 1993, s.9ff.

sentations- og transformationsperspektivet¹³. Lad os se nærmere på disse paradigmer og deres iboende antagelser omkring bevidsthedens struktur.

2.1 Kan computere tænke?

Det klassiske paradigmes syn på intelligens kan karakteriseres med udgangspunkt i den britiske matematiker Alan Turing's artikel "*Computing Machinery and Intelligence*"¹⁴, hvor Turing stiller spørgsmålet "kan en maskine tænke?". Turing giver med artiklen en beskrivelse af grundlaget for hvornår vi kan betragte en maskine for værende intelligent. For at kunne tale om intelligente maskiner må disse ifølge Turing passere en simpel test. Selve testen, der også kaldes Turing-testen, består af et menneske, der agerer dommer og som gennem en computer stiller spørgsmål til to andre forbundne enheder, hvoraf den ene er et menneske og den anden en computer¹⁵. Hvis dommeren ikke regelmæssigt kan skelne mellem svarene fra computeren til menneskets svar, så har computeren klaret testen og kan betegnes som intelligent.

Ifølge Turing kan man altså karakterisere en maskine som værende intelligent når den kan snyde en til at tro, at den er et menneske frem for en maskine. En af problematikkerne, der opstår heraf, er spørgsmålet om intelligensbegrebet kan beskrives udtømmende gennem adfærdsmønstre. For at Turing kan antage, at en maskine er intelligent, i det øjeblik denne via dens adfærd er i stand til at 'bedrage'¹⁶ en observant kræver det, at han sætter lighedstegn mellem computerens interne og eksterne tilstande¹⁷. Netop distinktionen mellem interne og eksterne tilstande viser den menneskelige/kognitive dimension. Det primært subjektive aspekt er ikke noget, der kan beskrives udefra, men må beskrives som en indre tilstand ved et system. Turing's behavioristiske intelligensfortolkning efterlader derfor i stedet flere spørgsmål om bevidstheden end den giver svar.

¹³ Kirkeby 1988, s.32.

¹⁴ Alan Turing: "*Computing Machinery and Intelligence*". Mind 1950, vol. LIX, no. 236.

¹⁵ Testen er her reformuleret med nutidige termer, for at vise at Turing's pointe fortsat har relevans.

¹⁶ Jeg mener ikke at computeren bevidst (intentionelt) bedrager observatøren, men nærmere at observatøren bedrager sig selv, hvad der spiller en væsentlig rolle i denne sammenhæng.

¹⁷ Altså skelne mellem, hvad computerens tilstand er i og for sig selv og tilstande som computeren tillægges.

Bevidsthedsopfattelsen, som Turing fremfører, træder endnu tydeligere frem med udformningen af Turing-Church tesen, der antager, at enhver databehandlingsproces kan udføres på en Turing-maskine¹⁸. Når alle databehandlingsprocesser kan udføres på en Turing-maskine og kognition antages at være lig databehandlingsprocesser, har man belægget for at kognition kan eftergøres på en sådan maskine¹⁹. Den moderne digitale computers programmeringssprog gør det i vore dage ud for en slags abstrakte maskiner, der kan sammenlignes med Turing-maskiner²⁰.

Forestillingen, at betragte en computer som analog til hjernen (*brain*) og et program eller software som værende konform til bevidstheden (*mind*), er med Turing og Churchs tese meget nærværende. Ikke mindst det klassiske paradigme har fundet denne sammenligning tiltrækkende, hvorfor vi nu vil se nærmere på denne retning indenfor kunstig intelligens.

2.2 Det klassiske paradigme

Det klassiske paradigme beror på antagelsen, at kognition betragtet som system kan side-stilles med databehandlingssystemer, som eksempelvis Turing-maskinen. Oveni kommer at man med Claude Shannons informationsteori, hvor information betragtes som værende uafhængigt af både indholdet af den overførte information, samt uafhængig af den materielle implementering af informationssystemet, finder yderligere belæg for, at muligheden for kunstig intelligens er indenfor rækkevidde. Det forhold at betragte information som værende uafhængigt af et subjekt²¹ og indhold er dog ikke uproblematisk. Teorien levner ikke plads til den subjektive fortolkning, men opererer med en form for objektivistisk viden, der muliggør den upersonlige og, fristes man til at sige, indholdsløse kommunikation²².

¹⁸ Tesen lyder: "For any deterministic automatic formal system whatever, there exists a formally equivalent Turing machine." Haugeland 1987, s.137.

¹⁹ Bernsen & Ulbæk 1993, s.14.

²⁰ Bernsen & Ulbæk 1993, s.40.

²¹ Vel vidende at Shannons terminologi ikke indbefatter et subjektivt synspunkt, har jeg valgt begrebet subjekt, da det er alment accepteret indenfor kommunikationsteorier, at subjektets rolle netop spiller en signifikant betydning.

²² Der er mange andre kritisable forhold ved denne informationsteori, som jeg dog af hensyn til opgavens karakter ikke vil inddrage nærmere.

Den klassiske opfattelse af intelligens som værende analog med informations- og databehandling medfører således en række konsekvenser for synet på kognition.

Menneskets kognition opfattes som et system, hvor dette systems virke består i at manipulere repræsentationer. Det afgørende for kognition er således ifølge det klassiske paradigme ikke det fysiske medium, hvori systemet er implementeret, men derimod den bestemte måde som systemet manipulerer repræsentationer på²³.

Menneskets intelligente virksomhed kan altså beskrives ud fra repræsentationsniveauet. Det intelligente udspringer så at sige af det ordnede bevidsthedsindtryk eller repræsentationsniveauets struktur. Repræsentationen indenfor det klassiske paradigme består af diskrete symboler.

Muligheden for at kunne tale om kunstig intelligens bliver altså virkeliggjort blandt andet gennem databehandlingsteorien og forestillingen om bevidsthedens symbolmanipulerende karakteristika, der antages at fungere gennem direkte repræsentation.

2.2.1 Direkte repræsentation

Naturlige kognitive systemer er interessante i repræsentationsmæssig sammenhæng i den forstand, at de betragtes som intelligente i kraft af deres egenskab som værende fysiske symbolsystemer af den *rette* slags²⁴. Bevidsthed opfattes altså med klassisk kunstig intelligens øjne således som værende skabt gennem direkte repræsentationer, også kaldet symbolsk repræsentation²⁵. Begrebet om repræsentationer skal opfattes som:

”..udtryk med sandhedsværdi, dvs. som en slags udsagn, bortset fra, at de er konkrete, analoge med fysisk foreliggende sætninger: De er sætningslignende, symbolske udtryk.”²⁶

²³ En forudsætning for at kunne tale om intelligens overhovedet er dog ifølge det klassiske paradigme, at denne knytter sig til et fysisk objekt.

²⁴ Newell og Simon udarbejdede i 1976 denne hypotese, som kognitionsforskningen har overtaget (kursivering er min egen).

²⁵ Modsætningen, den distribuerede repræsentation, vil blive diskuteret under det konnektionistiske paradigme.

²⁶ Bernsen & Ulbæk 1993, s.241.

Relationen mellem objekt og repræsentation er i den forstand en kausal relation, hvor objektet repræsenteres direkte i menneskets bevidsthed i kraft af et symbol.

Med repræsentationsteorien viser sig yderligere det forhold, at man indenfor det klassiske paradigme opfatter en abstrakt similaritet imellem tanke og sprog. Definitionen på tanke og hermed erkendelse bliver indenfor det klassiske paradigme således den, at:

”Distinct thoughts...are generally complex, constructed systematically (according to their content) from a comparatively modest kit of atomic constituents; in other words, the versatility of thought is attributed to a combinatorial structure essentially like the compositionality of symbolic systems”²⁷

Tankens alsidighed søges altså indkapslet i det symbolske systems komposition/grammatik. Denne kompositionelle betragtning gør at repræsentationsteorien kan udvides med betragtninger ud fra to retninger, en internalistisk og en eksternalistisk. Den internalistiske opfattelse ser repræsentationens indhold som værende determineret internt i forhold til systemet selv – altså alene bestemt af kompositionaliteten. Bevidsthed, og dermed det semantiske indhold af repræsentationerne er altså udelukkende afgjort af systemets selv og afhænger ikke, som for den eksternalistiske repræsentationsteori, af noget udenfor systemet²⁸. Det klassiske paradigmes opfattelse af repræsentation er eksternalistisk og inddrager konteksten i og med at den læner sig op ad denne i forklaringen af symbolernes indhold. Forståelsen er således i en grad båret af overensstemmelsen mellem symbol og det symboliserede. For eksempel er begrebet ’hund’ på denne måde afledt af den objektivt foreliggende hund. Observerer jeg følgelig en kat, men i stedet tænker begrebet hund i forstanden, at jeg misforstår det observerede, havner repræsentationsteorien i problemer. Teorien kan ikke forklare misforståelsen andet end gennem en korrektion af den oprindelige begrebslige repræsentation fra betydningen ’hund’ til disjunktionen ’hund og kat’ - deraf benævnelsen disjunktionsproblemet. Udvider vi begrebsdannelsen til de mere abstrakte begrebsdannelser som for eksempel klassen dyr, ser vi, at den direkte repræsentationsteori ikke umiddelbart kan redegøre for disse begrebsdannelser, da de er afledte, syntetiserede størrelser. Analogt til ’Fido’-Fido teorien, der må imødesee problemer når den eksempelvis

²⁷ Haugeland 1987, s.93.

²⁸ Dancy & Sosa (ed.) 1998, s.443.

skal redegøre for det meningsfulde i en sætning, som ikke har noget objekt, eller når det samme ord har forskellige betydninger, har repræsentationsteorien problemer med at redegøre for åbenbart indlysende forhold.

Problemerne for den klassiske opfattelse må tilskrives *skolens*²⁹ udgangspunkt – overbevisningen om den mekaniserbare fornuft.

2.2.2 Mekanisk fornuft – et paradoks

John Haugeland beskriver forsøget på at mekanisere den menneskelige kognition som *The Paradox of Mechanical Reason*³⁰. Den mekaniske redegørelse for den menneskelige erkendelse/kognition må, som vi har set det forsøgt indenfor den direkte repræsentationsteori, redegøre for hvordan og hvor meningsfuldhed opstår. Overordnet betragtet kan man inddele de mulige løsninger op i to filosofiske skoler – den dualistiske og den monistiske (*materialisme*). Ret beset er den dualistiske løsning ikke særlig hensigtsmæssig i et forsøg på at mekanisere den menneskelige erkendelse. Dualismen beskriver væsensforskellen mellem det fysiske (*res extensa*) og det psykiske (*res cogitans*), hvorfor en forening af disse i en mekanicistisk model synes håbløs. Sjæl-legeme problematikken omhandler netop det forhold, at disse væsensforskellige og adskilte verdener synes at have et samspil, som vi dog ikke kan redegøre nærmere for uden på det nærmeste at havne i en eller anden form for erkendelsesteoretisk skepticisme. Samspillet kan derfor ikke beskrives mekanisk uden at man er nødsaget til at forkaste den moderne fysiks grundlag, der netop beror på interaktionen mellem udstrakte, fysiske entiteter³¹.

Materialismen synes således umiddelbart mere attraktiv i forsøget på at mekanisere erkendelsen. Men netop gennem materialismen opstår paradokset om den mekaniske erkendelse, der i sin enkelhed går ud på, at:

²⁹ Begrebet skole karakteriserer de forskellige teorier, der indenfor det samme paradigme har søgt at løse bl.a. problemerne omkring repræsentation gennem forskellige teoretiske varianter. Jerry Fodor har bl.a. gennem begrebet om asymmetrisk afhængighed søgt at løse den direkte repræsentations problematikker. Fodor skelner i sin teori skarpt mellem hjernens neurologiske egenskaber og de kognitive egenskaber. Trods dette undgår han ikke den mekaniske problematik – fyldestgørende at redegøre for de kognitive egenskaber.

³⁰ Haugeland 1987a, s.36.

³¹ Haugeland 1987a, s.37.

”..if a process or system is mechanical, it can't reason; if it reasons, it can't be mechanical.”³²

Paradokset består af to horn eller to sider - en fornuftsside³³ og en mekanisk side³⁴. På fornuftssiden har man forsøgt at løse problemet ved at tilskrive bl.a. forståelsen eksistens til en slags indre 'manipulator' - en homunculus³⁵. I og med at man forsøger at redegøre for erkendelsen på en mekanisk måde og samtidig fastholder, at tankesymbolerne og reglerne skal give mening, må denne side indføre begrebet om en slags indre homunculus, hvortil denne mening viser sig gennem regelmanipulation.

Selve ideen om en indre homunculus leder dog hurtigt til en uendelig regres i og med at man nu er nødsaget til at forklare, hvordan tanken opstår her, og dette kan kun begrundes i en anden instans - en anden homunculus og så fremdeles³⁶. Thomas Hobbes var en af de første til at manifestere denne retning. Hobbes antagelse, at tanker kunne sidestilles med eksempelvis ord, sammenholdt med ideen om at den rationelle menneskelige erkendelse bestod i en slags regelbestemte manipulationer, ledte til, at disse fysiske symboler (tanker) måtte være manipuleret af noget. Dette *noget* tjener, som vi har set, ikke til nogen særlig god forklaring på hvorledes det mekaniske paradoks løses, men efterlader tværtimod Hobbes og dermed hele denne skole hensat til en uendelig regres i den mekaniske forklaring på erkendelsen.

En anden måde, som det mekaniske paradoks er søgt løst ved, er gennem den mekaniske side. Den mekaniske side kan karakteriseres med David Humes forsøg på at redegøre for erkendelsen. Hume søgte at beskrive erkendelsens grundlag videnskabeligt analogt til Newtons fysik. Hume antog således, at erkendelsens byggesten, *ideas*, var analogt til fysiske partikler, men alligevel forskellig fra disse. Hume antog, at erfaringen dannedes gen-

³² Haugeland 1987a, s.39.

³³ Jeg anvender her begrebet 'fornuftsside' eftersom begrebet 'erkendelsesmæssig side' virker malplaceret idet begge løsninger netop i søger at tage hånd om det erkendelsesmæssige i deres mekaniske forklaringer.

³⁴ Disse to sider har utallige filosofiske eksponenter. Ideen om kunstig intelligens er ikke ny, hvad fremgår af at fornuftssiden repræsenteres af blandt andre Thomas Hobbes. Den mekaniske side er blandt andre repræsenteret gennem David Hume.

³⁵ Lat.: små mennesker.

³⁶ Haugeland 1987a, s.41.

nem interaktionen af disse *ideas*. Problemet med Humes tilgang skal ses deri, at ikke-fysiske entiteter antages at agere i overensstemmelse med naturlige kræfter jf. Newtons fysik. Dertil kommer, at Hume ikke kan redegøre for, hvad der gør hans *ideas* til det, de er, og hvorfor deres interaktioner nødvendigvis tæller som tænkning³⁷.

Begge søgte løsninger på *den mekaniske fornufts paradoks* havner i en uholdbar situation, hvor det egentligt fornuftsmæssige, kognitionen, mister sit særkende. Den mekaniske tilgang afføder for begge retningers vedkommende, at kognitionen reduceres til noget fysisk, hvorved forskellen til andre fysiske entiteter bliver stadig mere utydelig. Hvor mennesket som udgangspunkt antages at besidde bevidsthed, til forskel fra andre fysiske entiteter, er endepunktet knapt så indlysende.

Ideen om at beskrive erkendelsen mekanisk, er som det fremgår, ikke ny. Alligevel synes det som om, at netop nyere forsøg udi denne retning havner i den samme paradoksale situation, hvor som vist:

”either meanings matter to the manipulations, in which case the processes aren’t really mechanical (they presuppose a homunculus); or else meanings don’t matter, in which case the processes aren’t really rational (they’re just some meaningless ‘machinelike’ interactions).”³⁸

Kunstig intelligens indenfor det klassiske paradigme træder lige ind i det mekanistiske paradoks i sit forsøg på at syntetisere den menneskelige erkendelse i en computer³⁹. Denne form for kunstig intelligens, der også af Searle er benævnt stærk kunstig intelligens⁴⁰, betaler således en høj pris i sit forsøg på syntetisere erkendelsen. Som udgangspunkt for at syntetisere erkendelsen, har man indenfor det klassiske paradigme beskæftiget sig med erkendelsen, som om disse kognitive mekanismer fungerede analogt med sprogets distinktion

³⁷ Haugeland 1987a, s.44.

³⁸ Haugeland 1987a, s.44.

³⁹ Syntetisering skal her forstås således at et ”*kunstigt intelligent system ikke blot efterligner naturlige systemers adfærd, men også producerer intelligent adfærd på samme måde som naturlige systemer.*”. Bernsen & Ulbæk 1993, s.31.

⁴⁰ Searle, John R.: “*Minds, Brains and Programs*” i Haugeland 1987 (b), s.282-306. Searles distinktion rubricerer teorier, der hævder at tilstrækkeligt programmerede computere i bogstaveligste forstand vil besidde mentale egenskaber, som stærk kunstig intelligens.

mellem indhold/udtryk. Begrebet om talehandlinger understøtter netop ideen om at sproglige udtryk modsvares af en korresponderende kognitiv tilstand. Skal talehandlingerne overhovedet have nogen betydning, må det kognitive indhold i én forstand være originalt. Men netop det originale, det kognitives *primære* funktion, bevidsthedens semantiske aspekt, synes at undslippe den klassiske skoles teorier. Syntetisering af menneskelig intelligens med udgangspunkt i de psykologiske mekanismer synes altså at komme til kort i og med, at man helt fjerner det subjektive psykologiske aspekt og dermed stadig mangler at redegøre for, hvordan systemet kan opfatte sig selv⁴¹. Dertil kommer, at den klassiske skole, i forsøget på at syntetisere intelligens benytter sig af digitale computere, der ifølge John Searle aldrig vil være i stand til at implementere en semantik. Det indlysende forhold er ifølge Searle, at:

”If it really is a computer, its operations have to be defined syntactically, whereas consciousness, thoughts, feelings, emotions, and all the rest of it involve more than a syntax”⁴²

Digitale computere vil ifølge Searle altid kun være i stand til at manipulere meningsløse symboler, hvor tankens egenart netop består i en manipulation af meningsfuldt semantisk indhold. Tanken om at syntetisere intelligens må altså opgives⁴³.

Hvor det klassiske paradigme tager udgangspunkt i de psykologiske mekanismer og semantikken i en syntetisering af den menneskelige intelligens, tager det konnektionistiske paradigme i stedet udgangspunkt i, hvorledes den menneskelige hjernes neurale netværk fungerer. Vi vil i det følgende se nærmere på dette paradigmes karakteristika.

⁴¹ Det er klart at en maskine som Deep Blue er intelligent i én forstand i og med at den mestrer skakspillet, men den er ikke tilnærmelsesvis intelligent i menneskelig forstand. Menneskelige egenskaber som det at bluffe synes ikke at være en egenskab som Deep Blue besidder. Skulle den på et tidspunkt formå at bluffe sin modspiller vil den stadig ikke selv være bevidst herom. Så længe Deep Blue ikke er bevidst om sine handlinger kan den ikke besidde intelligens i menneskelig forstand og i realiteten ikke bluffe i ordets forstand.

⁴² Searle 1984, s.37.

⁴³ Searle 1984, s.36-37.

2.3 Det konnektionistiske paradigme

Til forskel fra det klassiske paradigme hævder konnektionismen ikke at syntetisere kognitionen, men i stedet at simulere denne. Denne form for kunstig intelligens hævder, at den simulerede intelligens ikke nødvendigvis er direkte ækvivalent med den oprindelige menneskelige intelligens. Konnektionismen opererer dog stadig med et begreb om intelligens og har i de senere år vundet terræn i forhold til det klassiske paradigme.

Konnektionistiske netværk er som nævnt et forsøg på at kopiere strukturen i de neurale netværk. Udkommet er således, gennem brug af neurale netværk, at få maskiner til at udvise en adfærd, vi som mennesker vil betegne som intelligent⁴⁴. Det klassiske paradigme anvender som udgangspunkt for syntetiseringen serielle computermodeller, der som fælles-træk har programmeret alle aspekter af modellen og benytter eksplicitte symboler til at repræsentere begreber. Indenfor et konnektionistisk netværk er:

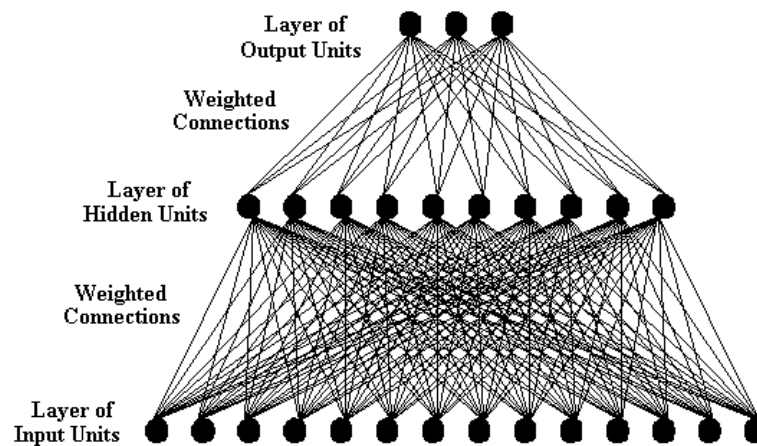
”..et begreb ikke repræsenteret i et enkelt knudepunkt eller enhed, men *fordelt over netværket som helhed*, uden eksplicitte regler eller symboler.”⁴⁵

Netværk kaldes derfor også typisk for ’parallelt fordelt processering’ (*Parallel Distributed Processing*). Det neurale netværks struktur består af et stort antal simple forbundne enheder, processorer, der indbyrdes justerer forbindelsen til hinanden. Det er gennem denne justering, at netværket lærer. Det konnektionistiske system er en idealiseret⁴⁶ model af det neurologiske netværk og kan eksempelvis tage sig ud på følgende måde:

⁴⁴ Poul René Bertelsen 1999, s.13.

⁴⁵ Gade 1998, s.107.

⁴⁶ Netværket er idealiseret, idet man ikke kender det fulde omfang af det neuronale netværk og dets virkemåder.



Figur 1⁴⁷

Et netværk kan altså bestå af nogle inputenheder, der på baggrund af inputtet beregner et output, som så sendes videre til næste lag af databehandlingsenheder, i dette tilfælde et skjult lag (se fig.1). Dette lag beregner ligeledes en outputværdi, som således sendes videre til det næste lag, hvorefter dette lag beregner systemets endelige output. Som det fremgår af figur 1 er der indbyrdes mellem lagene en række vægtede forhold. Netværket består altså af to typer bestanddele: *neuroner* og *forbindelser* mellem neuroner⁴⁸. Inputneuronerne er afhængige af input fra omverdenen. Det er således input fra den ydre verden, der bestemmer, om inputneuronerne skal sende signalet videre til det skjulte lag. En neuron i det skjulte lag fungerer ved at have en værdi, kaldet *tærskelværdien*, der afgør hvornår den skal sende et signal videre. Hvis signalinputtet når over denne værdi, vil neuronen således sende signalet videre. Denne transmission foregår populært sagt ved at neuronen fyrer. Outputneuronerne sender inputtet videre, såfremt dette overskrider tærskelværdien for disse. Forbindelserne mellem det skjulte neurale lag og input/output lagene er defineret ved en talværdi, kaldet en vægt. Vægten angiver kraften af det signal, som passerer forbindelsen⁴⁹. Det ovenfor beskrevne netværk er kun en modeltype af de mange netværkssystemer, der findes. Netværk med bestemte tærskelværdier og vægte er såkaldte 'færdige' netværk, hvor inputtet giver det ønskede output. Heroverfor står netværkstyper, hvor vægte ikke er givet på forhånd, men først skal oplæres. De forskellige netværkstyper kan inddeles i systemer med overvåget indlæring og systemer uden overvåget indlæring.

⁴⁷ Figuren er hentet fra artiklen *What is Connectionism?* af István S.N.

⁴⁸ Poul René Bertelsen 1999, s.16.

⁴⁹ Poul René Bertelsen 1999, s.17.

2.3.1 Overvåget indlæring

Et system med overvåget indlæring kan, på baggrund af et systematiseret inputmateriale, 'oplæres' til give korrekte svar på dette inputmateriale, samt fornuftige svar på inputmateriale, som netværket ikke har mødt før. Det væsentlige i denne sammenhæng er, at man ikke, som indenfor det klassiske paradigme, programmerer sammenhængen, men at netværket så at sige selv opdager denne sammenhæng.

Denne 'oplæring' af netværket kan for eksempel ske ved at man ud fra det fremkomne output arbejder sig baglæns i systemet, og her korrigerer vægtene, så outputtet bliver korrekt. Denne procedure kaldes også back-propagation. Netværk, der anvender denne form for beregningsprocedure, kaldes også for feed-forward netværk, idet processen kun går én vej: fra input til output. En ofte anvendt regel til denne procedure er Delta reglen, som regulerer netværket ved:

"...først at beregne fejlen på outputenhederne, dvs. deres afvigelse fra det ønskede output, og justere vægtene på forbindelserne til dem. Derpå beregnes fejlen på laget før outputlaget og vægtene på forbindelserne til dette lag justeres. Processen fortsættes ned gennem netværket til inputenhederne og gentages i mange træningscykler med små justeringer ad gangen."⁵⁰

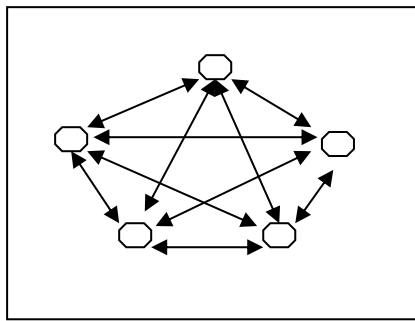
Netværket lærer altså via deltareglen, således at regulere sig til det ønskede output fremkommer. Netværkstyper af denne slags kan bl.a. udføre mønstergenkendelse og genkende bogstaver uvilkårligt i forhold til deres orientering⁵¹. Andre netværkstyper har gjort endnu større fremskridt på disse områder, så lad os forlade "feed-forward" netværkene, og i stedet se på netværk, som anvender "feed-back" i deres informationsbehandling.

2.3.2 Hopfield-netværket

Netværkstyper, som anvender "feed-back", har ikke den samme fastlåste arkitektur som "feed-forward" netværkene. Informationsbehandlingen kan her, som navnet antyder, både gå frem og *tilbage* mellem neuronerne. Hvert neuron fungerer både som input- og output-neuron, hvad der kan illustreres med Hopfield-netværket:

⁵⁰ Bernsen & Ulbæk 1993, s.216.

⁵¹ Bernsen & Ulbæk 1993, s.217.



Figur 2.

Giver man dette netværk et inputmønster, vil neuronerne fortsætte med at transmittere signaler frem og tilbage mellem hinanden og henholdsvis tænde og slukke hverandre alt afhængig af, om de i forvejen er tændte eller slukkede og i relation til de bestemte vægtningsforhold. Netværket vil arbejde sig hen imod en tilstand, hvor et endeligt mønster af tændte og slukkede neuroner er givet. Dette mønster er lig med outputtet⁵². Givet, at netværket er indlært et bestemt mønster, gælder der, at:

”Hvis man som input har en ufuldstændig eller til en vis grad fejlagtig kopi af dette mønster, vil outputtet være det korrekte mønster. Hvis inputtet er for ufuldstændigt og for fejlagtigt, risikerer vi, at outputtet er fejlagtigt, men der kommer et output.”⁵³

En af styrkerne ved denne netværkstype er, at det på mange områder afspejler de kognitive perceptuelle mekanismers egenart. Descartes eksempel med firkantede tårne, der på lang afstand fremstår som runde, er et eksempel herpå⁵⁴. På lang afstand får vi netop et ufuldstændigt sanseindtryk, som vi selv færdiggør (de runde tårne). Jo mere vi nærmer os tårnet, desto tydeligere bliver sanseindtrykkene i og med, at vi modtager flere og flere fragmenter af det korrekte mønster og vores fortolkning af det sete ændres pludselig fra rundt til firkantet. Husker vi på, at konnektionismens udgangspunkt netop var at simulere intelligens ved at få maskiner til at udvise en adfærd, vi som mennesker vil betegne som intelligent, er Hopfield-netværkets styrke, at det er i stand til at efterligne den menneskelige perceptions karakteristika.

⁵² Poul René Bertelsen 1999, s.26.

⁵³ Poul René Bertelsen 1999, s.26.

⁵⁴ Hartnack & Sløk 1991, s.183.

En sidste type af netværk, som jeg her vil nævne, er netværk af typen ikke-overvåget indlæring.

2.3.3 Ikke-overvåget indlæring

Netværk af denne type kræver ikke et ordnet eller struktureret input, men ordner derimod selve inputtet selv. Netværket afspejler på denne vis de reale menneskelige forhold, som erkendelsen må arbejde under idet, at vi som mennesker ofte ikke udsættes for struktureret input, men selv ordner dette. Kohonen-netværket, som er af denne type, er i praksis opbygget af to lag neuroner: et input og et output lag. Det signifikante ved denne netværkstype er den høje grad af forbindelser. Hver neuron i inputlaget har forbindelse med hver neuron i Kohonen-laget, der fungerer som output lag og hvor neuronerne indbyrdes har forbindelse til alle andre kohonen-neuroner i samme lag. Vægtene mellem inputneuronerne og kohonen-neuronerne varieres således, at når en inputneuron fyrer, efterfulgt af en kohonen-neuron, styrkes forbindelsen mellem disse⁵⁵. Omvendt vil forbindelsen mellem disse neuroner blive svækket, hvis dette forløb ikke er til stede. For kohonen-neuronerne gælder at:

”De vægte, der er knyttet til forbindelserne mellem kohonen-neuronerne indbyrdes, er fastlagt således, at vægtene mellem tætliggende neuroner er positive, dvs. neuronerne aktiverer hinanden. Vægtene aftager med afstanden, og mellem fjerntliggende neuroner er vægtene endog negative, dvs. fjerntliggende neuroner hæmmer hinanden.”⁵⁶

Dette medfører, at jo stærkere et signal, der passerer til en neuron i kohonen-laget, desto større aktivitet vil der være i det nærliggende lag, og samtidig vil forbindelsen til de fjerntliggende neuroner hæmmes. Dette princip bevirker, at netværket vil strukturere sig således, at et område eller en enkelt neuron med tiden vil repræsentere og svare til et bestemt mønster. Ved eksempelvis at introducere netværket for alfabetet, vil det med tiden lære dette i form af en repræsentation, givet ved et aktivt område eller neuron, for hvert givent bogstav. Netværket oplærer altså sig selv uden at inputtet er struktureret eller på forhånd ud-

⁵⁵ Dette princip kaldes også for Hebb'sk indlæring efter Donald Hebb.

⁵⁶ Poul René Bertelsen 1999, s.28.

valgt. På denne måde skaber netværket således et billede af virkeligheden relativt til de input, der er tilgængelige.

Det er nu tydeligt, hvorfor man kalder repræsentationen for distribueret. Afbildningen af omverdenen sker altså gennem en kontingent struktur repræsenteret ved netværkets struktur⁵⁷. Et forhold repræsenteres ikke direkte gennem en symbolsk enhed, men repræsenteres derimod ud af interaktionen af de forskellige neuroner, som indgår i et netværk.

2.4 Paradigmernes plausibilitet

Neurale netværk har vist, at de er i stand til at behandle mange former for perception på lige vilkår med de menneskelige erkendelsesbetingelser og følgelig med en psykologisk plausibilitet, der ikke kan siges at gælde for den klassiske opfattelse⁵⁸.

Forholdet omkring repræsentation synes at give et fortrin til den konnektionistiske skole, idet beskadigelse af netværket medfører en forringet repræsentation og ikke, som indenfor den klassiske repræsentation medfører, at repræsentationen helt falder bort. Neurale netværks evne til at ordne indtryk, og dermed 'lære', har ydermere vist en overlegenhed på sproglige områder⁵⁹, der af mange betragtes som den klassiske skoles kardinalområde. Tager man begge paradigmers udgangspunktet i betragtning, maskinelt at syntetisere eller simulere intelligens, viser det sig at:

”Indlæring synes at være en fundamental og uundværlig komponent i al intelligens. Den giver en tilpasningsevne til komplekse og skiftende omgivelser, uden hvilken kunstige systemer vanskeligt vil kunne operere i realistiske situationer. At give maskiner indlæringsevner som del af deres primitive, dvs. ikke-indlærte, viden og evner opfattes derfor som et hovedproblem i udviklingen af kunstig intelligens. Alternativet, som ikke er realistisk, er fra starten at programmere maskinen til at vise intelligent performans i enhver situation.”⁶⁰

⁵⁷ Kirkeby 1988, s.45.

⁵⁸ Med psykologisk plausibilitet mener jeg her den grad af adfærdsmæssig similaritet, der er imellem det oprindelige/naturlige system på den ene side og det kunstige system på den anden side.

⁵⁹ Blandt andet har T. Kohonen på baggrund af denne netværkstype i 1986 udviklet en såkaldt 'fonetisk skrivemaskine', hvor netværket transformerer lydligt input til skrift med 95% præcision.

⁶⁰ Bernsen & Ulbæk 1993, s.145.

For at en robot eller computers adfærd overhovedet skal komme i nærheden af at kunne betegnes som intelligent, må denne kunne agere passende i forhold til dens omgivelser. Neurale netværk kan siges at have et psykologisk forspring i forsøget på skabe kunstig intelligens, idet neurale netværkssystemer i langt højere grad end de klassiske systemer udviser en menneskelignende adfærd - uden på forhånd i hver mindste detalje at være programmeret dertil⁶¹. Derfra og så til helt at afskrive den klassiske skole, mener jeg ikke er rimeligt. Groft sagt kan man hævde, at neurale netværk reducerer intelligens til input/output regelmæssigheder i adfærd uden derved at forklare det egentligt subjektive - forståelsen. En del af menneskelig intelligent adfærd er, at denne opviser en forståelse. Som menneske er vi i stand til at reflektere over, om en adfærd er passende eller upassende – situationsfornemmelsen forudsætter således forståelse. Det er således svært at forestille sig, at et neuralt netværk på baggrund af sin klassificering af alfabetet skulle opnå nogen forståelse af dette – sammenhængen bør ekspliciteres nærmere. Tilmed har man søgt at indvende imod det konnektionistiske paradigme at det ikke er i stand til at repræsentere nogen syntaktisk struktur, hvorfor det i dybeste forstand heller ikke kan repræsentere tanken⁶². Denne indvending, som hyppigt kommer fra fortalere af den klassiske skole, beror på ideen om, at tanker nødvendigvis har den samme struktur som sproget. Det er tydeligt, at man i et konnektionistisk system vil have repræsentationer i form af komplicerede aktivitetsmønstre. Dette medfører at:

”Siden systemet repræsenterer begreber og deres relationer gennem komplicerede aktivitetsmønstre og deres sammenhæng, gælder de samme begrænsninger for beskrivelser af systemet på det konceptuelle eller begrebslige niveau:

⁶¹ Indenfor den klassiske skole har problemer med programmeringen af for eksempel ekspertsystemer givet konnektionismen et psykologisk forspring. Det problematiske aspekt er ikke bare udtømmende at programmere viden, men ligeledes udtømmende at programmere den kontekst hvorunder denne viden udfoldes. Kontekstproblemet viser sig også ifølge Hubert Dreyfus & Stuart E. Dreyfus i bogen *Mind Over Machine* bl.a. ved programmeringen af 'common sense' antagelser. Yderligere hævdes det i *Mind Over Machine*, at det er umuligt på forhånd at programmere hele den foranderlige verden med dens forskellige situationer og valg.

⁶² Bernsen & Ulbæk 1993, s.251.

...her vil beskrivelserne nødvendigvis blive... ufuldstændige, så de kun beskriver delaspekter af processerne...”⁶³

Vil man altså beskrive den menneskelige kognition gennem et konnektionistisk system, kan dette kun ske på et såkaldt subkonceptuelt niveau og ikke et konceptuelt niveau, eftersom disse kun vil blive ufuldstændige beskrivelser⁶⁴. Modsætningen til den klassiske skoles forsøg på en udtømmende beskrivelse af de kognitive processer er her indlysende, hvorfor de to retninger på flere måder står i skærende kontrast til hinanden.

Det er nu på sin plads at abstrahere nogle gennemgåede og modstillede træk ved de to bevidsthedsopfattelser repræsenteret i form af de to paradigmer.

2.5 Bevidsthedsopfattelser

Omdrejningspunktet for diskussionen mellem de to retninger kan siges at bestå i deres fortolkning af det mentales grundlæggende karakteristika. Forståelsen, og hvordan denne opstår, spiller her en vigtig rolle. Vi har i det foregående set, hvordan de to retninger søger at beskrive det mentale bl.a. gennem forskellige repræsentationsteorier. Hvad enten vi taler om symbolsk eller distribueret repræsentation, opviser repræsentationerne en forholden sig til verden. Gennem sproget udtrykker vi dette forhold, og således bliver sproget til det medium, hvorigennem erkendelsen og forståelsen udtrykkes. Det væsentlige i denne henseende er, at forståelsen, udtrykt gennem sproget, er givet for noget - et subjekt. Den klassiske opfattelse af intelligens antager, med overbevisningen om formaliserbarheden af det mentale, en slags regelfølge for sproget. Gennem systematiske manipulationer af symboler i overensstemmelse med en regel menes forståelsen at opstå⁶⁵. Searle har med sit ”Chinese Room” argument vist, at forståelse ikke er en nødvendig følge af regelmanipulation⁶⁶ - og netop denne indholdsløse, forståelsesforladte manipulation gør, at en computer ifølge Sear-

⁶³ Bernsen & Ulbæk 1999, s.254.

⁶⁴ Bernsen & Ulbæk 1999, s.255.

⁶⁵ Hvad enten vi her taler om scripts, frames eller miniuniverser (jf. Marvi Minsky) er essensen den samme – rammerne for den sproglige udfoldelse er programmeret. Man udtrykker sig altså i overensstemmelse med en regel. Her ses en tydelig parallel til Wittgensteins sprogspil.

⁶⁶ Searle, John R.: ”*Minds, Brains and Programs*” i Haugeland 1987 (b), s.282-306.

le aldrig være i stand til andet end at simulere intelligens⁶⁷. Searles argument angår umuligheden af den stærke kunstige intelligens og indbefatter således også konnektionismen. I et forsøg på at gendrive konnektionismen anvender Searle et tankeeksperiment, der i sin enkelhed går ud på at en person i et værelse ud fra en manual, som oversætter tegn fra kinesisk til manipulationsregler, kan manipulere med begreber. Spørgsmål på kinesisk kan således besvares af personen i værelset på baggrund af opslag i manualen. Det væsentlige ved dette tankeeksperiment er, at personen ikke har nogen mulighed for at forstå kinesisk. Hvor værelset udefra betraget synes at have forståelse og tale kinesisk forholder det sig anderledes set indefra. Værelset betraget som system har reelt ingen forståelse af kinesisk⁶⁸. Hans pointe med tankeeksperimentet er netop den, at selv om man simulerer den neurale struktur, er det kun den neurales formelle struktur, der simuleres – i Searles terminologi: skyggesiden af det mentale⁶⁹. Netop det forhold at kognition har et fundamentalt tilhørsforhold til den biologiske materie, søger Searle at vise gennem arbitrære tankeeksperiment - hvor for eksempel en gruppe af vandrer med deres indbyrdes vandstand logisk gengiver den formelle struktur. Gennem dette tankeeksperiment ser man ifølge Searle intuitivt, at det mentale forudsætter den rette materie - og kun den biologiske materie er af en sådan art⁷⁰.

For konnektionismen er det mentales egenart principielt ubeskrivelig og man er, som vi har set, overladt til en beskrivelse på et subkonceptuelt niveau. Dette kan måske forekomme tilfredsstillende i en søgen på at simulere intelligens set med udgangspunkt i intelligenskriteriet - den menneskelignende intelligente adfærd – men internalistisk set i forhold til at forklare det mentale, synes standpunktet at halte. Det konnektionistiske paradigme overlader os til en forståelse af det mentale, der på mange områder har paralleller til funktionalismen, som netop antager, at kognition er et produkt af de informationsbehandlingsfunktioner, et system fysisk implementerer⁷¹.

⁶⁷ Searle 1984, s.37. Searle forudsætter for dette standpunkt en biologisk naturalisme. Se Searle 1984, s.41.

⁶⁸ Jeg vender senere i opgaven tilbage til Searles Chinese Room.

⁶⁹ Searle, John R.: "*Minds, Brains and Programs*" i Haugeland 1987 (b), s.294-295.

⁷⁰ Searles biologiske naturalisme vil blive behandlet mere indgående senere i opgaven.

⁷¹ Bernsen & Ulbæk 1993, s.25.

Et generelt problematisk aspekt ved begge paradigmer viser sig ved, at de forlader sig på et eksternt kriterium for intelligens. Et system kan klassificeres som intelligent, hvis systemet opviser en intelligent adfærd. Samtidig postulerer kognitionsforskningen en form for definition af intelligens gennem begreber som perception, tænkning og handling – der i sidste instans menes at kunne fortolkes som informations- eller databehandlingssystemer⁷².

Problemet med denne fortolkning viser sig når teorierne skal redegøre for det mentale aspekt. Når man taler om intelligent adfærd, forudsættes forståelsen nærmest uproblematisk. Det, som kendetegner mennesket som et system af 'den rette slags', er netop den subjektive forståelse. De to retninger indenfor kunstig intelligens kredser om denne erkendelsesteoretiske problematik uden at komme en forklaring nærmere. Det forekommer altså som om, at der mellem definition og kriterium, immanent lægger en form for cirkularitet gemt, idet man ikke har adækvat viden om det naturlige system, men blot postulerer en teori om kognitionen ud fra adfærden. I dette postulat ligger gemt, at kognitionsforskningen opererer med et begreb om forståelse og bevidsthed. Disse begreber reduceres dog i sidste instans til et rent adfærdsmæssigt mønster. Betragter vi kognitionsforskningens definition på intelligens fremtræder det tydeligt:

”Den kombination af færdigheder i målrettet adfærd, som involverer hele det komplekse kognitive system, der kendetegner mennesket i højere grad end noget andet kendt system.”⁷³

Kognitionsforskningen forekommer at begå den fejl, at den opererer med et begreb om intelligens, som den selv anvender tvetydigt. I definitionen af intelligens refereres implicit til subjektive tilstande, systemets interne tilstand, som en forudsætning for en ekstern tilstand – den målrettede adfærd (intelligente adfærd). I redegørelsen for denne intelligente adfærd udelades det egentligt intelligente, det subjektive, for derefter at blive erstattet af et rent eksternt kriterium for intelligens - adfærden.

Inden vi afskriver muligheden for kunstig intelligens fuldstændig, er det dog på sin plads at undersøge det psykologiske intelligensbegreb nærmere for her at se, hvilken bevidsthedsopfattelse, der gemmer sig bag psykologiens begreb om intelligens.

⁷² Bernsen & Ulbæk 1993, s.27.

⁷³ Bernsen & Ulbæk 1993, s.11.

3 Det klassiske psykologiske intelligensbegreb

Vi har ovenfor set hvordan kognitionsforskningen opererer med begrebet om intelligens, og hvorledes dette udmøntes i de to undersøgte paradigmer. Indenfor psykologien hersker der ligeledes forskellige fortolkninger af begrebet intelligens. I det følgende skelner jeg mellem opfattelsen af intelligens som en statisk egenskab og som et dynamisk potentiale.

William Sterns *Intelligenskvotient* (IK), Charles Spearman's *Tofaktorteori* og L.L.

Thurstones *Flerfaktorteori* hører under den statiske opfattelse, mens Howard Gardners *Multiple Intelligences* (MI) og Daniel Golemans *Emotional Intelligence* tilhører den dynamiske intelligensopfattelse.

Inden en egentlig undersøgelse af Howard Gardners teori om mange intelligenser, forekommer det rimeligt først at betragte de statiske psykologiske teorier, som Gardners teori netop rummer et opgør med.

3.1 IK, g-faktor og flerfaktor

Denne klassiske opfattelse af intelligens tager udgangspunkt i psykologen Alfred Binets arbejde i Paris i begyndelsen af 1900-tallet. Den øgede opmærksomhed på elevernes færdigheder ledte Binet til udformelsen af en test, som kunne afgøre hvilke elever, der intellektuelt var svagt begavede⁷⁴. Binets test, som afgjorde forsøgspersonens *Intelligensalder*, kunne dog ikke anvendes til at sammenligne intelligensniveauet mellem børn på forskellige alderstrin. Tyskeren William Stern fandt derfor på at dividere intelligensalderen med levealderen. Herved fremkom intelligenskvotienten. Intelligenskvotienten er et generelt udtryk for intelligensniveauet og kan sammenlignes uanset forsøgspersonernes alder. Grundlaget for intelligenskvotienten er prøver, som hovedsageligt bygger på sproglige udtryk⁷⁵. Forsøgspersonens sproglige udvikling virker ikke alene ind på forståelsen af opgaven, men også på besvarelsen, hvorfor denne testmetode kan virke noget usikker. Dertil kommer, at disse tests ikke tager højde for intelligensens delfaktorer. Intelligensens delfak-

⁷⁴ Lehtovaara 1971, s.163.

⁷⁵ Howard Gardner kalder disse *paper-and-pencil tests*. Gardner 1993a, s.16.

torer findes ved at sammenligne scoren for forskellige tests – er denne korreleret, forudsætter testene samme slags præstationsevne.

Med Spearman's *tofaktorteori* deles intelligensen op i en såkaldt *g-faktor* (almen intelligensfaktor) og en *s-faktor* (specifik intelligensfaktor). Spearman antog, at *g-faktoren* var afhængig af hele hjernebarken, hvorimod *s-faktorerne* er lokaliserbare til bestemte delområder af hjernebarken. Spearman inddrager således en række fysiologiske faktorer i intelligensbegrebet. Spearman's *g-faktor* dækker over positive korrelationer og kan tages som et udtryk for, at den samme intelligens er indgået i forsøgspersonernes svar – deraf udtrykket generel faktor (*g-faktor*). Med Spearman's kobling af intelligensbegrebet og fysiologi opstår ideen om intelligens som en arvelig egenskab. G-faktoren antages at være et udtryk for en konstant i modsætning til de modificerbare s-faktorer, som kan ændres gennem påvirkning af miljø eller ved træning. For bedre at kunne tage højde for de individuelle forskelle i intelligensstrukturen udviklede Spearman senere *gruffefaktorteorien*. Her opererer Spearman med en korrelation mellem *s-faktorerne*, som derfor ikke længere betragtes som ensidigt autonome størrelser⁷⁶. Ideen om uafhængige intelligensfaktorer udvikles dog videre af L.L. Thurstone.

Thurstone's *flerfaktorteori* inddeler overordnet intelligensfaktorerne i tre grupper:

- 1) De sproglige eller verbale
- 2) Præstationer, baseret på *visuelle elementer; synsoplevelser*.
- 3) Præstationer, baseret på *numerisk-kvalitative* elementer. Disse bruges i forbindelse med numeriske, eller i anden henseende kvantitative symboler.

Disse grupper består yderligere af undergrupper i form af syv faktorer, der hver måler den individuelle intellektuelle præstationsevne. Af de ovenstående intelligensfaktorer fremgår det tydeligt, at *intelligenskvotienten* alene falder ind under den første gruppe. De individuelle forskelle i intelligensstrukturen/profilen kommer således bedre til udtryk gennem Thurstones *flerfaktorteori* end gennem *intelligenskvotienten*. Den samme intelligenskvotient kan således dække over forskellige intelligensprofiler, der indfanges med Thurstones teori. Til grund for Spearman og Thurstones teorier ligger matematiske metoder. Denne anvendelse af matematiske metoder medfører, at:

⁷⁶ Lehtovaara 1971, s.162ff.

”...given the same set of data, it is possible, using one set of factor-analytic procedures, to come up with a picture that supports the idea of a ‘g’ factor, using another equally valid method of statistical analysis, it is possible to support the notion of a family of relatively discrete mental abilities.”⁷⁷

For en teori om intelligensen må det siges at være centralt, om man forstår intelligens som én grundlæggende egenskab, der transcenderer forskellige begavelser, eller om man forstår intelligens som forskellige uafhængige egenskaber. Ikke mindst er det fatalt for Spearmans teori, idet han lokaliserer den generelle intelligens til hjernebarken. Spearmans intelligensbegreb har således et ontologisk fundament, som modsiges af Thurstones flerfaktorbegreb. Centralt for disse teorier er, at de har været meget nyttige til at forudsige elevens muligheder indenfor skolen, og korrigere skoleforløb. Dette skyldes i høj grad, at intelligens i de fremstillede teorier sidestilles med teoretisk begavelse⁷⁸. Og langt hen ad vejen har undervisningen i skolerne været lagt an på disse teoretiske færdigheder⁷⁹, som intelligenstestene netop måler⁸⁰. Der er her en tydelig forskel fra kognitionsforskningens begreb om intelligens. Kognitionsforskningen relaterer intelligensbegrebet til hele det komplekse kognitive system, hvorimod ovenstående psykologiske teorier relaterer intelligensbegrebet til teoretiske færdigheder.

Begge intelligensbegreber forudsætter, for overhovedet at tale om intelligens og begavelse, tanken. Intelligenstestene kan tages som et udtryk for et menneskes tankemæssige egenskaber. De tankemæssige egenskaber defineres med psykologien således, at:

⁷⁷ Gardner 1993a, s.16.

⁷⁸ Lehtovaara 1971, s.174.

⁷⁹ Her tænker jeg på de verbale og matematisk-logiske færdigheder.

⁸⁰ Prøver man i stedet at korrelere intelligenstests med post-skoleforløbet, tjener disse redskaber heller ikke som de bedste ledetråde. Det er alt andet givet, at en intelligent elev (intelligent i forstanden ovenfor) vil begå sig godt uden for skolen. Udviklingen indenfor undervisning i folkeskoler samt indenfor psykologien går dog i retning af ligeledes at inddrage den praktiske begavelse i en egentlig vurdering af intelligensen. Gardner 1993b, s.14.

”Med tænkning i egentlig forstand mener vi kun sådanne virksomheder, der ikke direkte er bundet til perceptionerne, eller som i hvert fald principielt kan foregå ved hjælp af begreber alene.”⁸¹

Tænkningens abstrakte natur er altså undersøgelsesområdet for intelligensteoriene. Tænkningen er relateret til konteksten efter som, at den direkte eller indirekte står i forbindelse med perceptionsvirksomheden og hukommelsen⁸². Den psykologiske definition af tankens begrebslige egenart ligner kognitionsforskningens antagelse om, at tankens struktur er på samme måde som sprogets. Kognitionsforskningen betoner dog i mindre grad de kontekstuelle faktorer, som psykologien forudsætter for tankevirksomheden. Tankevirksomheden foregår ifølge psykologien på begreber, men forudsætter blandt andet (1) reproduktion af tidligere erfaringer, (2) pålidelige perceptioner samt (3) udnyttelse af kendsgerninger, der er samlet ved enten sammenligninger, analyse, kontrol eller kombination. Disse forudsætninger transcenderer så at sige tankevirksomheden forstået på den måde, at det er kontekstuel afledte faktorer, som virker ind på tankens mulighed. Uden tidligere erfaringer og øjeblikkelige perceptioner at trække på er tanken indholdsløs⁸³.

Den centrale forskel mellem kognitionsforskningen og psykologien er, at kognitionsforskningen har påtaget sig den opgave at beskrive hele vort kognitive apparat, hvorimod psykologien med det traditionelle intelligensbegreb kun beskriver den del, som har at gøre med den teoretiske begavelse⁸⁴.

Distinktionen mellem de to intelligensopfattelser kan siges at ligge i måden, man søger at redegøre for tanken på. Psykologisk søger man en belysning af tanken ved en vurdering af denne udtrykt gennem testresultater og svar. På denne måde ligger fokus hos psykologien på den objektive side af tanken, hvilket samtidig muliggør en kvantitativ vurdering af denne. Indenfor kognitionsforskningen er fokus både på tankens kvantitative og kvalitative egenskaber. Kognitionsforskningens projekt kræver en teori for bevidstheden, der så at sige samler tankens objektive og subjektive egenskaber i én bevidsthedsopfattelse.

⁸¹ Lehtovaara 1971, s.145.

⁸² Lehtovaara 1971, s.151.

⁸³ Dette viser sig også indenfor kognitionsforskningen i og med at den kontekstuelle adgang, betragtes som væsentlig for indlæring m.m.

⁸⁴ Lehtovaara 1971, s.174.

Skal intelligensbegrebet i psykologisk forstand altså bringes i forbindelse med kognitionsforskningens intelligensbegreb, må det psykologiske intelligensbegreb udvides til at omfatte hele bevidstheden, såvel de praktiske, teoretiske som specielle begavelser. Det psykologiske intelligensbegreb skal således funderes i en teori om hele den menneskelige bevidstheds natur og struktur. En udvidelse af det psykologiske genstandsområde synes altså for nærværende projekt på sin plads.

Det er værd at bemærke, at psykologiens intelligensbegreb ikke udsiger noget om de menneskelige empatiske egenskaber eller forståelsens dybereliggende karakter. Grundlaget for intelligensbegrebet ligger i tankens mulighed, hvorom ikke siges andet end at denne består i begrebslig manipulation (symbolmanipulation) og at denne kan måles. Af denne begrebslige manipulation undersøger psykologien symbolmanipulationens udtryksside i form af den teoretiske begavelse⁸⁵. Skal intelligensbegrebet drages i relation til kognitionsforskningens intelligensbegreb, må den således også redegøre for symbolmanipulationens indholdsside. En udvidelse af den klassiske intelligensopfattelse synes altså uomgængelig, hvis intelligensbegrebet på nogen måde skal kunne appliceres på computere⁸⁶. Ydermere må man finde andre målemetoder, da en computer ikke kan siges at have en kronologisk alder i menneskelig forstand og dens intelligens derfor heller ikke vil kunne måles på traditionel vis med William Sterns definition af intelligenskvotienten⁸⁷.

Historisk betragtet går udviklingen indenfor psykologien da også hen imod en inddragelse af flere facetter af den menneskelige kognition i en bestemmelse af intelligensen.

En af de psykologiske teorier, der søger at udvide begrebet om intelligens og inddrage flere aspekter af det mentale system, er Howard Gardners teori om Multiple Intelligence.

⁸⁵ Med symbolmanipulationens udtryksside, mener jeg netop det forhold, at man ikke tager stilling til de erkendelsesteoretiske problemstillinger, som er forbundet med tanken. I stedet reduceres undersøgelsen af tanken udelukkende til en undersøgelse og vurdering af en persons teoretiske færdigheder, som de kommer til udtryk gennem empiriske tests.

⁸⁶ Forudsætningen for at måle intelligens er at 'emnet' besidder intelligens, som en egenskab. Denne forudsætning er ikke givet med computere, hvorfor intelligensbegrebets dybere betydning og struktur må fastlægges.

⁸⁷ Intelligenskvotienten fås ved at intelligensalderen divideres med levealderen og multipliceres med hundrede.

4 Multiple intelligence

Howard Gardner tager med sit begreb om intelligens afsæt i den klassiske intelligensopfattelse af intelligens således forstået, at han ligeledes vil nå frem til en karakteristik af et menneskes færdigheder, dets begavelse. Men her hører ligheden med den traditionelle opfattelse også op. Howard Gardner udvider begrebet om intelligens til også at omhandle kapaciteter, som hidtil har ligget udenfor intelligensbegrebets sfære. Howard Gardner definerer intelligens som:

”a biopsychological potential to process information that can be activated in a cultural setting to solve problems or create products that are of value in a culture.”⁸⁸

Intelligensbegrebet involverer altså de menneskelige problemløsnings- og skaberevners biologiske fundament, som vurderes i relation til det omgivende samfunds værdier. Udvidelsen af intelligensbegrebet således, at det ligeledes omfatter de skabende menneskelige aktiviteter, er væsentlig i forhold til den traditionelle opfattelse, der entydigt iagttog problemløsningsadfærden. Intelligensbegrebet udvides med Gardners definition også til et kulturrelativt fænomen sådan forstået, at forskellige kulturer værdsætter forskellige områder af begavelsen. I modsætning til den klassiske intelligensopfattelse tager Multiple Intelligence (MI) altså udgangspunkt i, at man i en beskrivelse af intelligensen ikke alene skal anvende tests og korrelationer mellem tests i en vurdering af intelligensen, men i stedet se på andre former, hvorigennem intelligensen udtrykkes. Disse former og deres betydning vil være relative i forhold til den omgivende kulturelle kontekst⁸⁹.

Samtidig med de kulturelle indvirkninger medfører MI's intelligensbegreb ligeledes et skift fra en ekstern og adfærdsmæssig vurdering af menneskets bevidsthed, forstået som symbolbehandlingssystem, til en direkte fokus på dette systems virkemåde. Det biologiske udgangspunkt for symbolsystemet er centralt for en undersøgelse af den menneskelige intelligens, idet at:

⁸⁸ Gardner 1999, s.33-34. Definitionen er en udvidelse i forhold til Gardners tidligere definition, hvor intelligens blev defineret som: ”*the ability to solve problems, or to fashion products, that are valued in one or more cultural community settings*”(Gardner 1993b, s.7.).

⁸⁹ Gardner 1993b, s.7.

”...what is distinctive about human cognition and information processing involves the deployment of these various symbol systems.”⁹⁰

Det forhold, at vores bevidsthed opererer gennem forskellige symbolsystemer, adskiller os fra andre organismer og er ifølge Gardner det egentlige særkende for vores bevidsthed, som specielt kognitionsforskningen og neurovidenskaben har beskæftiget sig med. Gardners projekt er med hans udvidelse af intelligensbegrebet at nå frem til den rette beskrivelse af intelligensen⁹¹. Ved at inddrage kontekstuelle, samt neurologiske faktorer i beskrivelsen af intelligensbegrebet, sammenholdt med opfattelsen af intelligens som et dynamisk begreb⁹², bliver intelligensbegrebet under MI til en beskrivelse af et biopsykologisk potentiale. Intelligensen er altså et potentiale, som aktualiseres i samspil mellem kontekstuelle såvel som biologiske faktorer. Det er væsentligt at fastholde at MI i forhold til klassisk intelligensopfattelse betragter det biologiske fundament for de menneskelige problemløsningsevner⁹³. Hvor den klassiske opfattelse traditionelt definerede intelligens i praksis som evnen til problemløsning, definerer Gardner intelligens som det, der medfører evnen til at løse problemer og/eller skabe produkter⁹⁴. Intelligens er altså ikke længere selve problemløsningsevnen, men det, som ligger bagved og muliggør denne evne. Disse bagvedliggende biologiske egenskaber kobles til de kontekstuelle faktorer ved at:

”...the biological proclivity to participate in a particular form of problem solving must also be coupled with the cultural nurturing of that domain. For example, language, a universal skill, may manifest itself particularly as writing in one culture, as oratory in another culture, and as the secret language of anagrams in a third.”⁹⁵

⁹⁰ Gardner 1993a, s.25.

⁹¹ Gardner 1993b, s.7.

⁹² Det dynamiske aspekt skal ses deri, at Gardner i modsætning til den traditionelle intelligensopfattelse betragter intelligens som flere faktorer, der kan udvikles gennem livet.

⁹³ Forskellen fra den traditionelle intelligensopfattelse er at Gardner inddrager neurovidenskabelige resultater, således at intelligensbegrebet også relateres til menneskets biologiske konstitution.

⁹⁴ Gardner 1993b, s.15.

⁹⁵ Gardner 1993b, s.16.

De evner, der anses for menneskeligt universelle, eksempelvis sproget, kommer altså til udtryk på forskellig måde gennem forskellige kulturer. MI søger *eo ipso* at opstille en teori, der redegør for intelligensbegrebet, så det kan anvendes på tværs af kulturelle kontekster. Affiniteten mellem biologi og bevidsthed er værd at undersøge nærmere. I det følgende vil jeg opstille Gardners terminologi og anskue på hvilken måde, begreberne om bevidsthed og hjerne spiller ind på intelligensbegrebet.

4.1 Intelligensbegrebets terminologi og analyseramme

En dybere forståelse af Gardners begreb om intelligens og bevidsthedens relation hertil fordrer, at vi ser nærmere på terminologien, som anvendes i teorien om MI. Terminologien falder i form af en analyseramme og de her forskellige tilknyttede begreber.

Howard Gardner tager som udgangspunkt for en analyse af det kognitive afsæt i, at enhver kognitiv handling er ledsaget af en agent, som udfører en handling, eller et sæt af handlinger indenfor et bestemt domæne⁹⁶. Disse handlinger er potentielt åbne for vurdering forstået således, at handlingerne kan vurderes af en sagkyndig⁹⁷. Det involverede analyseområde, intellektet, kan opdeles i tre analysefelter/perspektiver, et:

- Biopsykologisk perspektiv
- Domæneperspektiv
- Feltorienteret perspektiv.

For det *biopsykologiske* perspektiv gælder, at man her alene undersøger agenten og dennes kapaciteter, tilbøjeligheder, normer og målsætning. Dette analyseperspektiv involverer således en undersøgelse af såvel det genetiske og neurologiske grundlag, hvorpå adfærden er funderet og ligeledes en vurdering af personens psykologiske karakteristika⁹⁸.

Domæneperspektivet afdækker og undersøger på hvilken måde en agents handlinger er blevet udført i relation til det givne samfundsdomæne/disciplin. Med domæne menes:

”...an organized set of activities within a culture, one typically characterized by a specific symbol system and its attendant operations. Any cultural activity in which individuals participate on more than a casual basis, and in which de-

⁹⁶ De her anvendte begreber vil blive uddybet nedenfor.

⁹⁷ Gardner 1993b, s.50.

⁹⁸ Gardner 1993b, s.50.

grees of expertise can be identified and nurtured, should be considered a domain.”⁹⁹

Et domæne er følgelig karakteriseret ved et specifikt symbolsystem, som deltagere er underkastet givet deres deltagelse i specifikke kulturelle aktiviteter¹⁰⁰. Tidligere har filosoffer og eksperter indenfor de respektive domæner stået for disse undersøgelser, men efterhånden har computere overtaget denne funktion med fremhævelse af analyser af de strukturelle og processuelle egenskaber ved en opgave eller aktivitet¹⁰¹.

Det *feltorienterede* perspektiv vurderer og evaluerer handlingerne eller artefakterne, som udføres indenfor et givent domæne. Vurderingerne fremsættes her af personer, der regnes for kyndige indenfor dette område (*members of the field*). Sociologi og socialpsykologi er eksempler på discipliner, som tilhører dette område¹⁰².

Gardners opdeling i disse tre perspektiver tjener til at belyse den matrix, som intelligensbegrebet kan siges at karakteriseres ved. Med opdelingen af analyseområdet i disse tre perspektiver får intelligensbegrebet helt andre konnotationer end den traditionelle anvendelse. Gardners intelligensbegreb gør op med det uniforme intelligensbegreb¹⁰³. I stedet mener han, at vi taler om en mangetydig størrelse - nemlig *intelligenser*¹⁰⁴. Denne mangetydige størrelse lader sig ikke afdække alene gennem de såkaldte papir og blyant tests. Det biopsykologiske potentiale ses værende ækvivalent med intelligenser således forstået, at:

”...intelligences are not things that can be seen or counted. Instead, they are potentials – presumably, neural ones – that will or will not be activated, depending upon the values of a particular culture, the opportunities available in

⁹⁹ Gardner 1999, s.82.

¹⁰⁰ Om domænebegrebet skal forstås i retning af den sene Wittgensteins sprogspilbegreb herom er Gardner tavs.

¹⁰¹ Gardner 1993b, s.50.

¹⁰² Gardner 1993b, s.50.

¹⁰³ Med det uniforme intelligensbegreb mener Gardner det forhold at den traditionelle opfattelse, herunder også Thurstones flerfaktorteori, til stadighed har begrænset undersøgelsen af intelligens til et uniformt udtryk gennem verbale test eller papir-og-blyant tests. Gardner 1999, s.35.

¹⁰⁴ Gardner 1993b, s.6.

that culture, and the personal decisions made by individuals and/or their families, schoolteachers and others.”¹⁰⁵

Hvor man igennem den traditionelle opfattelse har et kvantificerbart udtryk for intelligensen gennem intelligenskvotienten, er intelligens med *teorien om mange intelligenser* ikke en kvantificerbar størrelse. Den rene intelligens, i form af det biopsykologiske potentiale, kan ikke måles eller vejes¹⁰⁶. Gennem en agents handlen udtrykkes flere intelligenser samtidig og kun i særtilfælde, hos blandt andre idiot savants, vidunderbørn og autister, kan man observere en ren form for intelligens komme til udtryk¹⁰⁷.

Af ovenstående perspektiver fremgår det, at intelligenser figurerer som distribueret og kontekstualiseret fænomen således, at:

”intelligences are always expressed in the context of specific tasks, domains, and disciplines.”¹⁰⁸

Intelligenserne manifesterer sig gennem en kontekstuel aktualisering i forhold til at løse en bestemt opgave eller en anden form for handlen fra agentens side. En beskrivelse af en persons intellekt sker således gennem en vurdering af agentens handlinger, som de kommer til udtryk indenfor flere forskellige domæner¹⁰⁹. Gardner gør følgelig op med antagelsen om, at intelligens kan betragtes som en enkelt egenskab, hvilket han mener opretholdes gennem sproglige konnotationer, og forkaster dermed opfattelsen af, at vi kan betegne personer som værende enten ’intelligente’ eller ’dumme’¹¹⁰.

Vi vil nu se nærmere på Gardners begreb om intelligenser og hvad dette begreb egentlig dækker over.

¹⁰⁵ Gardner 1999, s.34.

¹⁰⁶ Gardner 1993a, s.xx.

¹⁰⁷ Gardner 1999, s.39.

¹⁰⁸ Gardner 1993a, s.xx.

¹⁰⁹ Viser en person mere end almindelige evner indenfor et domæne betegnes han som lovende (*”gifted”*). En person, der er ekstremt lovende betegner man vidunderbarn (*prodigious*). Gardners begreber *ekspert*, *kreativitet* og *geni* vil jeg for nærværende ikke uddybe. Gardner 1993b, s.51.

¹¹⁰ Gardner 1999, s.34.

4.2 De mange intelligenser

Som udgangspunkt for en undersøgelse af intellektet og hvad, der kan regnes som intelligenser, opstiller Gardner en række kriterier. Gardner opererer med følgende otte af hinanden uafhængige kriterier på en intelligens¹¹¹:

1. Isolation ved hjerneskade¹¹²
2. Den evolutionære historie og evolutionens plausibilitet
3. En identificerbar kerneoperation eller sæt af operationer
4. Muligheden for at kunne kodes i et symbolsystem
5. En distinkt udviklingshistorie, sammen med et definerbart sæt af ekspert "end-state" handlinger.
6. Eksistensen af idiot savants, vidunderbørn og andre exceptionelle mennesker
7. Støtte fra eksperimentelle psykologiske opgaver
8. Støtte fra psykometriske opdagelser

I sit forsøg på at finde frem til antallet af intelligenser anvendes disse ovenstående kriterier som rettesnor. Er kriterierne i en rimelig grad imødekommet for et givent område, hævder Gardner, at vi kan tale om en intelligens¹¹³. Gardner giver ikke en egentlig uddybelse af hvornår disse kriterier er overholdt og man derfor kan tale om en rimelig grad af imødekommenhed.

Med udgangspunkt i disse kriterier underkaster Howard Gardner et omfangsrigt studiemateriale en såkaldt subjektiv faktoranalyse, eftersom materialets beskaffenhed ikke tillod en statistisk faktoranalyse. Ud af denne analyse fremkom der oprindeligt syv intelligenser, som med *Intelligence Reframed* blev udvidet til otte intelligenser. Ingen af intelligenserne er forfordelt frem for andre, men i stedet betragter Gardner alle intelligenser som rangerende på samme niveau¹¹⁴. De otte intelligenser er som følger¹¹⁵:

¹¹¹ Gardner 1999, s.36. Af pladshensyn har jeg ikke uddybet de forskellige kriterier yderligere.

¹¹² Gardner mener at man kan separere en intelligens ved at se på hjerneskadede. Det er væsentligt at denne separation udelukkende opfattes som et anvisende værktøj til stadfæstelse af en intelligens i begrebslig forstand og ikke som en egentlig lokalisering af denne. Problemer omkring sådanne lokaliseringer af kognitive funktioner (læsionsmetoden) uddybes i Gade 1997, s.66ff.

¹¹³ Gardner 1999, s.35.

¹¹⁴ Gardner 1993b, s.8.

- *Sproglig* intelligens betegner følsomheden overfor talt og skreven sprog, evnen til at lære sprog, samt evnen til at bruge sprog til at opnå bestemte mål.
- *Logisk-matematisk* intelligens betegner evnen til at analysere problemer logisk, udføre matematiske operationer og undersøge emner videnskabeligt.
- *Rumlig* intelligens betegner evnen til at genkende og manipulere mønstrene i det vide rum såvel som mønstrene i mere afgrænsede rum.
- *Musisk* intelligens betegner evner i udførelsen, kompositionen og anerkendelsen af musiske mønstre.
- *Kropslig-kinæstetisk* intelligens betegner den potentielle evne til at bruge ens krop eller dele deraf i løsningen af en opgave eller skabelsen af et produkt.
- *Interpersonel* intelligens betegner en persons evne til at forstå intentioner, motivationer og begær hos andre mennesker og samtidig at arbejde effektivt med andre.
- *Intrapersonel* intelligens betegner evnen til at forstå sig selv – at have et retvisende billede af sig selv deriblandt ens begær, frygt og kapaciteter – og at bruge denne information på en effektiv måde til at regulere ens eget liv.
- *Naturalistisk* intelligens betegner evnen til at kunne genkende og skelne mellem forskellige arter – inden for såvel flora og fauna – i det omgivende miljø.

Gardner udelukker ikke, at der i fremtiden vil dukke nye intelligenser op. Det er selvfølgelig ikke med nogen logisk eller videnskabelig nødvendighed, at han opererer med lige netop otte intelligenser frem for et andet antal¹¹⁶. En beskrivelse af intellektet ud fra disse otte intelligenser skaber et mere nuanceret syn på den menneskelige erkendelse. Hvor den traditionelle intelligensopfattelse havde svært ved at redegøre for forskellene i psykologisk profil, synes teorien om mange intelligenser at forudsætte netop en sådan forskellighed. Forskelligheden kommer til udtryk ved at:

“We all possess the same ensemble of intelligences – in one sense, they represent our species’ intellectual heritage – but we do not exhibit equal strengths or similar profiles.”¹¹⁷

¹¹⁵ Gardner 1999, s.41ff.

¹¹⁶ Gardner 1999, s.47.

¹¹⁷ Gardner 1999, s.166.

Vi besidder alle det samme antal intelligenser, mens forskelle i den psykologiske profil viser sig gennem forskelle i styrken af de enkelte intelligenser. I og med at teorien ikke har et ensidigt, men netop flersidigt syn på intelligensen, træder disse forskelle i den psykologiske profil tydeligere frem gennem en beskrivelse af intellektet ud fra teorien om mange intelligenser.

Anvender man Gardners egen terminologi på hans beskrivelse af intelligenser, kan man sige, at intelligenserne som beskrevet ovenfor er betragtet ud fra et domæne perspektiv. Skal vi se nærmere på det biologiske tilhørsforhold, hvad Gardner understreger som en af teoriens mange fordele til forskel for det klassiske perspektiv, er vi nødt til at gøre det ud fra et biopsykologisk perspektiv. Det er også her, at forholdet mellem hjerne og bevidsthed træder tydeligere frem, hvorfor en undersøgelse af dette område synes uomgængelig.

4.3 Det biopsykologiske potentiale

Som vi så tidligere, sidestiller Gardner begrebet om intelligens med det biopsykologiske potentiale. Dette forhold gør, at teorien skal ses i lyset af den biologiske oprindelse til hver problemløsningsfærdighed¹¹⁸. Det biologiske fundament taget i betragtning gør, at kun færdigheder, som har universel betydning for menneskeslægten tages i betragtning, hvorfor ovenstående intelligenser har et universelt præg og dermed omfatter hele menneskeslægten¹¹⁹. Når en agent løser en opgave eller handler på anden vis, optræder flere af intelligenserne samtidig – intelligenserne samarbejder. Dette kan illustreres med, at en person i besvarelsen af en opgave, hvor eksempelvis den logisk-matematiske intelligens testes, vælger at ledsage bevarelsen med en uddybende og måske perspektiverende kommentar og således anvender den sproglige intelligens. Om denne perspektivering falder uden for intelligensbedømmelsen, er egentlig uinteressant. Det interessante er for så vidt, hvis man skal

¹¹⁸ Gardner 1993b, s.15-16.

¹¹⁹ Evolutionen taget i betragtning ville være mere retvisende om man her undgik de tidlige konnotationer som begrebet menneskeslægten rummer og i stedet anvendte et mere abstrakt begreb som menneskeheden. Netop teoriens biologiske udgangspunkt gør, at det bør anses som relativt til et givent tidspunkt i menneskets udviklingshistorie, hvornår en given intelligens er opstået. Det synes altså som om Gardner mener at samtlige intelligenser er opstået samtidig, hvad der biologisk set ikke entydigt synes at være belæg for eftersom neo-cortex har udviklet sig med tiden.

følge Gardner, at vi trækker på forskellige intelligenser igennem vores problemløsende og skabende adfærd.

4.3.1 Subintelligenser

Hver enkelt intelligens kan yderligere opdeles i subintelligenser¹²⁰. Subintelligenserne repræsenterer kernefunktioner, som er fundamentale for den overordnede intelligens¹²¹. Kerneoperationerne er det egentlige biologiske fundament for intelligenserne, hvor:

”These capacities are likely to be mediated by specific neural mechanisms and triggered by relevant internal or external types of information.”¹²²

Kernefunktionerne kan altså lokaliseres til bestemte neurale netværk i hjernen. Forholdet mellem disse kernefunktioner og intelligenserne består følgelig i, at:

“Fundamentally, any intelligence refers to a biopsychological potential of our species to process certain kinds of information in certain ways. As such, it clearly involves processes that are carried out by dedicated neural networks”¹²³

Givet at hver form for intelligens kan underinddeles i kernefunktioner, har hver enkelt intelligens altså et biologisk fundament. Omvendt forudsætter Gardner, at disse neurale funktioner kan sammenfattes under ét i en slags metafunktion – den respektive intelligensform¹²⁴. Intelligensernes biologiske tilhørsforhold er altså betinget af, at man medgiver, at de respektive subintelligenser med rimelighed kan henføres til én bestemt intelligens. Om kernefunktionerne er uafhængige eller interdependente, betragter Gardner ikke som væsentligt eftersom han finder det sandsynligt, at de optræder sammen i en metafunktion (den specifikke intelligensstype) og dermed også uproblematisk kan grupperes som sådan¹²⁵.

¹²⁰ Gardner 1993b, s.9 og Gardner 1999, s.37.

¹²¹ Gardner mener eksempelvis, at den sproglige intelligens består af kernefunktioner såsom fonetisk diskrimination, beherskelse af syntaks og tilegnelse af ords betydninger. Gardner 1999, s.37.

¹²² Gardner 1999, s.37.

¹²³ Gardner 1999, s.94.

¹²⁴ Gardner 1999, s.37.

¹²⁵ Gardner 1999, s.37.

Tager vi afsæt i Gardners intelligensbegreb, fremkommer der altså et modulært syn på hjernen, hvor:

”A mind consists of a number of fairly specific and fairly independent computational mechanisms. Common properties and common regions may well exist, but they are surely not all of the story and may well not even be its most pertinent...part.”¹²⁶

Gardner mener følgelig, at hjernen er opdelt i forskellige funktionelle afdelinger, hvortil intelligensbegrebet relaterer sig. De forskellige intelligenser er qua deres kernefunktioner altså funderet i biologien.

At kernefunktionernes lokalisbarhed skulle medføre, at intelligenserne på samme måde er lokaliserbare, synes ikke at følge med nødvendighed. Alligevel hævder Gardner, at:

“Until it becomes possible to designate neural circuitry as representing one or another intelligence in action, we cannot know for sure which intelligence or intelligences are being invoked on a specific occasion.”¹²⁷

og

“...we now understand that the human mind, reflecting the structure of the brain, is composed of many separate modules or faculties.”¹²⁸

Ved at sidestille hjernen med bevidstheden begår Gardner den fejl at sidestille intelligenserne med subintelligenserne – men subintelligenserne tegner ikke det komplette billede af intelligenserne, i givet fald synes det at være irrelevant at operere med otte kriterier.

Gardner synes ydermere at begå en cirkulær antagelse idet, han opstiller kernefunktionerne som et kriterium på intelligens (det tredje kriterium) for herefter i selv samme kriterium at forudsætte denne intelligens.

Forholdet mellem kernefunktionerne og de respektive intelligenser afspejler sig i en dikotomi omkring intelligensernes status, som Gardner ikke entydigt undslipper. Gardner siger om intelligensernes status at:

¹²⁶ Gardner 1993a, s.56.

¹²⁷ Gardner 1999, s.95.

¹²⁸ Gardner 1999, s.3.

“...they exist not as psychically verifiable entities but only as potentially useful scientific constructs.”¹²⁹

og

“An intelligence is a new kind of construct, one that draws on biological and psychological potentials and capacities. It should not be confused with domains or disciplines, which are socially constructed human endeavors.”¹³⁰

Gardners begreb om intelligens er altså ikke andet end et konstrukt. Intelligensbegrebet som sådan har ikke nogen egentligt ontologisk status, men er blot ment som et redskab, der er baseret på biologiske og psykologiske potentialer og egenskaber. Howard Gardners skelnen mellem konstrukt og socialt konstrukt bevirker, at han synes at operere med en slags objektivitetsskala. På denne skala figurerer intelligensbegrebet som et videnskabeligt konstrukt over for de sociale menneskelige aktiviteter, der beskrives gennem konstrukter som domæner og discipliner. Når man på denne måde tager udgangspunkt i en form for konstruktivisme kan det forekomme problematisk efterfølgende at operere med et sådant objektivistisk skel. Denne konstruktivisme medfører ydermere, at:

“...it becomes necessary to say, once and for all, that there is not, and there can never be, a single irrefutable and universally accepted list of human intelligences.”¹³¹

Det ligger så at sige immanent i intelligensbegrebet i kraft af, at det er et konstrukt, at man derfor ikke kan udlede en finit liste over universelle menneskelige intelligenser. Men det er vel egentlig det projekt, som Gardner har sat sig for alt efter, som han siger:

”The science in this enterprise, to the extent that it exists, involves trying to discover the *right* description of the intelligences.”¹³²

Denne formulering tilskriver nærmere intelligenserne en ontologisk status, hvilket der er i direkte modsætning til han konstruktivistiske syn. At Gardner går ud fra en form for onto-

¹²⁹ Gardner 1993a, s.70.

¹³⁰ Gardner 1999, s.82.

¹³¹ Gardner 1993a, s.59.

¹³² Gardner 1993b, s.7.

logisk status, fremkommer endnu mere tydeligt i diskussionen om nye tilkomne intelligenser. Under diskussionen om en mulig niende intelligens, den spirituelle intelligens, fremhæver Gardner som en af grundene til ikke at betragte denne som en egentlig intelligens, at:

“Many of us do not recognize the Spirit as we recognize the mind and the body, and many of us do not grant the same ontological status to the transcendent or the spiritual as we do to, say, the mathematical or the musical.”¹³³

Her fremgår det tydeligt at grunden til at forkaste den spirituelle sfære som en intelligens er, at den ikke nyder samme ontologiske status som eksempelvis den matematiske eller den musiske intelligens. Hvordan Gardner kan opretholde et begreb om intelligenserne som et konstrukt og så samtidig kategorisere dem som havende ontologisk status er mig uvist, men uproblematisk er det i hvert fald ikke.

Tager man Gardner til indtægt for et konstruktivistisk synspunkt, hvad der kan synes mere end rimeligt på baggrund af ovenstående, bliver det tilmed endnu mere vanskeligt at se hvordan det biopsykologiske perspektiv skal kunne adskilles fra domæne- og feltperspektivet. I bund og grund må det biologiske og psykologiske område også betragtes som et konstrueret område og dermed som et domæne og et felt, hvorfor sondringen i tre perspektiver synes at smuldre.

Lad os forlade denne diskussion og i stedet vælge at anskue Gardners teori fra oven for hermed se på hans sigte med teorien, samt hans antagelser omkring bevidstheden.

4.4 Bevidsthedsopfattelsen

Som vi har set det, ønsker Gardner med sin teori at beskrive det menneskelige intellekt. Selv om Gardner medgiver, at intelligensbegrebet er et konstrukt, er hans foretagende at afdække de universelle menneskelige egenskaber, som kan henføres til ethvert intellekt. Teorien som helhed skal ifølge Gardner opfattes som:

”...an account of human cognition in its fullness – I put forth the intelligences as a new definition of human nature, cognitively speaking.”¹³⁴

¹³³ Gardner 1999, s.53.

¹³⁴ Gardner 1999, s.44.

og

“MI theory is about the intellect, the human mind in its cognitive aspects”¹³⁵

Med teorien om mange intelligenser søger Gardner følgelig at beskrive den menneskelige erkendelse i sin helhed. Gardner beskuer den menneskelige erkendelse i forening med evolutionen således, at:

“...cognition has evolved so that we can make sense of human beings (self and others) who have and experience emotions. Emotions do accompany cognition, and they may well prove more salient under certain circumstances.”¹³⁶

Ud fra denne opfattelse af den menneskelige kognition synes det eo ipso rimeligt ikke blot at antage, at han med sin teori må forventes at redegøre for hvorledes mening opstår, men ligeledes må redegøre for de subjektive aspekter som akkompagnerer mening. Gardner tager dog en omvej for netop at undgå denne erkendelsesteoretiske diskussion. Som udgangspunkt betragter han intelligens som et potentiale, hvorved han forstår den menneskelige hjerne, og dermed faktisk også bevidsthed, som bestående af interdependente funktionelle afsnit, der hver især behandler en specifik form for information. Gardner betragter det derfor således, at:

”...the mind is better viewed in a vertical way, as a set of faculties geared to particular contents in the external world and in human phenomenal experience.”¹³⁷

Denne funktionsopdeling og modulære syn på bevidstheden medfører, at teorien om mange intelligenser ikke ser intellektet givet via en singular bagvedliggende instans, en såkaldt central processorenhed, hvortil meningen opviser sig. At Gardner rent terminologisk anvender begrebet adfærd (behavior) som baggrund for diskussionen om eksistensen af en enkelt udøvende instans gør ikke, at han kan undgå at forholde sig til adfærdens subjektive

¹³⁵ Gardner 1999, s.89.

¹³⁶ Gardner 1999, s.206.

¹³⁷ Gardner 1999, s.104. Gardner skelner her ikke mellem hjerne og bevidsthed. At bevidstheden er afledt af eller i det mindste relaterer sig til hjernen er antaget hos bl.a. John Searle i *Minds, Brains and Science* s.39. Distinktion er vigtig at fastholde især når man, som Gardner, har sat sig for at redegøre for menneskets erkendelse i sin helhed.

aspekt. Netop omkring eksistensen af en singulær udøvende instans og dennes implikationer for en teori mener Gardner, at:

“A theory that does not posit an executive function has certain advantages over one that does. It is simpler, and it avoids the specter of infinite regression – the “homunculus” question of who or what is in charge of the executive. Nor does effective work require an executive”¹³⁸

Gardner er altså opmærksom på den uendelige regres, som man havner i ved at indføre en sådan singulær udøvende instans – en homunculus. At teorien er mere simpel, og dermed skulle optræde som valid af selv samme grund, er ikke på nogen måde bevist. Ockhams rasekniv synes her nærmere at glatbarbere teorien end at få den til at fremtræde som en redegørelse for hele intellektet. Gardner mener da også i stedet, at der eksisterer flere udøvende og styrende instanser, som vokser frem i og med:

”...as we become more competent in any domain, we eventually acquire the capacity to reflect upon that domain. Metacognitive capacity in one domain says nothing about comparable capacities in other domains.”¹³⁹

og

“As individuals become expert in various domains, they do become explicitly aware of what they are doing. So in that sense, they are conscious of their activities. This form of consciousness has a domain-specificity which the term "executive function" usually denies.”¹⁴⁰

Gardners betragter således det bevidsthedsmæssige (awareness) som en erhvervet evne. Hvordan Gardner undgår den uendelige regres, som ledsager homunculus tanken, er langt fra tydeligt. Ved i stedet at fortolke bevidstheden som sammensat af flere udøvende funktioner, bliver problemet ikke mindre eftersom disse funktioner, for så vidt de antages at styre og kontrollere antagonistiske informationsprocesser fri af hinanden¹⁴¹, blot bevirker, at der

¹³⁸ Gardner 1999, s.105.

¹³⁹ Se brevveksling med Howard Gardner s.89ff.

¹⁴⁰ Se brevveksling med Howard Gardner s.89ff.

¹⁴¹ Gardner 1999, s.105.

nu opstår flere uendelige regresser. Ved ensidigt at koncentrere sig om informationsbehandlingen, som varetages af det biopsykologiske potentiale mener Gardner, at kunne undgå den erkendelsesteoretiske diskussion, idet:

”As long as intelligences are restricted to the processing of ”contents in a world,” we avoid epistemological problems. So it should be. The concept of “intelligence” should not be expanded to include personality, motivation, will, attention, character, creativity, and other valued human capacities.”¹⁴²

Ved alene at betragte intelligenserne som udtryk for specifikke informationsbehandlings-systemer, der er rettet mod en bearbejdelse af objektiv information mener Gardner at undgå de erkendelsesteoretiske problemer. Det forekommer dog noget letkøbt, at en teori, der prætenderer at redegøre for det menneskelige intellekt i sit fulde omfang, ikke ønsker at inddrage dette forhold i en egentlig redegørelse¹⁴³. Problemet bliver ikke mindre omfangsrigt når man inddrager Gardners egen opfattelse af informationsbehandlingssystemet som et symbolmanipulerende organ, hvorved forståelsen opstår. Forståelsen tilvejebringes via de computationelle egenskabers manipulation med symboler, hvor:

”I conceive of a symbol as any entity (material or abstract) that can denote or refer to any other entity...so long as it is used (and interpreted) as representing some kind of information.”¹⁴⁴

Forståelsen er altså ikke andet og mere end denne repræsentations overensstemmelse med det repræsenterede aktualiseret gennem en potentiel egenskab, det biopsykologiske potentiale. Symbolmanipulationens relation til biologien ser Gardner således, at:

“...the forms and kinds of symbolization in which humans participate may also be guided by biological processes.”¹⁴⁵

¹⁴² Gardner 1999, s.204.

¹⁴³ At Gardner ikke har medtænkt dette forhold fremgår yderligere af brevvækslingen, hvori han siger: ”I did not realize that you expected of me a whole cognitive/mind theory. While I have certain ambitions this is well beyond my scope or capacity”. Se brevvæksling med Howard Gardner s.89ff.

¹⁴⁴ Gardner 1993a, s.302.

¹⁴⁵ Gardner 1993a, s.314.

Foreningen mellem de biologiske neurale processer og symbolmanipulation er til en vis udstrækning altså ledet af biologiske processer sådan forstået, at de menneskeskabte symbolkomplekser beror på disse primære symbolmanipulerende egenskaber¹⁴⁶.

Med neurovidenskabens resultater og om muligt kortlægning af neurale processer mener Gardner, at:

“If one has indeed specified the nature of human intelligences – the raw materials for cognition – on the one hand, and the range of human cultural roles and functions, on the other, one ought to be able to generate a list of all possible symbol systems and, if you like, all the domains in which human beings can become intellectually engaged.”¹⁴⁷

Gardner finder det altså muligt at kortlægge samtlige tænkelige symbolsystemer, som vi kan deltage i¹⁴⁸. Med en hypotetisk kortlægning af symbolsystemerne og mulighederne indenfor kognitionsforskningen synes det altså som om, at man kan skabe en komplet forening mellem teorier om kunstig intelligens og Gardners teori om mange intelligenser. En udtømmende beskrivelse af de tænkelige symbolsystemer beror til en vis udstrækning på en komplet beskrivelse af netop det biopsykologiske potentiale, hvilket Gardner opfatter som 'computational capacities'¹⁴⁹. Det er disse computationelle egenskaber, man på forskellig vis opererer med som udgangspunkt indenfor kunstig intelligens. Synspunktet forudsætter dog yderligere muligheden af at fremkomme med en udtømmende liste over de symbolsystemer som mennesket kan interagere igennem. En sådan opstilling af mulige symbolsystemer mener Gardner principielt vil være muligt¹⁵⁰. Hvad angår kunstig intelligens mener Gardner, at computere vil være i stand til at kunne overgå os indenfor nogle intelligenser, men han mener samtidig, at:

¹⁴⁶ Gardner 1993a, s.314.

¹⁴⁷ Gardner 1993a, s.303.

¹⁴⁸ Gardner er forventningsfuld på dette område eftersom han andetsteds fremhæver, at: "Until it becomes possible to designate neural circuitry as representing one or another intelligence in action, we cannot know for sure which intelligence or intelligences are being invoked on a specific occasion", Gardner 1999, s.95. Min kursivering.

¹⁴⁹ Gardner 1993a, s.66.

¹⁵⁰ Gardner 1993a, s.304.

“While artificial-intelligence programs can no doubt trounce us logically – and may soon surpass us in many spatial and linguistic feats – I view the notion of the interpersonal intelligences of a machine as a “category error.”¹⁵¹

En computer vil ifølge Gardner ikke kunne tilskrives de specifikt menneskelige intelligenser. Grunden til at Gardner ikke mener, at vi vil kunne tilskrive en computer menneskelige egenskaber, skal ses i sammenhæng med Gardners egen distinktion mellem menneske og computer hvorved Gardner mere udførligt finder, at:

“Whether computers can also develop personal intelligences is a subject of considerable dispute. Many experts in artificial intelligence believe that it is just a matter of time before computers exhibit intelligence about human entities...I feel that this is a category error: One cannot have conceptions of persons in the absence of membership in a community with certain values, and it seems an undue stretch to attribute such a status to computers¹⁵²”

Ifølge Gardner rummer vort begreb om mennesket en opfattelse af, at mennesket så at sige er indlejret i et samfundsmæssigt tilhørsforhold, med dertil hørende værdier. Umiddelbart kan vi ikke tilskrive en computer de samme egenskaber, hvorfor man må forkaste tanken om kunstig intelligens. At det netop er hele tanken om kunstig intelligens, der her må forkastes, ses yderligere af følgende forhold:

“What we usually call Consciousness- with a capital C--does seem to me to be an outgrowth of Intrapersonal intelligence, in that we cant be Conscious without a sense of the "I".”¹⁵³

og

“intrapersonal intelligence grows in part out of emotional reactions which, I believe, are not undergone by computers as we now know them. In contrast,

¹⁵¹ Gardner 1999, s.45.

¹⁵² Gardner 1999, s.111.

¹⁵³ Se brevveksling med Howard Gardner s.89ff.

emotions are irrelevant to logical-mathematical reasoning, as I understand it.”¹⁵⁴

Eftersom en computer ikke kan tilskrives de specifikke menneskelige intelligenser, kan den heller ikke tilskrives en menneskelig bevidsthed ud fra den betragtning, at denne ifølge Gardner er givet med den intrapersonelle intelligens. Hvorledes denne *consciousness with a capital C* nærmere er grundet i den intrapersonelle intelligens, kommer Gardner ikke nærmere ind på. Det i denne sammenhæng relevante for Howard Gardner er, at mennesket indgår i en kulturel sammenhæng, hvorigennem vi behersker forskellige symbolsystemer, og bedømmes via symbolsystemets rette placering i perspektiverne domæne og felt på vores brug af disse¹⁵⁵.

Når Gardner afskriver kunstig intelligens med henvisning til de personlige intelligensers karakteristika, synes det at forrykke balancen hvad angår status i de sideordnede intelligenser¹⁵⁶. Hertil kommer, at Gardners begreb om emotioner på baggrund af ovenstående, fremstår som noget tvetydigt, eftersom han andetsteds akkurat fremhæver:

“Also, I now consider emotional facets of each intelligence rather than restrict emotions to one or two personal intelligences.”¹⁵⁷

Hvordan emotioner kan være irrelevante for den matematisk-logiske intelligens og samtidig være et islæt ved disse er svært at se. På baggrund af de her gennemgåede perspektiver kan det synes urimeligt at tilskrive Howard Gardner en egentlig epistemologisk afklaret bevidsthedsopfattelse. Omvendt kan man hævde at i og med at Gardner beskæftiger sig med intelligens, og dermed forståelse, så må han i en eller anden forstand have et dybere begreb om bevidsthed. At det er relevant at inddrage det filosofiske begreb om bevidsthed, kan Gardner ikke være uenig i, eftersom han selv mener, at:

¹⁵⁴ Se brevveksling med Howard Gardner s.89ff.

¹⁵⁵ Der er her, om noget utydelig, så dog alligevel en parallel til Wittgensteins begreb om regelfølge. Den offentlige bedømmelse af i hvor stor udstrækning en regel er fulgt eller ikke kan ses analog med regelfølgeregabet, som forudsætter en objektiv kontekst, hvorved alene overensstemmelse eller uoverensstemmelse med reglen kan vurderes.

¹⁵⁶ Gardner 1993b, s.8.

¹⁵⁷ Gardner 1999, s.43.

“What it means to be intelligent is a profound philosophical question, one that requires grounding in biological, physical, and mathematical knowledge.”¹⁵⁸

At Howard Gardner så i praksis mener, at dette er uden for hans rækkevidde¹⁵⁹ modsiger egentlig hans eget udgangspunkt, hvor han netop understreger, at han søger at give den *rette*¹⁶⁰ og *fulde*¹⁶¹ beskrivelse af intelligenserne.

Det fremstår efterhånden evident at teorien om mange intelligenser er funderet på en implicit antagelse om, at bevidstheden knytter sig til hjernens computationelle egenskaber (computational capacities). Disse computationelle egenskaber er identiske med hjernens neurale netværk¹⁶². Eftersom den intrapersonelle intelligens ifølge Gardner antages at rumme jeg’et til forskel fra de andre intelligenser, synes det rimeligt at formode, at jeg’et på en eller anden måde er relateret til en specifik konstellation af neuroner. Hvorvidt jeg’et skal forstås som værende knyttet til den intrapersonelle intelligens uafhængig af erfaringen (a priori), eller i stedet er erfaringsbaseret (a posteriori), herom er Gardner utydelig.

Howard Gardners teori om mange intelligenser er på mange måder mangelfuld idet han undlader at gøre rede for forholdet mellem hjernens computationelle egenskaber (de neurale netværk) og bevidsthedsbegrebet. Ud fra den betragtning, at teorien om mange intelligenser søger at beskrive det menneskelige intellekt, synes det at være en undergravende fejl ved teorien, at den baserer sig på et begreb om forståelse og bevidsthed, som ikke i sig selv uddybes nærmere. At Gardner yderligere henfører en egentlig vurdering af intellektet til kontekstuelle forhold i kraft af menneskets indplacering i en given kulturel kontekst, undtager ham ikke mindre for at redegøre for det subjektives karakter. Det er jo ikke de neurale netværk, man vurderer på, men den forståelse, som et menneske opviser indenfor et givent intelligensområde. At man så i henhold til forskellige kulturelle kontekster har forskellige værdisæt, som udmønter sig i, at man eksempelvis i en kultur sætter sproglig intelligens over kropslig-kinæstetisk intelligens, hvor forholdet er modsat i en anden, mindsker ikke disse krav til teorien. Den mangler at redegøre for det aspekt, som er univer-

¹⁵⁸ Gardner 1999, s.22.

¹⁵⁹ Se brevveksling med Howard Gardner s.89ff.

¹⁶⁰ Gardner 1993b, s.7.

¹⁶¹ Gardner 1999, s.44.

¹⁶² Gardner 1993a, s. 63.

selt for alle kulturer, nemlig forståelsen, og dermed også bevidsthedens relation til de computationelle egenskaber.

Begrebet om computationelle egenskaber viser sig følgelig at være centralt både for teorier om kunstig intelligens, men også for Howard Gardners teori om mange intelligenser. Det står nu klart, at forholdet mellem bevidsthed og materie, hvad enten vi taler biologisk (neuroner) eller materielt (neurale netværk) indenfor begge teoretiske retninger søges redegjort gennem et begreb om computationer. Et mere indgående kendskab til disse computationers relation til bevidstheden og tilhørende problemstillinger finder vi i David Chalmers *enhedsdualisme*, hvorfor vi nu vil se nærmere på dennes bevidsthedsopfattelse.

5 Enhedsdualisme

Begrebet om computationelle egenskaber indgår, som vi har set det, på en central måde i teorier om kunstig intelligens og ligeledes i Howard Gardners intelligensteori. Begrebet om computationelle egenskaber og deres relation til de menneskelige kognitive egenskaber, er dog ikke uproblematisk. En bevidsthedsteori, der anvender begrebet om computationelle egenskaber centralt, og som samtidig forsøger at tage højde for de tilhørende problematiske aspekter, udfolder sig i den australske matematiker og filosof David J. Chalmers forfatterskab. Chalmers adresserer med hans enhedsdualisme bevidsthedens store problem (*'the hard problem of consciousness'*)¹⁶³, der netop vedrører, hvorledes fysiske processer i hjernen kan give anledning til subjektive oplevelser¹⁶⁴. Chalmers teori er i kraft af at tage fat på bevidsthedens store problem (*'hard problem'*) således rettet mod at forklare en ellers forklaringsmæssig afgrund (*explanatory gap*) - sammenhængen mellem hjerneprocesser og de tilhørende bevidsthedskomponenter¹⁶⁵.

5.1 Teorien om implementation

Eftersom den matematiske teori om computationer betegner et abstrakt forhold og teorier om kunstig intelligens beskæftiger sig med konkrete systemer, forudsætter en applikation af begrebet om computationelle egenskaber en teori om disse to sfæres relation. Teorien om implementation beskriver forholdet imellem det abstrakte matematiske begreb og det fysiske system, således at systemet kan anskues som en realisation af computationen, og computationen kan anskues som en beskrivelse af systemet¹⁶⁶. Teorien om implementation skal altså kunne gøre rede for svaret på spørgsmålet: *Under hvilke betingelser implementerer et fysisk system en given computation?*¹⁶⁷

Chalmers definition på en egentlig teori om implementation lyder således:

¹⁶³ I modsætning til 'the easy problems of consciousness', der omhandler spørgsmål som vedrører de objektive mekanismer af et kognitivt system jf. Chalmers: *The Puzzle of Conscious Experience*.

¹⁶⁴ Chalmers: *The Puzzle of Conscious Experience*.

¹⁶⁵ Chalmers: *The Puzzle of Conscious Experience*.

¹⁶⁶ Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁶⁷ Chalmers: *A Computational Foundation for the Study of Cognition*.

”A physical system implements a given computation when the causal structure of the system mirrors the formal structure of the computation”¹⁶⁸

Forudsætningen for i et system at implementere en given computation, skal altså findes i en antagelse af sammenlignelighed/isomorfisme (isomorphic) i den kausale topologi. Det generelle begreb om computationer kommer gerne til udtryk i en formalisme, og hvad enten formalismen er specificeret som Turing maskinen, Pascal programmer eller neurale netværk, deler de alle det samme grundlag – at de kan subsumeres under klassen af *combinatorial-state automata* (CSA). Alle formalismer vil derfor have det til fælles, at:

”...the formal state-transitional structure of the computation mirrors the causal state-transitional structure of the physical system.”¹⁶⁹

En sådan beskrivelse kan ikke alene anvendes på formalismer, men har langt mere vidtrækkende konsekvenser i og med, at:

“A CSA description specifies a division of a system into parts, a space of states for each part, and a pattern of interaction between these states. This is precisely what is constitutive of causal topology”¹⁷⁰

Implementationsteorien for sådanne *combinatorial-state automata* danner således basis for den generelle teori om implementation. I og med at et system kan opdeles og tilskrives bestemte tilstande (states) - og mønstret af disse tilstande, i kraft af deres relationer, kan udledes - har Chalmers påvist muligheden for en kausal topologi. Med kausal topologi forstår Chalmers blot den abstrakte kausale organisation, som systemet udfolder. Denne topologi er uafhængig af det, hvoraf systemets enkelte dele består og hvori de enkelte dele er implementeret¹⁷¹. Forbindelsen mellem computation og kognition ligger nu ligefor, hvis det kan vises, at kognitive systemer besidder mentale egenskaber i kraft af deres kausale organisation. Skulle dette være tilfældet, mener Chalmers at have bevist tesen om compu-

¹⁶⁸ Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁶⁹ Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁷⁰ Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁷¹ Chalmers: *A Computational Foundation for the Study of Cognition*.

tationel tilstrækkelighed¹⁷² (*computational sufficiency*) og kunstig intelligens er derfor en mulighed. Det er værd at bemærke, at Chalmers søger at danne en generel referenceramme og ikke umiddelbart udsiger noget om den ene eller anden slags formalisme, er korrekt i sin beskrivelse af bevidstheden¹⁷³.

5.2 Organisational invariants

Begrebet om organisational invariants er tæt knyttet til begrebet om kausal topologi således, at en egenskab kan betegnes som invariant, hvis denne ikke varieres i relation til et systems kausale topologi. Sammenhængen mellem en egenskab og et systems kausale topologi kan derfor udtrykkes således, at:

“Call a property *P* an organizational invariant if it is invariant with respect to causal topology: that is, if any change to the system that preserves the causal topology preserves *P*.”¹⁷⁴

Enhver ændring, der foretages til et system og som samtidig bevarer den kausale topologi bevarer altså ligeledes de organisationelle invariante egenskaber. Chalmers nævner fem forskellige typer ændringer, som man kan foretage i henhold til et system¹⁷⁵. Den mest interessante af ændringerne er tanken om at det er muligt at erstatte af dele af et system med andre dele, som udfører den samme logiske funktion, hvorunder systemet som helhed altså bevarer sin kausale topologi. Erstattes en neuron med en silicium chip (silicon chip), ændrer dette altså intet i forhold til systemets kausale topologi ifølge Chalmers. Ovenstående betyder, at princippet om organisational invariants kan defineres på følgende måde:

¹⁷² Tesen lyder: “...the right kind of computational structure suffices for the possession of a mind, and for the possession of a wide variety of mental properties.” Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁷³ Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁷⁴ Chalmers: *A Computational Foundation for the Study of Cognition*.

¹⁷⁵ Chalmers: *A Computational Foundation for the Study of Cognition*.

”...given any system that has conscious experiences, then any system that has the same functional organization at a fine enough grain will have qualitatively identical conscious experiences”¹⁷⁶

Inden vi omsætter begrebet om organisationel invarians til et kognitivt system i sit fulde omfang, er det nødvendigt først at udspecificere, hvoraf et kognitivt systems mentale tilstande består.

5.3 Fænomenale og fysiske egenskaber for bevidsthed

Når vi er bevidste om noget, er vi som regel i stand til at handle på dettes foranledning eller verbalt udtrykke dette. Omvendt kan man sige, at når en sådan information er tilgængelig, er vi som regel bevidste herom, blot i en anden forstand. Chalmers drager derfor en atypisk distinktion mellem consciousness (bevidsthed) og awareness (opmærksomhed). Opmærksomhed er en psykologisk tilstand, som er objektiv og fysisk i og med, at den dækker over hjernens tilstande¹⁷⁷. Bevidsthed er derimod en fænomenal subjektiv tilstand. Det ovenfor beskrevne forhold mellem bevidsthedens relation til foreliggende information er derfor således, at:

”Access consciousness is defined in terms of the causal role that a state plays, whereas phenomenal consciousness is defined in terms of the way the state feels.”¹⁷⁸

Opmærksomhed er med andre ord en tilstand, hvori en mængde information er direkte tilgængelig for tilsigtet kontrol, og kan karakteriseres ved sin kausale rolle – ved det, hjernen gør. Fænomenal bevidsthed er derimod egenskaber, som kan karakteriseres ved måden hvorpå disse tilstande erfares (experience) bevidsthedsmæssigt – måden som det at have tilstandene, føles på ¹⁷⁹. De to tilstande er forbundne således, at:

¹⁷⁶ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*.

¹⁷⁷ Chalmers: *The Puzzle of Conscious Experience*.

¹⁷⁸ Chalmers: *What is the unity of Consciousness?*

¹⁷⁹ Chalmers: *A Computational Foundation for the Study of Cognition*

“When a state is phenomenally conscious, it is access conscious, and vice versa.”¹⁸⁰

Relationen mellem de to former for mentale egenskaber, de fænomenale og de psykologiske, er således givet via en lovbunden kohærens – når der er bevidsthed, er der opmærksomhed og omvendt. Chalmers går derfor et skridt videre, eftersom han mener, at der eksisterer en række strukturelle omstændigheder ved de to tilstande, som er kohærente således, at:

”We can recast the underlying concept as a principle of structural coherence: the structure of conscious experience is mirrored by the structure of information in awareness, and vice versa.”¹⁸¹

Det psykofysiske forhold kan altså betegnes gennem princippet om strukturel kohærens, der netop anser forholdet mellem de psykologiske egenskaber og de fænomenale, som værende kohærent. Det kan diskuteres, om Chalmers strukturelle kohærens er en nødvendig eller en tilstrækkelig forbindelse. Er kohærens en nødvendig forbindelse, må der nødvendigvis være givet en bevidsthedstilstand (*consciousness*) til hver kognitiv tilstand (*awareness*), men det er ikke nødvendigvis tilfældet, idet:

”...we don’t quit believing what we believe when we are asleep (think of your beliefs about yourself such as your name, phone number, etc.), yet we are not then conscious of them.”¹⁸²

Kognitive tilstande er altså ikke nødvendigvis ledsaget af bevidsthed¹⁸³. Chalmers korrelation mellem *access consciousness* og *phenomenal consciousness* antyder, at det netop er en nødvendig forbindelse, han opererer med, hvorfor antagelsen om, at enhver kognitiv tilstand er ledsaget af en fænomenal tilstand, synes at være et en diskutabel præmis. I og med, at Chalmers definerer begrebet om *access consciousness* som værende den kausale

¹⁸⁰ Chalmers: *What is the unity of Consciousness?*

¹⁸¹ Chalmers: *The Puzzle of Conscious Experience*.

¹⁸² Harnish 2001, s.242.

¹⁸³ Harnish 2001, s.242. Daniel C. Dennett har endvidere mere udførligt diskuteret problemet omkring drømme og bevidsthed i bogen *Brainstorms*.

rolle en mental tilstand besidder indenfor et kognitivt system¹⁸⁴, så må han forholde sig til disse kognitive tilstande, der ikke umiddelbart synes at have en bevidsthedsmæssig karakter. Det gør Chalmers ved at medgive at definitionen på *access consciousness* til en vis grad er diskutabel og variabel¹⁸⁵, men det ændrer ikke ved det forhold for bevidstheden, som er opstillet mellem de to aspekter gennem princippet om strukturel kohærens. Den strukturelle kohærens mellem de to mentale aspekter – det psykologiske og det fænomenologiske – kan også udlægges som en relation, hvor den kognitive tilstand er principielt tilgængelig for bevidstheden, hvad John Searle kalder for konnektionsprincippet¹⁸⁶. Searle argumenterer imod konnektionsprincippet. Udgangspunktet for argumentet er følgende: At alle intentionelle tilstande, hvad enten de er bevidste eller ubevidste, har en aspektuel form ('aspectual shapes') - en måde at repræsentere verden på. Hvis man ser bort fra den introspektive beskrivelse, kan aspektuelle former karakteriseres på to måder - gennem adfærd eller neurofysiologisk. En sådan adfærdsmæssig eller neurofysiologisk beskrivelse vil dog aldrig være udtømmende. På baggrund af ovenstående argumenterer Searle, herefter¹⁸⁷:

1. Antag at der eksisterer 'dybe' ubevidste intentionelle tilstande.
2. Disse tilstande må nødvendigvis have en aspektuel form.
3. Til disse ubevidste intentionelle tilstande, såfremt de eksisterer, ville der ikke høre andet end en neurofysiologisk beskrivelse eller deres adfærdsmæssige konsekvenser.
4. Neurofysiologi og adfærd kan ikke beskrive den aspektuelle form.
5. Dybe ubevidste tilstande har derfor ingen aspektuel form.
6. Alle intentionelle tilstande har en aspektuel form.
7. Derfor eksisterer dybe ubevidste intentionelle tilstande ikke.

Searles argument belyser det åbenlyst urimelige i at hævde, at ubevidste intentionelle tilstande ikke besidder intentionalitetens fundamentale rettethed, og derfor i sidste instans slet ikke eksisterer. Argumentationen hviler dog på det grundlag, at man kun kan beskrive en mental tilstand gennem neurofysiologi eller adfærd, når man ser bort fra den introspek-

¹⁸⁴ Chalmers & Bayne: *What is the Unity of Consciousness?*

¹⁸⁵ Chalmers & Bayne: *What is the Unity of Consciousness?*

¹⁸⁶ Harnish 2001, s.242.

¹⁸⁷ Harnish 2001, s.243.

tive karakteristik. Denne antagelse er ikke nødvendigvis korrekt. Chalmers fremfører jo netop en computationel beskrivelse som en tredje mulighed, hvortil Searle ikke forholder sig, hvad der svækker hans argumentation noget. Dertil kommer, at Searle ikke forholder sig til muligheden for at nogle aspektuelle former ikke nødvendigvis har en udpræget kvalitativ karakter. I det øjeblik man eksempelvis associerer to forskellige ting med den samme kvalitative egenskab, har den aspektuelle form mistet sin udprægede kvalitative karakter. Ser vi bort fra denne kritik, er det ifølge Chalmers givet, at der består et sådant strukturelt bånd mellem de to mentale egenskaber, hvorfor de fleste mentale egenskaber må betragtes som organisationelle invarianter. Konsekvensen heraf bliver, at Chalmers kan hævde:

”A given functional organization can be realized by diverse physical systems. For example, the organization realized by the brain at the neural level might in principle be realized by a silicon system.”¹⁸⁸

I og med at de mentale egenskaber anskues som organisationelle invarianter, vil et system, som realiserer samme funktionelle struktur, besidde de samme mentale karakteristika. Som yderligere argumentation for princippet om organisationel invarians benytter Chalmers sig af en række tankeeksperimenter, hvoraf *Fading Qualia* her vil blive gennemgået.

5.4 *Fading Qualia*

Chalmers tankeeksperimenter har alle den samme argumentatoriske form som et *reductio ad absurdum*. Ved at tage et standpunkt og vise dets urimeligheder argumenterer Chalmers for det modsatte standpunkt. *Absent Qualia*, *Fading Qualia* og *Dancing Qualia* er alle sådanne tankeeksperimenter¹⁸⁹. Med *Fading Qualia*¹⁹⁰ tankeeksperimentet søger Chalmers at vise det mulige i, at to isomorfe systemer begge besidder bevidsthed.

Vi antager som udgangspunkt, at to systemer – ét neuronalt og ét baseret på silicium chips - besidder den samme funktionelle organisation, men hvor systemet baseret på silicium

¹⁸⁸ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁸⁹ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁹⁰ Tankeeksperimentet er her fremstillet med henblik på dets overordnede perspektiver og er således ikke gengivet med alle detaljer. For yderligere detaljer se Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

chips (vi kalder det for Robot) ikke er bevidst grundet nogle ikke-organisationelle faktorer¹⁹¹. På denne baggrund kan der konstrueres en række mellemtilstande de to systemer indbyrdes, hvor små ændringer foretages under bevarelse af den funktionelle organisation. Ved således gradvist at erstatte neuroner med silicium chips, der besidder de samme input/output funktioner, bevares systemets overordnede funktionelle organisation. Under en sådan ombytning melder spørgsmålet sig: Hvad vil det sige at være et system i en mellemtilstand? (*What is it like to be a system in between?*). Som svar på dette mener Chalmers, at der eksisterer to muligheder: bevidstheden udfases (*Fading Qualia*) eller bevidstheden forsvinder pludseligt (*Suddenly Disappearing Qualia*). Skulle konsekvensen være, at bevidstheden pludselig forsvinder, må bevidstheden i sidste instans bero på en enkelt neuron således, at når denne substitueres, forsvinder hele bevidstheden¹⁹². Denne mulighed synes stærkt implausibel ikke mindst i kraft af, at den modsiger hele vores opfattelse af naturen som kontinuert og ikke diskontinuert. Tilbage er muligheden, at bevidstheden udfases. Vi forestiller os nu et system, der besidder en degraderet form for bevidsthed uden at denne er helt forsvundet. Hvor det oprindelige bevidste system har en oplevelse af en ting, oplever det muterede system noget helt andet på trods af, at systemet har den samme adfærd og herunder verbale udsagn, som det bevidste system. Systemet i mellemtilstanden, hvis vi skal undgå den pludselige kollaps af bevidsthed, må således være karakteriseret ved:

“This is a being whose rational processes are functioning and who is in fact *conscious*, but who is completely wrong about his own conscious experiences.”¹⁹³

Systemet er altså bevidst, men samtidig ude af kontakt med dets bevidste erfaringer, eftersom erfaringerne udlægges som noget ganske andet end det, de egentlig er. Chalmers argumenterer derfor, at:

¹⁹¹ Her antager Chalmers som en forudindtaget præmis at Absent Qualia er muligt. Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁹² Udskiftningen behøves ikke at foregå på det neuronale niveau, men kan også foregå på eksempelvis det molekulære niveau. Konsekvenserne er de samme.

¹⁹³ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

”For a sentient, rational being that is suffering from no functional pathology to be so systematically out of touch with its experiences would imply a strong dissociation between consciousness and cognition.”¹⁹⁴

Ud fra præmissen om at systemets funktionelle organisation er bevaret og ikke besidder nogen dysfunktion, har vi altså at gøre med en position, hvor vi må antage, at forbindelsen mellem bevidsthed og kognition er strengt adskilte. Denne adskillelse er så urimelig at Chalmers mener, at udgangspunktet – eksistensen af ikke-organisationelle faktorer - må være fejlagtigt. Chalmers mener derfor, at:

“A much more reasonable hypothesis is therefore that when neurons are replaced, qualia do not fade at all...If so, then our original assumption was wrong, and the original isomorph, Robot, has conscious experiences.”¹⁹⁵

Det forekommer derfor mere plausibelt at hævde, at to systemer, som er forskellige med hensyn til bestanddelene, men identiske med hensyn til den funktionelle organisering, har identiske bevidsthedsoplevelser. At en udfasende bevidsthed er logisk mulig, berører ikke det faktum, at den forekommer empirisk umulig. Chalmers hævder således, at bevidsthedsbegrebet supervenerer naturligt og ikke logisk på det fysiske begrebsniveau¹⁹⁶. Selv betegner Chalmers denne position som nonreduktiv funktionalisme, hvor bevidstheden altså er bestemt af den funktionelle organisation, uden nødvendigvis at være reducerbar til denne¹⁹⁷. I bogen *The Rediscovery of the Mind* retter John Searle følgende indvending mod tankeeksperimentet. Searle forestiller sig, at ens bevidsthed udfases samtidig med, at ens adfærd er uanfægtet af dette. Betraget udefra synes der altså ikke at være noget galt, mens man indefra langsomt er ved at dø¹⁹⁸. At denne konsekvens ikke er utænkelig, begrundet Searle i at:

¹⁹⁴ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁹⁵ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁹⁶ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁹⁷ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

¹⁹⁸ Searle 1992, s.67.

”...we may suppose in this case that these regions¹⁹⁹ are gradually deteriorating to the point where there is no consciousness in the system...the silicon chips are able to duplicate the input-output functions of the whole central nervous system, even though there is no consciousness left in the remnants of the system.”²⁰⁰

Chalmers forkaster Searles argumentation, eftersom han ikke mener, at der er plads til dannelsen af nye overbevisninger (beliefs). Overbevisninger, og dermed bevidsthed, afspejles gennem systemets funktionelle organisation, hvorfor Searles indvending ifølge Chalmers må bero på en dualisme, hvor bevidstheden:

”...would float free of internal functioning like a disembodied mind.”²⁰¹

Dette dualistiske standpunkt er så uplausibelt, at Chalmers igen vender tilbage til, at der må være en forbindelse mellem bevidstheden og hjernens funktionelle organisation. Tanken om udfasende bevidsthed forkastes dermed til fordel for tanken om, at begge systemer besidder bevidsthed – Robot har altså bevidsthed i modsætning til vores indledende antagelse.

Tankeeksperimentet understøtter i en forstand selve ideen om, at der eksisterer psykofysiske love som eksempelvis princippet om organisationel invarians og princippet om strukturel kohærens. Disse psykofysiske lovmæssigheder kan dog ikke udgøre fundamentet for en generel bevidsthedsteori, men er rettere principper udledt på baggrund af en fundamental forudsætning. Spørgsmålet som melder sig er således, hvori disse lovmæssigheder er funderet.

5.5 Informationstilstande

I artiklen *The Puzzle of Conscious Experience* fremkommer Chalmers selv med et bud på, hvad der kunne gøre det ud for et teoretisk grundlag. Claude E. Shannons²⁰² matematiske

¹⁹⁹ De områder Searle her henfører til er bestemte neurale områder, nærmere betegnet den retikulære formation.

²⁰⁰ Searle 1992, s.67.

²⁰¹ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*

²⁰² Shannon: *A Mathematical Theory of Communication*.

teori for information karakteriserer begrebet om information som den mængde muligheder en given tilstand indeholder og udsiger dermed noget om, hvilke potentielle forskelligheder et system indeholder²⁰³. Disse informationstilstande kan indlejres i den fysiske verden – et fysisk realiseret informationsrum - ud fra den betragtning, at tilstanden tændt lys svarer til en bestemt spænding²⁰⁴. Den fænomenale farvedimension kan ligeledes ses som et fænomenalt realiseret informationsrum, hvor denne fænomenale realisering har en tilsvarende struktur realiseret fysisk²⁰⁵. På denne måde kan Shannons begreb om information danne fundament for Chalmers teori²⁰⁶. Ontologisk betragtet er det fænomenale og det fysiske således forskellige værensdimensioner, mens de realiserede informationsrum værensmæssigt er identiske. Begge dimensioner er altså udtryk for én og samme informationstilstand i informationsrummet. Perspektiverne for Chalmers bevidsthedsteori er altså i sidste instans:

”...that a theory of physics and a theory of consciousness could eventually be consolidated into a single grander theory of information.”²⁰⁷

Den overordnede teori om information kan altså ses som havende et fysisk og et fænomenalt aspekt. Hvorledes man ontologisk skal karakterisere informationsbegrebet i denne forstand, som en abstrakt og egenskabsløs substans, er uvist²⁰⁸.

For indeværende opgave er det mest interessante, at denne antagelse leder til det åbenlyse kontraintuitive forhold, at Chalmers bliver nødt til at fremstille ovennævnte lyskontakt

²⁰³ Eksempelvis indeholder en lyskontakt kun to tilstande, hvad lyset angår – tændt eller slukket. Et sådan informationsrum med to tilstande kaldes en *bit*, hvorfor systemet rummer information svarende til én bit.

²⁰⁴ Chalmers: *The Puzzle of Conscious Experience*.

²⁰⁵ Dette fremgår af forholdet omkring farveoplevelser, hvor: ”The three-dimensional encoding of color spaces... suggests that the information state in a color experience corresponds directly to an information state in the brain.” Chalmers: *The Puzzle of Conscious Experience*.

²⁰⁶ Hubert L. Dreyfus betoner det forhold som Shannon selv understregede - nemlig at *information* og *mening* ikke må forveksles. Informationsbegrebet i Shannons teori er tømt for det semantiske aspekt som meningen opviser. Ved at transformere informationsteorien til en meningsteori – som Chalmers – antager man fejlagtigt at ”*experience can be analyzed into isolable, atomic alternative choices*”, hvad der ikke følger med nødvendighed af Shannons informationsteori. Dreyfus 1979, s.165.

²⁰⁷ Chalmers: *The Puzzle of Conscious Experience*.

²⁰⁸ De her anførte overvejelser er nogle foreløbige overvejelser, som alene har til formål at afstikke teoriens generelle natur. Chalmers: *The Puzzle of Conscious Experience*.

eller en termostat som bevidst i og med, at disse systemer rummer en mængde information²⁰⁹. Reelt er der ifølge Chalmers to muligheder: enten har systemet en slags prototypisk bevidsthed eller også må man antage at kun noget information – eksempelvis komplekse informationsmønstre – har dette bevidsthedsmæssige (experientielle) aspekt. At Chalmers er tilhænger af bevidsthed for alle systemer fremgår tydeligt af artiklen *Absent Qualia, Fading Qualia, Dancing Qualia*.

Det kontraintuitive forhold, som her opvises, er netop spørgsmålet omkring implementati-
on og dennes problematik. Jeg vil i det følgende berøre denne diskussion og herunder ind-
drage aspekter fra teorier om kunstig intelligens, som vi så i det første afsnit.

²⁰⁹ En anden tilhænger af dette synspunkt er John McCarthy, der tilskriver en termostat bevidsthedstilstandende: ””The room is too cold”, ”The room is too hot”, and ”The room is OK””. John McCarthy *Ascribing Mental Qualities to Machines*.

6 Implementation

Hvad enten vi med Searle drager en distinktion mellem stærk og svag kunstig intelligens eller som Chalmers ser dette skel som nedbrudt - som en konsekvens af at begge paradigmer beror på en computationel fortolkning – så er der en fælles problematik, som tegner sig. Problematikken vedrører spørgsmålet om implementation. Tager vi den paradigmatiske distinktion mellem klassisk og kunstig intelligens frem, viser det sig med Chalmers, at vi ikke længere er i stand til entydigt at opretholde dette skel. Det konnektionistiske paradigme må i lyset af Chalmers begreb om computationalismens anvendelighed i sidste instans beløbe sig til at syntetisere den menneskelige bevidsthed. Men hvad enten konnektionismen andrager syntetiserende perspektiver eller ej, er spørgsmålet stadig, hvorledes det er muligt at forbinde bevidsthed med en materie, som er ganske forskellig fra den biologiske. Et spørgsmål, som også klassisk kunstig intelligens teorier må besinde sig på at afdekke. Begge paradigmer – konnektionistisk som klassisk - tager del i den opfattelse, som Chalmers ovenfor har fremvist - at bevidsthed er uafhængig af materien hvori den er implementeret. Overfor dette standpunkt står bl.a. John Searle, som mener, at biologien har mere at tilføje end ovennævnte standpunkt tillader.

Spørgsmålet om implementation vedrører i høj grad spørgsmålet om forholdet mellem syntaks og semantik. I bogen *Minds, Brains and Science* opstiller Searle argumentet *det kinesiske værelse* (The Chinese Room²¹⁰), som skal vise det åbenlyst kontraintuitive i forholdet at syntaks er tilstrækkelig for semantik. Scenariet er som følger: Searle forestiller sig en person er låst inde i et værelse med en manual - et sæt formelle regler for hvordan han skal manipulere med kinesiske symboler. Personen selv forstår ikke et ord af kinesisk og derfor heller ikke, hvad symbolerne betyder, men kan kun skelne symbolerne ud fra deres forskellige form. Reglerne udspecificerer blot hvilke symboler, der skal anvendes, og hvordan disse skal manipuleres med. Ved at skyde symboler ind under døren og lade personen med manualen i hånden finde frem til det symbol, der skal skydes ud af døren igen, kan man efterligne en situation, hvor eksempelvis spørgsmål bliver besvaret – vel at mærke uden at personen i værelset forstår noget som helst. Givet at manualen er god nok, kan man ikke

²¹⁰ Argumentet udvides med artiklen *Is the Brain's Mind a Computer Program?*

skelne mellem svarene fra det kinesiske værelse og fra en indfødt kineser. Forskellen er dog væsentlig idet, at:

”by virtue of implementing a formal computer program from the point of view of an outside observer, you behave exactly as if you understood Chinese, but all the same you don’t understand a word of Chinese.”²¹¹

Personen i det kinesiske værelse forstår ikke og har endsig ingen mulighed for at lære kinesisk. Searle betragter situationen som værende analog til den at implementere et program på en digital computer²¹². Derfor, hvis personen i det kinesiske værelse ikke forstår kinesisk, så gør en digital computer heller ikke. Bevidsthed og herunder forståelse af sprog kræver ifølge Searle mere end blot det at manipulere formelle symboler – det kræver fortolkning og tilskrivelse af mening til disse symboler. At en digital computer ikke kan besidde bevidsthed, fremgår af Searles argumentation, der i korte træk kan formuleres således:

- 1. aksiom: Hjerner forårsager bevidsthed
- 2. aksiom: Bevidsthed har mentalt indhold; mere præcist semantisk indhold
- 3. aksiom: Syntaks er ikke tilstrækkelig for semantik.
- 4. aksiom: Computerprogrammer defineres ved deres formelle eller syntaktiske struktur.
- Konklusion: Programmer er hverken konstituerende eller tilstrækkelige for bevidsthed²¹³.

I og med at computerprogrammer er defineret syntaktisk og det kinesiske værelse viser, at dette ikke er tilstrækkeligt for semantik, så er stærk kunstig intelligens en umulighed. Ud fra ovenstående aksiomer mener Searle desuden at kunne udlede, at den måde, som hjernefunktioner forårsager bevidsthed på, ikke alene kan være i kraft af et program²¹⁴. For at noget andet skal kunne forårsage bevidsthed, kræver det kausale kræfter, der er ækvivalen-

²¹¹ Searle 1984, s.32.

²¹² Searle 1984, s.33.

²¹³ Searle 1984, s.39.

²¹⁴ Searle 1984, s.40.

te med hjernens. I yderste instans konkluderer Searle derfor, at implementeringen af et program ikke er tilstrækkeligt for at en artefakt vil kunne besidde bevidsthed. Derimod kræves det, at denne artefakt besidder de samme kausale egenskaber som den menneskelige hjerne. Searles biologiske naturalisme viser sig her tydeligt, idet han altså hævder, at:

”...mental states are biological phenomena. Consciousness, intentionality, subjectivity and mental causation are all part of our biological life history...”²¹⁵

Searle betragter altså mennesket som tænkende biologiske maskiner²¹⁶. Standpunktet er derfor klart, at det, der afgør om noget er bevidst, altså ikke blot afhænger af den implementerede computationelle algoritme, men ligeledes af den materie, hvori computationen foretages. Chalmers og funktionalismens problem er altså ifølge Searle at den ikke medtænker biologiens engagement, hvorfor Searle hævder, at:

”...the computational properties of the brain are simply not enough to explain its functioning to produce mental states.”²¹⁷

Searle medgiver faktisk at en computationel beskrivelse af hjernen er mulig, men samtidig, at denne ikke er tilstrækkelig i en beskrivelse af de mentale egenskabers opståen. Som modsvar til det kinesiske værelse fremfører Chalmers det såkaldte *systemsvar*. Denne retning hævder, at det er hele systemet taget under ét, som rummer bevidsthed. Ifølge denne fortolkning viser kinesiske værelse blot, at en homunculus – manden i værelset – ikke forstår kinesisk, men derimod er der intet i vejen for at hævde, at hele systemet som sådan forstår kinesisk²¹⁸. I sidste instans vil dette, indbefattet Chalmers computationelle begreb, dog medføre, at man kan implementere en computation i en hel befolkning²¹⁹ for eksempel ved et system bestående af papirlapper, som overbringes i overensstemmelse med en formel regel og på denne måde vil man være nødsaget til at hævde, at denne konstellation

²¹⁵ Searle 1984, s.41.

²¹⁶ Searle 1984, s.40.

²¹⁷ Searle 1984, s.40.

²¹⁸ Chalmers: *A Computational Foundation for the Study of Cognition*.

²¹⁹ Chalmers: *Absent Qualia, Fading Qualia, Dancing Qualia*.

besidder bevidsthed. Chalmers hævder, at Searle begår en fejl ved i sit ræsonnement at sætte lighedstegn mellem program og implementation af et program; forskellen er dog, at:

”While programs themselves are syntactic objects, implementations are not; they are real physical systems with complex causal organization, with real physical causation going on inside.”²²⁰

Det afgørende er altså ifølge Chalmers, at ved implementation af et program forstummer diskussionen om syntaksens utilstrækkelighed for semantik. Searles argumentation strækker sig da også kun til at argumentere for at den implementerede algoritme må opvise de samme slags kausale kræfter som hjernen, uden at han dermed giver noget kriterium for at afgøre mellem hvilke ting, der har og hvilke ting, der ikke har denne kausale kraft til at skabe bevidsthed. Argumentationen glider her over i et forsvar af den biologiske naturalisme, hvor det springende punkt hvad angår bevidsthed ifølge Searle er, at:

”...mental states are both *caused by* the operations of the brain and *realized in* the structure of the brain (and the rest of the central nervous system).”²²¹

Hvad Searle forstår ved disse to forhold mellem det mentale og det fysiske træder endnu tydeligere frem når Searle lige ud siger, at:

”Mental phenomena are caused by neurophysiological processes in the brain and are themselves features of the brain.”²²²

Der er følgelig to relationer som gør sig gældende – en kausal interaktion og en identitetsrelation²²³. Disse to relationer er dog ikke entydige i og med, at Searle andre steder begrundet forholdet mellem mentale tilstande og neurobiologiske tilstande som et supervenient forhold, hvor det mentale supervenerer på det neurobiologiske, således at:

²²⁰ Chalmers: *A Computational Foundation for the Study of Cognition*.

²²¹ Searle 1983, s.265.

²²² Searle 1992, s.1.

²²³ Tim Van Gelder: *Monism, Dualism, Pluralism*

”...the supervenience of the mental on the physical is marked by the fact that physical states are causally sufficient, though not necessarily causally necessary, for the corresponding mental states.”²²⁴

Hvordan identitetsrelationen kan opretholdes samtidig med at de mentale egenskaber supervenerer på det fysiske uden dog at være ontologisk reducerbare²²⁵ hertil, er et åbent spørgsmål.

Fælles for Searle og Chalmers er, at de begge ser en forbindelse skabt mellem det mentale og det fysiske gennem kausation. Searle hævder, at det er den egentlige realisation i hjernen, der resulterer i bevidsthed, hvorimod Chalmers hævder, at det er det computationelle aspekt - uanset naturen af det, hvori denne realiseres – som er det væsentlige.

Overordnet betragtet er begge filosofers problem at undgå det dualistiske dilemma, der opstår når man på denne måde inddrages i sjæl-legeme problematikken gennem en påstået kausalitet mellem det mentale og det fysiske. Dilemmaet synes at have følgende to horn:

”Either you have dualism and an unintelligible account of causation or you have an intelligible account of causation and abandon the idea of the causal efficacy of the mental in favor of some version of the identity thesis with an attendant epiphenomenalism of the mental aspect of psycho-physical events.”²²⁶

En kausal forbindelse mellem de to sfærer leder altså i yderste konsekvens frem til, at man enten havner i mysticisme eller en form for epifænomenalisme²²⁷. Både Searle og Chalmers fremhæver det mentales nonreduktive karakter, den kausale forbindelse mellem mentale tilstande og neurobiologiske tilstande og bevidsthedens supervenation på det fysiske. De to filosoffer deler altså mange opfattelser sammen på trods af deres forskellighed. Metafysisk betragtet kan Searles *biologiske naturalisme* og Chalmers *non-reduktive funktionalisme*, da også begge subsumeres under samme metafysiske position - non-reduktiv fysika-

²²⁴ Searle 1992, s.125.

²²⁵ Reducerbar skal her forstås i ontologisk forstand. Searle benytter sig af en kausal reduktionisme, der efter hans mening ikke fører en ontologisk reduktion. Searle 1992, s.114f.

²²⁶ Searle 1983, s.264f.

²²⁷ Searle 1983, s.264.

lisme²²⁸. Selv Thomas Nagel anser det for umuligt at skelne Searles antireduktionistiske position fra egenskabsdualisme²²⁹.

Det problematiske i Chalmers *naturalistiske dualisme* og Searles *biologiske naturalisme* består således - med den non-reduktive fysikalisme – i at begge søger at opretholde en værenskategorisk forskel – en art ontologisk skel - mellem det mentale og det fysiske, samtidig med, at de herefter vil forklare og bygge bro mellem disse. Begge filosofers problem – det gælder også de to paradigmer indenfor teorier om kunstig intelligens - er i grunden at forklare, hvorledes dette tilsyneladende værenskategoriske skel mellem subjekt og objekt kan overvindes i en teori om bevidstheden, der forbinder de to sfærer. Alle har de det samme udgangspunkt i formodningen om, at det mentale, er forskelligt fra det fysiske, hvorefter de så reelt søger at forklare det ene ved det andet. Teoriernes egne forudsætninger står på denne måde i vejen for deres forehavende.

Selve det videnskabelige i projektet at forklare det mentale synes altså her at blive anfægtet. Spørgsmålet er blevet taget op af filosofen Thomas Nagel, hvorfor vi her vil se nærmere på hans betragtninger omkring det mentales relation til det fysiske.

²²⁸ Tim Van Gelder: *Monism, Dualism, Pluralism*

²²⁹ Nagel 1995, s.96.

7 Den naturvidenskabelige begrænsning

Diskussionen omkring det mentales relation til det fysiske er i høj grad et spørgsmål om relationen mellem karakteregenskaberne ved det subjektive og det objektive. Med *Mortal Questions* fremkommer Thomas Nagel netop med tanker, som vedrører dette forhold og hvorledes distinktionen mellem subjektivt og objektivt perspektiv virker ind på sjællelegeme problematikken²³⁰.

Nagels begreb om subjektivitet er i første omgang central for en forståelse af hans distinktion mellem subjektivt og objektivt perspektiv. Af samme grund vil vi her begynde med kaste et blik på det subjektive perspektiv.

7.1 *What is it like?*

Begreberne om objektivt og subjektivt perspektiv skal ses i den forstand at Nagel mener, at det, som karakteriserer det subjektive, netop er en bestemt form perspektiv:

”...fundamentally an organism has conscious mental states if and only if there is something that it is like to *be* that organism – something it is like *for* the organism.”²³¹

Det subjektive skal altså forklares ved, hvad det vil sige for en organisme at være denne bestemte organisme²³². Thomas Nagel bruger en flagermus til at belyse hans begreb om subjektivitet. Eftersom en flagermus er et pattedyr, virker det rimeligt at tilskrive den en form for bevidsthed (*experience*). En flagermus perciperer ved hjælp af en slags ekkolod. Når flagermusen perciperer sender den højfrekvente lydssignaler ud og ved at opfange reflektionerne, får den et billede, en perception af dens omgivelser. Gennem denne korrelation af udsendte og opfangede signaler bliver flagermusen i stand til at skelne afstand, størrelse, form, bevægelse og teksturer, der er ækvivalent til det menneskelige syn. I et forsøg på at finde ud af hvad for en slags oplevelser denne flagermus har, er vi begrænset af vores

²³⁰ Nagel 1979, s.198-213.

²³¹ Nagel 1979, s.166.

²³² Standpunktet er ofte blevet til parafrasen 'what it is *like*'. Nagel understreger at det ikke betyder, hvad det vil sige for os at være eksempelvis en flagermus, men betoner at det vigtige i denne sammenhæng er, hvad det vil sige for subjektet selv ('how it is for the subject himself'). Nagel 1979, s.170.

egen perceptuelle tilgang til omverden, hvorfor vi ikke kan forestille os andet end hvad det vil sige for et menneske at være en flagermus. Det centrale er dog, hvad det vil sige for en flagermus at være en flagermus. Ekstrapolationen fra os selv til hvad det vil sige at være en flagermus vil aldrig kunne danne andet end en form for skematisk begreb om hvad det vil sige. Den specifikke subjektive karakter af oplevelsen vil vi ikke kunne begribe²³³. Det centrale er ifølge Nagel derfor relationen mellem virkeligheden (*facts*) på den ene side og de begrebslige skemaer eller systemiske repræsentationer af denne virkelighed på den anden side.

Forestiller man sig, at man er en flagermus kommer man hurtigt frem til, at det er uendeligt svært at begribe, hvad det vil sige for en flagermus at være en flagermus²³⁴. Det subjektive perspektiv unddrager sig her altså en beskrivelse. Enten kan man derfor forkaste ideen om at flagermus har en slags mentale oplevelser eller man må erkende, at:

”...the subjective domain in alle its forms implies a belief in the existence of facts beyond the reach of human concepts.”²³⁵

Det er Nagels opfattelse, at vi må erkende at det subjektive rækker ud over hvad vi kan begrebsliggøre. Dette betyder dog ikke, at vi ikke kan tale om det subjektive - blot er vi nødsaget til at gøre det ud fra generelle observationer og ikke ud fra individuelle, personlige perspektiver²³⁶. Det subjektive aspekt - her kortlagt som et første-personligt perspektiv – som Nagel forfægter, skal ikke forstås som en forfægtelse af det subjektives principielt private karakter. I stedet mener Nagel, at:

”The point of view in question is not one accesible only to a single individual. Rather it is a *type*. It is often possible to take up a point of view other than one’s own, so the comprehension of such facts is not limited to one’s own case.”²³⁷

²³³ Nagel 1979, s.168f.

²³⁴ Nagel 1979, s.169.

²³⁵ Nagel 1979, s.171.

²³⁶ Nagel 1979, s.171.

²³⁷ Nagel 1979, s.171.

Ud fra den betragtning, at vi anskuer det subjektive som en type, en generalisation, kan vi tale om den fænomenologiske virkelighed (*phenomenological facts*) på en objektiv måde. Vi kan tilskrive en anden person subjektive fænomener og kvaliteter. Tilskrivelsen er objektiv, men samtidig subjektiv i den forstand, at den forudsætter en sammenlignelighed mellem de to subjekter²³⁸. Derfor er dét at begribe de subjektive kvaliteter hos en person som er født døv og blind ikke muligt for personer uden disse medfødte kvaliteter og omvendt. At vi ikke kan begrebsliggøre denne subjektivitet, leder os dog ikke til at forkaste troen på eksistensen af denne. Med udgangspunkt i flagermusen kan man derfor hævde:

”Reflection on what it is like to be a bat seems to lead us, therefore, to the conclusion that there are facts that do not consist in the truth of propositions expressible in a human language. We can be compelled to recognize the existence of such facts without being able to state or comprehend them.”²³⁹

Der eksisterer altså et noget, et *what it is like*, hvormed vi ikke kan begrebsliggøre os endsi-ge forstå dette. En forkastelse af subjektiviteten ville her medføre en forkastelse af vores egen oplevelse af virkeligheden:

“And to deny the reality or logical significance of what we can never describe or understand is the crudest form of cognitive dissonance.”²⁴⁰

Ved at forkaste det subjektive på baggrund af dets uforklarlighed, dets resistens overfor begrebsliggørelse og dermed fratage det subjektive sin plads i virkeligheden, havner man i en uoverstigelig konflikt. Nagels beskrivelse af det subjektives karakter kræver, at han definerer forholdet mellem subjekt og objekt på en signifikant måde, hvorfor vi nu vil se nærmere på denne distinktion.

7.2 Subjekt og objekt virkelighed

Ifølge Thomas Nagel er der en tendens til at søge en objektiv beskrivelse af alt før man godskriver dette som værende virkeligt²⁴¹. Det objektive perspektiv ligger immanent i na-

²³⁸ Nagel 1979, s.172.

²³⁹ Nagel 1979, s.171.

²⁴⁰ Nagel 1979, s.171.

²⁴¹ Nagel 1979, s.196.

turvidenskabens måde at beskrive virkeligheden på – man fristes til at sige, at virkeligheden er objektiv. Men som vi har set det ovenfor undslipper det mentale, det subjektive sig netop en objektiv beskrivelse. Hvad enten vi opfatter bevidstheden som et resultat af homunculustanken eller systemtanken, udfolder denne diskussion sig under naturvidenskabelige præmisser i kraft af deres begrebslige tilgang til kausalitet. Hvor systemtanken henfører bevidstheden til hele systemet i kraft af dettes kausale relationer, henfører homunculustanken det til en form for agent. Herved vinder man intet i forklaringen af det mentale eftersom:

”...by substituting an agent as the cause, it avoids the alternative that actions is something that just happens. It is a doomed attempt to capture the doings of the action in a new kind of causation.”²⁴²

Det utilfredsstillende i denne forklaring skal ses deri, at vi havner i den samme problematik, der angår begrebet om kausation og hvorledes bevidsthed opstår i relation hertil²⁴³. Det naturvidenskabelige begreb om kausation forstået som et årsags-virkningsforhold er det egentlige problem. Problemet ved at betragte det mentale som et naturvidenskabeligt kausalt forhold er, at:

”So long as mental states are looked at objectively, in their causal relations to stimuli and behavior, no special issues arise which do not arise about the physical analysis of other natural phenomena.”²⁴⁴

En beskrivelse af eksempelvis hjernens neurale funktioner er ikke forskellig en beskrivelse at ethvert andet fysisk fænomen, men det mentales karakter lader sig heller ikke åbenbare gennem denne objektive kausalrelationelle beskrivelse. Det synes altså umuligt gennem en fysisk beskrivelse at redegøre for, hvad det vil sige for et væsen at besidde disse mentale tilstande²⁴⁵. Det antagonistiske forhold mellem subjektivitet og objektivitet kan belyses ved forskellene i perspektiv således, at:

²⁴² Nagel 1979, s.198.

²⁴³ Sammenligneligheden med den mekaniske fornufts paradoks ligger ligefor.

²⁴⁴ Nagel 1979, s.198f.

²⁴⁵ Frank Jackson fremhæver med artiklen *Epiphenomenal Qualia* det samme forhold - bevidstheden unddrager sig en fysisk beskrivelse.

”...subjective aspects of the mental can be apprehended only from the point of view of the creature itself...whereas what is physical is simply there, and can be externally comprehended from more than one point of view.”²⁴⁶

Forskellen er altså at objektiviteten tillader, at man kan anskue et forhold fra mange forskellige synsvinkler - en upersonlig anskuelsesform -, hvorimod subjektiviteten er anskuelser fra en bestemt synsvinkel - en personlig anskuelsesform. Netop den personlige anskuelsesform er umulig at bringe ind under den upersonlige til trods for, at vi som væsener er en del af den verden, som netop kan anskues upersonligt og objektivt²⁴⁷. Vi befinder os altså i en art kontradiktorisk forhold. Som bevidste væsener besidder vi mennesker mentale egenskaber, der kun kan beskrives ud fra os selv samtidig med, at vi er en del af en objektiviserbar virkelighed. Det mest nærværende virkelige - mine mentale tilstandes egenart - er bare ikke beskrivelige i det objektiviserende sprog, som med naturvidenskabens paradigmatiske karakter tegner virkeligheden.

Nagel fremhæver, at kontradiktionen mellem det subjektive og objektive ikke er kategorisk, men nærmere en begrebslig kontradiktion indenfor en skala, hvor det subjektive og objektive er polære modsætninger²⁴⁸. I kraft af denne polære begrebslige anskuelse transcenderer han den naturvidenskabelige begrænsning, som netop oppebærer den kategoriske distinktion mellem subjekt og objekt. Det polære forhold mellem subjektiv og objektiv spænder på en skala lige fra det enkelte individs synsvinkel i den ene ende til det objektiviserende ideal, der netop er unddraget en bestemt synsvinkel, i den modsatte ende. Ser man individet under dette objektiviserende ideals 'synsvinkel', er individet derfor blot et af mange elementer²⁴⁹. På skalaen er forholdet mellem subjektiv og objektiv relativt de to poler imellem således forstået, at:

”A general point of view is more objective than the view from where you happen to be, but less objective than the viewpoint of physical science.”²⁵⁰

²⁴⁶ Nagel 1979, s.201.

²⁴⁷ Nagel 1979, s.201.

²⁴⁸ Nagel 1979, s.206.

²⁴⁹ Nagel 1979, s.206.

²⁵⁰ Nagel 1979, s.206.

Dette relative forhold svinder af syne indenfor sjæl-legeme diskussionen, hvor det kontradiktoriske forhold og det objektive ideal har vundet overtaget og dermed positionen som den rette beskrivelse af virkeligheden. Nagels polære distinktion kan nemt opfattes som at han forfægter subjektivitetens karakter som principielt privat. Nagel mener dog ikke, at det subjektive nødvendigvis skal betragtes som privat i modsætning til den offentlige objektive tilgang. Derimod betragter Nagel disse subjektive tilstande som intersubjektivt tilgængelige, idet han forudsætter:

”...that the subjective ideas of experience, of action, and of the self are in some sense public or common property.”²⁵¹

Denne offentlige og intersubjektive opfattelse af det mentale deler Nagel med Ludvig Wittgensteins senfilosofiske opfattelse i *Philosophische Untersuchungen*²⁵². De subjektive fænomenologiske begrebers offentlige status er dog en anden end den offentlighed, som fysiske beskrivelser besidder²⁵³. For at forstå dette skelner Nagel mellem fremtræden (*appearance*) og egenskaber (*properties*). Det mentales intersubjektive aspekt viser sig med Wittgenstein ifølge Nagel ved at:

”Wittgensteins position on sensations is that they just *are* appearances, so their properties are not the properties of objects which appear to whoever has them, and similarity in their properties is not similarity in the properties of such objects. Rather it is a similarity in appearances. That is a similarity between irreducibly subjective phenomena.”²⁵⁴

Denne opfattelse af det mentales intersubjektive karakter som en sammenlignelighed i fremtræden har dog en række konsekvenser. Wittgensteins opfattelse medfører i yderste konsekvens i kraft af hans privatsprogsargument, at vi ikke kan forstå det mentale uafhængigt af dets ydre kriterier²⁵⁵. Wittgenstein karakteriserer selv bevidsthedsfænomener - som eksempelvis smerte (*Empfindung*) - på følgende måde:

²⁵¹ Nagel 1979, s.207.

²⁵² Nagel 1979, s.207.

²⁵³ Nagel 1979, s.207.

²⁵⁴ Nagel 1979, s.207.

²⁵⁵ Gorzelak 1995, s.148.

”Sie ist kein Etwas, aber auch nicht ein Nichts! Das Ergebnis war nur, daß ein Nichts die gleichen Dienste täte wie ein Etwas, worüber sich nichts aussagen läßt.“²⁵⁶

Bevidsthedsfænomener er følgelig ikke et noget, men heller ikke et intet. Det centrale er, at de objektive kriterier nødsager os til ikke at kunne udtale os om disse bevidsthedsfænomener andet end gennem det objektiverende sprog. Sprogligt betraget vil et intet (*Nichts*) derfor have den samme funktion som et noget (*Etwas*), hvorom vi alligevel ikke kan udtale os²⁵⁷. Det subjektives ontologiske status synes, gennem en accept af privatsprogsargumentets konklusioner, derfor at unddrage sig en forklaring²⁵⁸. Men hvor Wittgenstein ikke fikserer det subjektives ontologiske status, gør Thomas Nagel det derimod.

Konsekvensen er, at Nagel forkaster det objektive syn på verden, ikke fordi han anser det som ufuldstændigt, men derimod fordi dette syn kun danner et partielt billede af virkeligheden²⁵⁹. Med denne opfattelse stiller Nagel sig i kontrast til fænomenologien og materialismen, som begge deler det samme udgangspunkt:

”a conviction that a single world cannot contain both irreducible points of view and irreducible objective reality – that one of them must be what there *really* is and the other somehow reducible and dependent on it.”²⁶⁰

Overbevisningen indenfor de to retninger er, at et man for at tegne et uniformt billede af virkeligheden havner i en beskrivelse af denne, hvor det ene synspunkt kommer til at repræsentere virkeligheden, som den *virkelig* tager sig ud, og det andet, på ene eller anden led, er reducibelt og afhængig af dette. Nagel søger at undgå denne uniforme tankegangs problemer i og med, at han ser det som en opgave at acceptere begge synspunkter – objektivt og subjektivt – som et ufravigeligt grundvilkår, hvorigennem vi betragter verden. Opgaven ligger derfor ligefor:

²⁵⁶ Wittgenstein 1995, s.376.

²⁵⁷ Wittgenstein kommer med *Philosophische Untersuchungen* derfor frem til den samme konklusion som i *Tractatus*, nemlig at givet sprogets natur er visse regioner unddraget dette og dermed uudsigelige. Gorzelak 1995, s.149.

²⁵⁸ Gorzelak 1995, s.148.

²⁵⁹ Nagel 1979, s.212.

²⁶⁰ Nagel 1979, s.212.

“The task of accepting the polarity without allowing either of its terms to swallow the other should be a creative one. It is the aim of eventual unification that I think is misplaced, both in our thoughts about how to live and in our conception of what there is. The coexistence of conflicting points of view, varying in detachment from the contingent self, is not just a practically necessary illusion but an irreducible fact of life.”²⁶¹

Med dette synspunkt forkaster Nagel ikke naturvidenskaben, men nærmere dennes iboende ideal om, at virkeligheden kan beskrives gennem en singularer fremtrædelsesform. Som en konsekvens rejser Nagels synspunkt tilligemed det problem, at der ikke synes at være tale om en enkelt måde, hvorpå ting er i sig selv²⁶². Selv om Nagel fastholder, at en beskrivelse af virkeligheden må indbefatte begge synspunkter, synes dette dog at stå overfor uoverstigelige problemer. De epistemiske konsekvenser ved at acceptere Wittgensteins privatsprogsargument er at netop sætninger om det egenpsykiske ikke er kognitive, men derimod ekspressive²⁶³, hvorfor Nagel som en konsekvens heraf synes at bevæge sig væk fra sit udgangspunkt, nemlig det subjektive synspunkts særegne karakter. En egentlig karakteristik af bevidsthedsfænomener må, qua accepten af privatsprogsargumentet, bero på offentlige og objektive kriterier, hvad der netop unddrager sig de karakteristiske subjektive aspekter. Nagel forfægter heller ikke en overbevisning om, at vi overhovedet vil komme i nærheden af en beskrivelse af virkeligheden, som ikke er partikulær idet han anfører:

”...a general conception of reality would require a general conception of experience which admitted our own subjective viewpoint as a special case. This is completely beyond us and will probably remain so for as long as human beings continue to exist.”²⁶⁴

I og med at en general beskrivelse af virkeligheden forudsætter en generel beskrivelse af bevidsthed (*experience*), som netop inkluderer alle former for bevidsthed, hvorunder menneskets kun er et tilfælde, synes ideen om en uniform virkelighedsbeskrivelse at fortone

²⁶¹ Nagel 1979, s.213.

²⁶² Nagel 1979, s.212.

²⁶³ Gorzelak 1995, s.149ff.

²⁶⁴ Nagel 1979, s.213.

sig. Med en noget romantisk talemåde kunne man sige, at virkeligheden er mangfoldig og givet vort grundvilkår ikke lader sig beskrive udtømmende på en entydig måde.

I det følgende vil jeg sammenfatte de i opgaven udlagte iagttagelser og analyser med henblik på at vende tilbage til opgavens udgangspunkt og problemstilling.

8 Version 3.0?

Som jeg indledningsvis anførte har formålet været at undersøge, om Howard Gardners psykologiske teori om mange intelligenser kunne perspektivere området for kunstig intelligens.

Det er skæbnens ironi at Howard Gardners problemer i bund og grund er de samme som vi har set diskuteret indenfor området for kunstig intelligens. Gardners teori bygger immanent på en opfattelse af forståelse, som udspringende af biologien. Med Gardners opfattelse af hjernen som bestående af funktionelt opdelt computationelle områder, hvorudfra forståelse skulle udspringe, synes det samme problem at opstå som gør sig gældende for Chalmers computationelle tilgang: at påvise relationen mellem de subjektive kvaliteter og de objektive neurale aktiviteter på en adækvat måde. Gardner kaster efter min mening på dette område ikke nyt lys over sagen.

I relation til kunstig intelligens kan man yderligere tilføje, at Gardner på den ene side med sine betragtninger åbner op for at vi kan tale om computere som værende intelligente²⁶⁵, og at han på den anden side udelukker muligheden af kunstig intelligens. Vi tvinges nærmest til at antage, at en computer har en art bevidsthed, eftersom den kan klassificeres under den samme intelligensteori som et menneske, men hvad der reelt gemmer sig bag betegnelsen af en computer som intelligent er svært at se. Howard Gardners ræsonnementer på dette område er præget af en række inkonsistente og modstridende betragtninger. På den ene side forkaster han homunculustanken for herefter, at indføre denne, uden de større betænkeligheder, på området for de personlige intelligenser om end på en noget uklar facon. Det er min klare overbevisning, at en række af Gardners inkonsistente betragtninger er afstedkommet af et ønske om i sit fulde at redegøre for de menneskelige kognitive egenskaber - uden dog herunder at medtænke en egentlig epistemisk afklaring. Et sådant projekt mener jeg i grunden er inkonsistent. Det kan ikke lade sig gøre at fundere en intelligensteori i det biologiske og medregne forståelsen uden at redegøre for, hvorledes bevidsthed opstår i relation hertil. Gardner rammes af denne kritik, idet han prætenderer at give en komplet beskrivelse af de menneskelige kognitive egenskaber.

²⁶⁵ De besidder matematisk-logisk intelligens. Dertil medgiver han at computere med tiden vil overgå os indenfor de verbale og rumlige intelligenser.

At Gardner som udgangspunkt søger at skabe en dynamisk opfattelse af intelligens med baggrund i biologien er prisværdigt. Relationen mellem intelligens og det biopsykologiske potentiale forekommer dog noget mudret, idet han på den ene side fastholder at intelligensbegrebet er et konstrukt, samtidig med at intelligensen defineres ækvivalent til det biologiske korrelat; hvad han modsat ikke anser som et konstrukt, men nærmere en art ontologisk realitet. Som Gardner erkender i brevvekslingen²⁶⁶ ligger de erkendelsesteoretiske overvejelser uden for hans rækkevidde, hvad jeg mener er en grundlæggende fejl, hvorfor jeg her vil konkludere at *teorien om mange intelligenser* ikke på centrale områder kan belyse aspekter ved kunstig intelligens. Det forekommer nærmere som om, at kognitionsforskningen kan belyse og berige Howard Gardners teori end omvendt.

De her fremførte overvejelser er dog ikke uden relevans i og med at de har opvist en central problemstilling, som både må imødekommes af Howard Gardners teori om mange intelligenser og teorier om kunstig intelligens. Det centrale i diskussionen om kunstig intelligens har hidtil været en kamp mellem de to paradigmer - mellem syntese eller simulering. Men både klassisk kunstig intelligens og konnektionismen må forholde sig til et og samme problem, nemlig relationen mellem det fysiske og mentale, uanset om man prætenderer en syntese eller en simulering af bevidstheden. Og netop i denne relation opvises det egentlige kardinalpunkt for diskussionen – bevidsthedens karakter eller egenart. Hele diskussionen om muligheden for kunstig intelligens forudsætter vel egentlig den præmis, at vi kender originalen - at vi ved, hvad intelligens, og dermed bevidsthed og subjektivitet, er for en størrelse. Men som vi har set er det langt fra indlysende, hvad bevidsthed er for en størrelse. Bevidsthedens synes ikke at være kortlagt adækvat og udtømmende. Kognitionsforskningen tydeliggør dette for så vidt, at dens fundament beror på et tvetydigt begreb om intelligens.

Kunstig intelligens er som konsekvens heraf konfronteret med det skisma, som har præget filosofiens historie – forholdet mellem sjæl og legeme. Findes der en vej ud af dette skisma synes det derfor mere rimeligt at betragte kunstig intelligens primære mål som et forsøg på at afdække bevidsthedens natur, og *der næst* som et forsøg på at udforske muligheden for simulering eller syntese af denne. Kunstig intelligens og dennes mulighed beror altså i lige så høj grad på, at den afdækker det, den forudsætter; før vil det ikke være muligt at have et

²⁶⁶ Se brevveksling med Howard Gardner s.89ff.

begreb om det mulige - endsige umulige i at projektet når sit mål. Det i opgaven indledende citat af Shakespeare synes derfor at rumme en optimisme, som ikke er åbenbar. I bund og grund *ved* vi ikke hvad vi er – netop derfor er diskussionen om det mulige eller umulige ved kunstig intelligens en pseudodiskussion. Førend vi afdækker vores egen natur kan vi end ikke diskutere muligheden af at syntetisere eller kopiere denne.

I og med at meget af diskussionen om muligheden for kunstig intelligens tager udgangspunkt i en fundamentalt set uafklaret natur kommer diskussionen, som vi har set det, mere til at dreje sig om et trosspørgsmål, som gennem hele filosofihistorien har været til debat. Diskussionen eksemplificeres ved implementationsspørgsmålet, hvor den centrale debat udmøntes i en kamp på sandsynligheder mellem et biologisk og et fysisk standpunkt. Dette skal ikke forstås således, at jeg mener, at diskussionen er nyttesløs, men tværtimod, at jeg ser denne diskussion som udspringende af det faktum, at bevidsthedens karakter på mange måder stadig er os fremmed. Det at bevidsthedens karakter ikke er entydigt fastlagt medfører at implementationsspørgsmålet i høj grad er et spørgsmål om tro. På baggrund af intuitive og kontraintuitive betragtninger afstikkes mulighederne for kunstig intelligens og dermed forholdet mellem det subjektive og objektive.

Thomas Nagels tanker omkring det subjektives natur fremstår ikke som det endegyldige svar på dette spørgsmål, men nærmere som en anvisning af, at det subjektive per definition synes at have en kvalitativ karakter, som i en vis forstand er unddraget den objektiverende sprogbrug. Problemerne omkring Thomas Nagels objektive fænomenologi skal jeg ikke genfremstille her, blot vil jeg fremføre, at den centrale pointe synes at bestå i, at både teorier om kunstig intelligens og også Howard Gardners teori om mange intelligenser i grunden forudsætter en opfattelse og beskrivelse af virkeligheden, som fjerner sig fra det centralt subjektive – nemlig perspektivet fra første-person. Opfatter man naturvidenskabens objektiverende perspektivløse ideal, som svaret på hvordan virkeligheden *virkelig* er, synes man at afskære sig fra at inkludere det subjektive i en beskrivelse af denne, fordi det subjektive om noget kun adækvat kan karakteriseres ved ét signifikant perspektiv og ikke et perspektivløst syn.

Thomas Nagels synspunkter kan i en forstand ses som et opgør med rationalitetens princip, der udfoldes i naturvidenskabens grundsyn og som gennemsyrrer hele vores vestlige tankegang: et objektivistisk ideal, hvor virkeligheden og sandheden ses som en objektiv stør-

relse, der kan tilegnes gennem blotlæggelse og dermed forklares udtømmende og adækvat. At det objektive eksisterer, selv når det ikke erkendes af et subjekt, indebærer ikke nødvendigvis at al realitet på samme måde skal kunne afskæres fra et erkendende subjekt, som naturvidenskaben i en eller anden henseende forstår det. I og med at man ikke sætter spørgsmålstejn ved naturvidenskaben bliver denne et dogme, som ligger i vejen for udvikling på dette område. Selvfølgelig ideen om det rationelle i videnskaben afspejler implicit forestillingen om at udvikling sker i en kontinuerlig lineær proces hen imod en mere sand og korrekt beskrivelse af virkeligheden. Denne myte eller antagelse afskærer sig selv fra sin egen udvikling, hvad der tydeliggøres indenfor området for kunstig intelligens, hvor antagelsen netop møder sine egne begrænsninger ved forsøget på at objektivere det umiddelbart ikke objektivbare, første-personlige synspunkt, hvorom vi ikke synes at kunne begrebsliggøre os andet end gennem generaliseringer.

Vort videnskabsteoretiske ideal foreskriver os netop at medtage flest mulige forklaringer i forsøget på at beskrive eller forklare fænomener. Idet naturvidenskaben forudsætter sit eget fundament og ikke stiller sig kritisk overfor sig selv, sætter den ikke følgende til debat: Findes der andre mulige forklaringsmodeller til at beskrive virkeligheden med, end det grundlæggende rationelle videnskabssyn, som hidtil har præget kognitionsforskningen? Nu er det ikke kun kunstig intelligens og Gardners intelligensteori, der rammes af denne begrænsning af det objektivierende ideal. Filosofiens egen gyldighed beror i vid udstrækning på kravet om validitet – gyldiggørelse og bevisførelse. Eksemplificeret med Chalmers og Searle synes filosofiske teorier med indbyggede aksiomer at falde for deres egne forudsætninger når de behandler relationen subjekt – objekt. Kravet om bevisførelse synes dog ikke anfægtet indenfor logikken, eftersom denne er i overensstemmelse med de naturvidenskabelige idealer eller vice versa.

Nu er det ikke filosofiens selvberettigelse som står til diskussion, men derimod dennes relation til dette emneområde. Filosofien har selv gennem årtusinder søgt svaret på erkendelsens karakter – epistemologisk såvel som ontologisk. Vel at mærke uden at komme til et endegyldigt svar. Om svaret på menneskets natur kan fremføres under et uniformt udtryk kan, som vi har set det med Thomas Nagel, diskuteres. Alligevel består der en opgave i at udforske menneskets natur nærmere.

Det må derfor i fremtiden være gældende, at filosofien hånd i hånd med kognitionsforskningen udforsker bevidstheden. De faglige grænser bør nedbrydes, således at målet – at afdække menneskets inderste natur - sker i et sammenspil mellem fagområder. Alt for ofte sker det, som eksemplet med Howard Gardner viser, at en faggruppe ensidigt undersøger et felt, uden at lade sig berige af perspektiver udarbejdet inden for et andet område eller disciplin. Det mest prægnante udtryk for dette ses i den stående forherligelse af netop kunstig intelligens på baggrund af vores tekniske formåen. Ray Kurzweils *The Age of Spiritual Machines* er bare et af mange eksempler på den teknologiske forherligelse i betragtningen af det mentales natur. Diskussionen om robotter i fremtiden vil kunne besidde bevidsthed er en diskussion, der må foregå på ganske andre præmisser end hidtil – ikke forstået således, at jeg her parteciperer i Thomas Nagels synspunkt, men jeg mener grundlæggende, at det subjektives egenart, på baggrund af de i opgaven fremlagte betragtninger, synes at unddrage sig en objektiverende forklaring.

Det er ikke min hensigt at forfægte en erkendelsesteoretisk skepticisme i denne diskussion og postulere, at vi ikke kan nå til en dybere forklaring af det subjektive. Jeg finder det modsat meget sandsynligt at bevidstheden er knyttet til biologiske-fysiske processer. Men det er min overbevisning, at vi er langt fra en egentlig *erkendelse* af bevidsthedens grundlæggende natur.

Det er tvingende nødvendigt, at vi for at komme nærmere en dybere forståelse af bevidstheden – de ufattelige tanker - vælger en ny måde at anskue denne på. Denne nye optik må bestå i, med Alwyn Scotts ord, at:

”Forskerne bør være mere ydmyge, når de står over for sådan en væld og vigtighed, som livets og sindets natur udgør.”²⁶⁷

Denne ydmyghed bør dog ikke begrænses til forskere indenfor naturvidenskabelige discipliner alene, men gælder i lige så høj grad forskere indenfor filosofien. En sådan ny optik vil rydde en række antagelser af vejen som åbenbart grundløse, hvorefter man kan håbe på mere klart at kunne fastholde fokus på de mere indlysende problemer omkring kunstig intelligens. Kun således vil det være muligt engang i fremtiden at gøre sig forhåbninger om

²⁶⁷ Berlingske Tidende, 8/12-2001. 5. sektion, bagsiden.

at kunne betragte homo sapiens version 3.0 –en menneskelig hybrid af maskine og *bevidst* bevidsthed.

Kevin Warwicks postulat, som jeg indledningsvis berørte, synes ikke mindst at understrege behovet for et tættere samarbejde mellem filosoffer og forskere. Warwick antager netop uproblematisk, at overførsel af nerveimpulser fra ham selv til hans kone vil svare til en identisk oplevelse af smerte, hvad der under ingen omstændigheder synes at være belæg for. Det er i én forstand intelligent, at han i fremtiden vil være i stand til at kunne åbne døre via elektroniske impulser fra et implantat i hans arm. At dét skulle åbne døren til bevidstheden, stiller jeg mig tvivlende overfor.

9 Summary of The Inconceivable Thoughts

We know what we are – but not what we may become

Shakespeare

In Ray Kurzweil's book *The Age of the Spiritual Machines* this passage appears as an introduction to a future society in the year 2099. In this society, man- and machine intelligence cannot be distinguished. The possibility of artificial intelligence (AI) is seen as a matter of having sufficient computational power. Kurzweil's thoughts seem to comprehend intelligence and consciousness in an unduly unproblematic way. As many scientists share his views on AI, I find it necessary to investigate whether AI is as unproblematic as it seems to be. The basis for these investigations is the notion of intelligence used by Howard Gardner in his theory of Multiple Intelligences (MI). My hope is that this psychological notion of intelligence will display a different picture on consciousness and as a result of AI.

Artificial intelligence can be divided into two paradigms - classical AI and connectionism. These paradigms have differing opinions on the perspectives of AI. While classical AI claims to synthesize the human consciousness, connectionism takes a more subtle approach, only claiming a possibility of simulation. Classical AI perceives intelligence as a matter of symbolic representation. Symbolic representation exhibits numerous problems that can be seen as results of the attempt to mechanize reason. The Paradox of Mechanical Reason emphasizes that *...if a process or system is mechanical, it can't reason; if it reasons, it can't be mechanical*²⁶⁸. Furthermore, an explanation of consciousness by means of symbolic representation leads to the problematic notion that once syntax is established, semantics will emerge. On the other hand, connectionism does not see representations as a direct symbolic unit, but instead as distributed through the interaction of neurons. This complicated pattern of conceptual representation and their relations forces us to a description of understanding on a subconceptual level, and therefore it does not seem sufficient as an explanation of consciousness. Both paradigms appear to rely on some external criteria for intelligence, and in the end, this results in the reduction of intelligence as a matter of

²⁶⁸ Haugeland 1987a, s.39

behavioral patterns. The question that comes into mind is whether this psychology of the mind, represented by any of the two paradigms, seems adequate when the psychological conception of consciousness as exemplified by Howard Gardner is included. In contrast to the hitherto dominant paper and pencil tests, MI seeks to establish a biological foundation for the notion of intelligence. Gardner sees the brain as a biopsychological unit that processes information. Gardner exhibits eight intelligences that he imagines belong to specific neural parts of the brain. At the same time, he seeks to come to terms with the notion of intelligence as a construct. The notions of intelligence thus give the impression to be blurred. Furthermore, MI presents a hazy notion of the way that meaning comes off. On one hand, meaning occurs as a result of simple manipulation with symbols related to the computational capacities of the mind - on the other hand, Gardner stresses that his theory avoids the philosophical epistemic problems. As we have seen, the notion of symbol manipulation does not conceal these problems. The possibility of AI is rejected on an explanation exhibiting the inconsistency of MI.

On the assumption that the notion of computations is a feature of both AI and MI, a discussion of David Chalmers's (DC) notion of consciousness appears to transpire naturally. With a notion of the mental properties as organizational invariants (OI), the *hard problems* of consciousness are dealt with. Given the properties of OI, we can describe the structure of the brain and due to the principle of structural coherence, we can describe the consciousness. In all, consciousness is only a matter of a functional organization and is not dependent upon the substance in which this organization is implemented. DC imagines that the principles are consolidated into a single theory of information. As such, conscious properties and material properties can be seen as aspects of information. Chalmers' theory brings up a number of matters. In relation to the scope of this paper, I find the discussion on implementation most relevant. As an opponent to DC, John Searle (JS) stresses that the substance exposes an importance in the matter of consciousness. The Chinese Room argument serves to underline the idea that semantic *does not* take care of itself just by implementing a formal rule. Observed from the outside, the answers from within the room give the impression that the room understands Chinese even though this is impossible according to JS. JS does not reject that the brain possesses computational properties but simply emphasizes that this does not sufficiently explain how the brain comes to produce mental states. As an

answer, DC argues that programmes can be seen as syntactic objects, whereas implementations cannot. On the contrary, JS argues that an implementation must possess the same causal abilities as the brain but he does not give any criteria of the way this is to come about. According to JS, the explanation is to be found in the fact that consciousness and mental states are biological phenomena. Mental states are both seen as caused by the neurophysiological processes and as features of the brain. These conceptions reveal a causal interaction and at the same time an identity relation between the mental properties and the brain. However, JS states that the mental states supervene on the material states. While at the same time being seen as an identity relation, the idea of this is obscure. The problem of causality between the mental states and the material states seems to lead either to a kind of dualism and as a result of this to an incomprehensible notion of causality or if rejecting this causal relation, to a form of identity thesis with attendant epiphenomenalism²⁶⁹. Both DC and JS can be characterized as nonreductive physicalists. As such, both theories and as we have seen the AI and MI are faced with the problem of avoiding to reduce the mental states to the material states while simultaneously explaining the psycho-physical relation in a comprehensible manner. Thomas Nagel (TN) beholds the psycho-physical relation in a quite different way. Subjectivity and objectivity are seen as diametrical perspectives. The subjective aspect of the mental states is only comprehensible from the creature itself. Thus, *What it is like to be a bat* cannot be comprehended from a human point of view. Science aims at giving a description of reality as an objective reality. This objective ideal is unification in a point of view from nowhere. The essential subjective character, however, is the specific point of view. As a consequence, we are faced with a need for a new understanding of reality that illuminates instead of eliminates the subjective point of view. Though the objective phenomenology of TN in many ways is problematic, I find his considerations on the objective ideal rewarding in this discussion on consciousness.

With this paper, my initiate aim was to enlighten the discussion of AI through the findings on consciousness in MI. Sad to say, MI does not bring anything new to the discussion. In trying to explain the impossibility of AI, It rather seems that the theory is stretched in such a way that it provides more questions than answers. My investigations have not been in vain. As we have seen, the question of AI embarks on a comprehensible notion on con-

²⁶⁹ Searle 1983, s.264f.

sciousness. As it has turned out, the initiate quote of Shakespeare is by no means obvious. A new perspective is called for. These considerations stress the need for collaboration between different disciplines in a search for an answer. In humble appreciation of the powers of the mind, we must face these future examinations on consciousness.

10 Appendiks

10.1 Brevveksling med Howard Gardner

Følgende brevvekslinger (tre breve – tre svar) er afstedkommet på baggrund af utydeligheder i Howard Gardners teori Multiple Intelligences. For så vidt muligt at undgå misfortolkning udbad jeg ham svar på en række centrale områder.

Mine breve er i det følgende gengivet med normal skrift mens Howard Gardners svar er fremhævet med fed skrift. Brevene er ordnet efter kronologisk rækkefølge, hvorfor hvert enkelt svar kan ses som respons på det forestående brevets spørgsmål.

I'm a Danish student studying trying to finish my degree in philosophy.

My paper is concerning Artificial intelligence. I'm trying to find out if a more broad definition of intelligence will have affect on the findings in cognitive science and Artificial intelligence.

In my project I am using your findings, represented through MI, as a major turning point in the concept of intelligence.

I have recently stumbled across some problems concerning your theory - and I would appreciate if you could help me clear this matter up?

In Intelligence Reframed you say:

“Considerable evidence suggests that such executive functions are handled by structures in the frontal lobe. The “modeler” must then decide whether to consider this a separate intelligence, or an entity that emerges from other intelligences, such as the intrapersonal intelligence. At present, I lean toward the latter alternative.”

p.105 (US edition, 1999)

What is your distinction between subintelligences and these referred executive functions?

Are these executive functions working on behalf of all intelligences?

Earlier on the same page you are referring to the advantages of a theory that does not posit a single executive function - due to the homunculus question.

How do you explain avoiding the homunculus question when you at the same time posit a number of executive functions?

In appreciation of your help.

Yours sincerely

Torsten Bo Hansen

Thanks for your query. I wish that I could give you a simple and straightforward answer. There is no question that human beings stand out in their capacity to plan, and to engage in meta-cognitive activity. And so, for you as an AI student, and for me as a psychologist, the question is how best to model these activities. Robert Sternberg, David Perkins and others interested in "thinking skills" would have no problems in simply asserting that we have such an executive intelligence or capacity. Nor do I think that they feel that they are giving into huncular thinking. I have been influenced by the writings of Annette Karmiloff-Smith (see BEYOND MODULARITY). She argues that as we become more competent in any domain, we eventually acquire the capacity to reflect upon that domain. Metacognitive capacity in one domain says nothing about comparable capacities in other domains. I am more sympathetic toward that line of argument, which seems less general and less homocular to me. You can see that I am still uncertain about these issues and would welcome your own thoughts.,
With best wishes, Howard Gardner

Dear Mr. Gardner

I'm very thankful to receive your reply in such a short time.

I'm not quite sure if you understood the deeper meaning with my questions - My fault.

My concern is how you can account for the problem of consciousness?

You say that your theory is concerning the entire human mind - a complete cognitive theory in other terms.

Due to my philosophical perspective I seek to find out how you make a distinction between subintelligences and these 'executive' functions and if these functions are seen as correlated to the human conscious sphere.

Is your concept of 'executive' functions only meant to be referred to the intrapersonal intelligence - or do you operate with a number of 'executive' functions referring to each specific intelligence? Or do you conceive of the metacognitive skills as a kind of unified layer behind all the different intelligences providing the conscious aspect of the human mind?

Does these 'executive' functions have a conscious aspect or is your concern only directed at one side of the computational aspect - the physical?

Do you conceive the 'metacognitive' skills as equivalent to your own terminology of 'executive' functions?

I'm aware that you say that the problem of consciousness does not arise - I'm sorry to say that I disagree. Your own concept of the homunculi problem is as I see it is based on a notion of consciousness directly linked to these 'executive' functions. If by 'executive' you only mean the pure computational capacity and thereby exclude the consciousness aspect - you never get a homunculi problem. So when you do address this problem I must think that you do include the consciousness aspect. And then I'm really confused - so my last question is:

How do you account for avoiding the problem of the homunculi when you operate with a number of executive functions?

All in all - if you have a notion of how consciousness is to be related to your theory I would be more than willing to lend you an ear.

Once again I'm very thankful.

Yours sincerely

Torsten Bo Hansen

I did not realize that you expected of me a whole cognitive /mind theory. While I have a certain ambition, that is well beyond my scope or capacity. I will try to give you a brief and honest answer to the two principal questions you raise, though I know neither response is adequate:

1. As individuals become expert in various domains, they do become explicitly aware of what they are doing. So in that sense, they are conscious of their activities. This form of consciousness has a domain-specificity which the term "executive function" usually denies. What we usually call Consciousness- with a capital C--does seem to me to be an outgrowth of Intrapersonal intelligence, in that we cant be Conscious without a sense of the "I".

2. I have never written about Consciousness/consciousness because I dont feel that I have enough of a grasp on the issue to make a contribution. My own intuition is that Consciousness arises because humans, unlike other species, can represent information in a number of ways. The consistency and inconsistency in these modes of representation are what "nudges" us up

to consciousness. Put differently, if humans were restricted to one mode of representation, however powerful, I don't think that they would attain Consciousness.

Again, these are only impressions. Given my current preoccupations (I am studying Good Work and will in fact be attending a conference on that topic in Denmark at the end of January-- for details, contact knoop@dpu.dk), I don't think that I will ever have the time or motivation to turn my attention to the questions that you raise, and I hope that you will be able to make discernible progress.

With best wishes, Howard Gardner

Dear Mr. Gardner

A couple of questions has entered my mind. I hope that you will be so kind to offer a few moments on these questions.

First:

You say that AI is a category error and computers will never be able to participate in the sphere of the interpersonal intelligence. Why is this intelligence outstanding? Every intelligence is valued in a community (your argument against AI)but you don't seem to have any problem ascribing mathematical-logical intelligence to computers.

Second:

In 'Intelligence Reframed' You say that the realm of the spiritual is concerning matters that we will not allow the same ontological status as the other intelligences. Do you hereby conceive the status of the intelligences as ontological?

Yours sincerely

Torsten Bo Hansen

I see one more set of questions from you. You are certainly stretching me in ways that are unusual for me. I fear I don't have time to engage in more discussions of these issues, and I concede that is my failing. Brief comments:

1. I single out intrapersonal intelligence as distinctly human-- and not replicable by AI. I do this because intrapersonal intelligence grows in

part out of emotional reactions which, I believe, are not undergone by computers as we now know them. In contrast, emotions are irrelevant to logical-mathematical reasoning, as I understand it.

2. My remark about the unique nature of spirituality stems from my belief that it deals with realms that are in principle unknowable. But you are right to indicate that, for many people, the spiritual is as real as the musical or the logical mathematical. It's just that we can reach consensus on music or logic in ways that seems inherently impossible for spiritual matters-- faith and reason are incommensurate.

With best wishes.

Howard Gardner

11 Litteraturliste

11.1 Bøger

Bernsen, Niels Ole og Ulbæk, Ib: *Naturlig og kunstig intelligens*. Nyt Nordisk Forlag Arnold Busck A/S. København, 1993.

Bertelsen, Poul René: *Hjernens filosofi* (kompendium). Odense Universitet Trykkeri. Odense, 1999.

Boden, Margaret A.: *Minds and Mechanisms – Philosophical psychology and computational models*. The Harvester Press Limited. Sussex, 1981.

Dancy, Jonathan & Sosa, Ernest (ed.): *A Companion to Epistemology*. Blackwell Publishers Ltd., Oxford, 1998.

Dennett, Daniel C.: *Kinds Of Minds*. Weidenfeld & Nicolson. London, 1996.

Dreyfus, Hubert L.: *What Computers Can't Do – The limits of artificial intelligence*. Revised edition. Harper Colophon Books, 1979.

Dreyfus, Hubert L & Dreyfus, Stuart E.: *Mind Over Machine. The power of human intuition and expertise in the era of the computer*. Basil Blackwell Ltd., 1986.

Gade, Anders: *Hjerneprocesser – kognition og neurovidenskab*. Frydenlund Grafisk. København, 1997.

Gardner, Howard: *Frames of Mind*. Second edition. Fontana Press. London, 1993a.

Gardner, Howard: *Multiple Intelligences - The theory in practice*. Basic Books. New York, 1993b.

Gardner, Howard: *Intelligence Reframed – Multiple Intelligences for the 21st Century*. Basic Books. New York, 1999.

Gleerup, Jørgen: *Organisationskultur som læreproces og kommunikation*. Odense Universitetsforlag, 1998.

- Goleman, Daniel: *Følelsernes intelligens (Emotional Intelligence)*. 2. udgave, 1. oplag. Borgens Forlag. Gylling, 1997.
- Gorzalak, Flemming: *Det filosofiske begrundelsesproblem*. Phd.-afhandling. Odense Universitet. Filosofisk institut, 1995.
- Harnish, Robert M.: *Minds, Brains, Computers – An Historical Introduction to the Foundations of Cognitive Science*. Blackwell Publishers. Oxford, 2001. – (Af bogen fremgår det, at den først er udgivet i år 2002. Kan det tolkes som et udtryk for at computere er intelligente og ved at overtage magten...? Det tror jeg ikke! Jeg tror nærmere på en menneskelig fejl og har derfor tilladt mig at rette det til 2001.)
- Hartnack, Justus & Sløk, Johannes (red.): *De store tænkere – Descartes*. 2. udgave. Munksgaard. København, 1991.
- Haugeland, John: *Having Thought*. Harvard University Press. Cambridge, 1998.
- Haugeland, John: *Artificial Intelligence – The very idea*. 3. printing. Bradford Books. The MIT Press, Massachusetts, 1987a.
- Haugeland, John. (edt.): *Mind Design*. 5. printing. A Bradford Book. The MIT Press. Massachusetts, 1987b.
- Kirkeby, Ole Fogh: *Kroppens tanke – om kunstig intelligens – det umuliges kunst*. Samfundslitteratur. 1988.
- Klawonn, Erich G.: *Jeg'ets ontologi – en afhandling om subjektivitet, bevidsthed og personlig identitet*. Odense Universitetsforlag. Odense, 1991.
- Kurzweil, Ray: *The Age of Spiritual Machines*. Orion Business Books. London, 1999.
- Lehtovaara, Arvo: *Lille lærebog i psykologi*. Nyt Nordisk Forlag. København, 1971
- Luhmann, Niklas: *Social Systems*. Stanford University Press. Stanford, Californien, 1995.
- Nagel, Thomas: *Other Minds*. Oxford University Press. New York, 1995.
- Nagel, Thomas: *Mortal Questions*. Cambridge University Press, New York, 1979.

Penfield, Wilder: *The Mystery of the Mind – a critical study of consciousness and the human brain*. Princeton University Press. Princeton, New Jersey, 1975

Peter Menzel and Faith D'Aluisio: *Robo Sapiens*. MIT Press. Massachusetts, 2000.

Qvortrup, Lars: *Det Hyperkomplekse Samfund*. 2. udgave. Gyldendal. København, 2000.

Searle, John R.: *The Rediscovery of the Mind*. Third printing. The MIT Press. Massachusetts, 1992.

Searle, John R.: *The Mystery of Consciousness*. The New York Review of Books. New York, 1997.

Searle, John R.: *Minds, Brains and Science*. British Broadcasting Corporation. London, 1984.

Wittgenstein, Ludwig: *Philosophische Untersuchungen*. 10 oplag, Werkausgabe Band 1. Suhrkamp Taschenbuch Verlag. Frankfurt am Main, 1995.

Ølgaard, Bent: *Kommunikation og økomentale systemer ifølge Gregory Bateson*. 2. udgave, 2. oplag. Akademisk Forlag. København, 1993.

11.2 Artikler

Chalmers, David J.: Facing Up to The Problems of Consciousness.

<http://www.u.arizona.edu/~chalmers/papers/facing.html>

Chalmers, David J.: *What is a Neural Correlate of Consciousness?*

<http://www.u.arizona.edu/~chalmers/papers/ncc2.html>

Chalmers, David J.: *The Puzzle of Conscious Experience*

<http://www.u.arizona.edu/~chalmers/papers/puzzle.html>

Chalmers, David J.: *On The Search for The Neural Correlate of Consciousness*.

<http://www.u.arizona.edu/~chalmers/papers/ncc.html>

Chalmers, David J.: *A Computational Foundation for the Study of Cognition*.

<http://www.u.arizona.edu/~chalmers/papers/computation.html>

Chalmers, David J.: *Absent Qualia, Fading Qualia, Dancing Qualia*

<http://www.u.arizona.edu/~chalmers/papers/qualia.html>

Chalmers, David J. & Bayne, Tim: *What is the Unity of Consciousness?*

<http://www.u.arizona.edu/~chalmers/papers/unity.html>

Chalmers, David J.: Availability: The Cognitive Basis of Experience?

<http://www.u.arizona.edu/~chalmers/papers/availability.html>

Gelder, Tim Van: *Monism, Dualism, Pluralism.*

<http://www.arts.unimelb.edu.au/~tgelder/papers/MDP.html>

Haugeland, John: *The Nature and Plausibility of Cognitivism.* The Behavioural and Brain Sciences, vol.1 (1978). Cambridge University Press. Artikel fra: Haugeland, John. (edt.): *Mind Design.* 5. printing. A Bradford Book. The MIT Press. Massachusetts, 1987.

Jackson, Frank: *Epiphenomenal Qualia.* Philosophy Quarterly vol.32, 1982. s.127-136.

McCarthy, John: *Ascribing mental qualities to machines.*

<http://www-formal.stanford.edu/jmc/ascribing/ascribing.html>

McCarthy, John: *Artificial Intelligence, Logic and Formalizing Common Sense.*

<http://www-formal.stanford.edu/jmc/ailogic/ailogic.html>

McCarthy, John: *A logic AI approach to context.*

<http://www-formal.stanford.edu/jmc/logical/logical.html>

McCarthy, John: *Free will even for robots.*

<http://www-formal.stanford.edu/jmc/freewill/freewill.html>

McDermott, Drew: *Artificial Intelligence Meets Natural Stupidity.* SIGART Newsletter. No. 57 (1976). Artikel fra: Haugeland, John. (edt.): *Mind Design.* 5. printing. A Bradford Book. The MIT Press. Massachusetts, 1987.

- Nagel, Thomas: *Subjective and Objective*. Artikel fra: Nagel, Thomas: *Mortal Questions*. Cambridge University Press, 1979.
- Nagel, Thomas: *What is it like to be a bat?* Artikel fra: Nagel, Thomas: *Mortal Questions*. Cambridge University Press, 1979.
- Putnam, Hillary: *Reductionism and the Nature of Psychology*. *Cognition*, 2 (1973). Artikel fra: Haugeland, John. (edt.): *Mind Design*. 5. printing. A Bradford Book. The MIT Press. Massachusetts, 1987.
- Salovey, Peter & Sluyter, David J.: *What Is Emotional Intelligence?*. Artikel fra: Salovey, Peter & Sluyter, David J. (red.): *Emotional Development and Emotional Intelligence*. BasicBooks. New York, 1997.
- Searle, John R.: *Is the Brain's Mind a Computer Program?* *Scientific American* (1990), 262, nr.1, s.20-25
- Searle, John R.: *Minds, Brains and Programs*. *The Behavioural and Brain Sciences*, vol. 3. 1980. Cambridge University Press. Artikel fra: Dennet, Daniel C. & Hofstadter, Douglas R (edt.). *The Mind's I – Fantasies and Reflections on Self and Soul*. Penguin Books. Basic Books, London 1982.
- Shannon, Claude E.: *A Mathematical Theory of Communication*. *Bell Systems Technical Journal* 1948, nr. 27. s.379-423.
- Turing, Alan: *Computing Machinery and Intelligence*. *Mind* 1950, vol. LIX, no. 236. Artikel fra: Dennet, Daniel C. & Hofstadter, Douglas R (edt.). *The Mind's I – Fantasies and Reflections on Self and Soul*. Penguin Books. Basic Books, London 1982.
- Weng, Juyang m.fl.: *Autonomous Mental Development by Robots and Animals*. *Science* vol. 291, 2001. s.599-600.