

Det ubevisste i lys av nyere teori og forskning

Rolf Marvin Bøe Lindgren

Semesteroppgave høsten 1992

Sammendrag

Det er i den senere tid skjedd en fornyelse i interessen for det ubevisste. Dette er kanskje særlig i forbindelse med PDP-teorier om bevissthet, som kan gi et utvidet perspektiv på studiet av det ubevisste, både i relasjon til emosjoner, og til psykonalytisk forståelse av det ubevisste. Det later til at det ubevisste er mer omfattende enn tidligere antatt, og også at det ubevisste tar en større del i rasjonell tenkning enn tidligere antatt.

Innhold

Innledning	1
Hva er det ubevisste?	1
Bakgrunn	2
Freuds ubevisste	2
Broadbents filtermodell	4
The Hidden Observer	5
Modeller fra «cognitive science»	5
Ubevisst kognitiv prosessering	7
Emosjoner og det ubevisste	10
Fortrengning og det kognitivt ubevisste	15
Konklusjon	16
Tillegg: Church-Turing tesen	18
Beregnbarhet	18
Kompleksitet	18

Innledning

“Education is what remains when we have forgotten all that we have been taught.”

– George Saville, Marquis of Halifax (1633-1695).

Hva er det ubevisste?

Visse mentale prosesser kan påvirke vår bevissthet og være av betydning for våre handlinger. Disse prosessene kan vi vanskelig, eller ikke i det hele tatt, gjøre verbalt rede for, eller på annet vis hente fram eller få innsikt i. En fellesbetegnelse på disse prosessene er *det ubevisste* ((Kihlstrom 1990, s. 446 og 448); (Lazarus 1991, s. 361)).

Den empiriske psykologiens opprinnelige prosjekt var studiet av de mentale prosesser som er tilgjengelig for bevisstheten. Wundt, Titchener, og andre strukturalister som var med på å legge grunnlaget for den tidligste psykologiske forskning antok at sinnet¹ rent generelt er i stand til å observere sitt eget indre arbeid (Kihlstrom 1990, s 445).

Imidlertid fantes det samtidig en sterk filosofisk tradisjon som fokuserte på mentale prosesser som *ikke* er tilgjengelige for bevisstheten. Arthur Schopenhauer (Schopenhauer, 1819, referert i Hogan (1985, s. 48–49)) hevdet at mennesker er grunnleggende irrasjonelle, og at de er drevet frem av krefter i sitt indre som de bare vagt er klar over. Den viktigste av disse er trangten til formering. Nietzsche (*ibid.*) hevder at enhver følelse, holdning, handling og personlig prinsipp² har sitt grunnlag i selvbedrag og er tuftet på en ubevisst løgn—de er sublimerte uttrykk for seksualitet og aggressivitet.

Freud hevdet at vårt bevisste mentale liv er bestemt av ubevisste idéer samt impulser og emosjoner, såvel som forsvarsmekanismer som ubevisst blir iverksatt mot dem. Helmholtz hevdet, på den annen side, at bevisst persepsjon er produktet av ubevisste slutninger basert på individets kunnskaper om verden og minner om tidligere erfaringer, og at f.eks. visuell persepsjon delvis er må være basert på rasjonelle, ubevisste slutninger som gir seg bevisst uttrykk i f.eks. dybdesyn (Johnson-Laird 1988, s. 15, s. 354). Helt siden dengang har studiet av det ubevisste foregått i to forskjellige leier. Det dynamisk og det kognitivt ubevisste er blitt studert hver for seg

¹Jeg oversetter «the mind» med «sinnet»

²expressed virtue

og uavhengig av hverandre (Boston, 1986, referert i Kihlstrom (1990, s. 446)).

Spørsmålet som har opptatt forskere og tenkere siden dengang er hvordan tanker og impulser blir og forblir ubevisste. Jeg vil konsentrere meg om forklaringsmodeller fra dynamiske og kognitive tradisjoner.

I det følgende vil jeg konsentrere meg om det *dynamisk ubevisste* slik det omtales i bl.a. Pervin (1989) og i Gullestad (1992), og det *kognitivt ubevisste* slik det omtales i bl.a. Kihlstrom (1990) og i Lazarus (1991).

Jeg ønsker å ta spesielt for meg PDP-modeller fordi det her vært en del diskusjon rundt dem i de senere år, og fordi jeg mener at det er viktig å vise det historiske og formelle grunnlaget for PDP-modellene dersom man skal drøfte forholdet mellom PDP-modeller og det ubevisste.

Selv om psykologer har vært med på å forme «cognitive science», f.eks. R. Newell eller R. P. Abelson, så er innflytelsen av logikere og matematikere som Frege, Gödel, Wittgenstein, Turing, Shannon og Russel sterk. En relativt teoretisk gjennomgåelse er lagt ved som etterskrift.

I det følgende vil jeg ta for meg det ubevisste i lys av nyere teori og forskning. Det jeg spesielt ønsker å belyse, er følgende:

1. Hvilke oppgaver har det ubevisste?
2. Hva er forholdet mellom det ubevisste og emosjoner?
3. Hva er forholdet mellom det dynamisk ubevisste og det kognitivt ubevisste?

Bakgrunn

Freuds ubevisste

Freud så på det mentale som et energisystem, hvor energien strømmer, ledes ut på sidespor, eller blir demmet opp. Energien er begrenset, slik at dersom mye energi blir brukt på en ting blir det tilsvarende mindre energi til overs til andre ting. All adferds mål er nytelse, i form av redusert spenning eller frigjøring av energi (Pervin 1989, s. 77).

Det mentale liv kan beskrives i henhold til i hvilken grad vi er klar over fenomener: Det *bevisste* henviser til hva vi er klar over i enhver gitt situasjon, det *forbevisste* henviser til det vi ikke for øyeblikket er oppmerksomme på men som vi kan bli bevisst dersom vi leder vår oppmerksomhet mot det, og det *ubevisste* henviser til de fenomenene som vi ikke er klar over og

som vi ikke kan bli klar over unntatt under visse helt spesielle betingelser (Pervin 1989, s. 80).

Det ubevisste er alogisk og tar ikke hensyn til tid eller rom. Det blir ikke lagt merke til gjennom mesteparten av vårt rasjonelle, våkne liv, men kan bli gjort tydeligere under drømmer eller i idéene til psykotikere.

Freud fant evidens for at det ubevisste i sitt kliniske arbeid. Pasienter ble, i løpet av den terapeutiske prosessen, klar over minner og ønsker som de tidligere ikke hadde vært klar over. Slike oppdagelser ble gjerne assosiert med svært vonde følelser. Dette satte Freud på tanken at det ubevisste inneholder minner og ønsker som ikke er en del av vår bevissthet, men «aktivt gjemt» i det ubevisste (Pervin 1989, s. 83).

The theory of repression is the corner-stone on which the whole structure of psycho-analysis rests (*Freud, 1914, referert i Gullestad (1992, s. 29)*)

Fortrengningsbegrepet i dag har utviklet seg siden Freuds tid. Det antas fortsatt at fortrenghing er et forsvar for ego-et mot *indre fare*, analogt med at flukt er et forsvar mot *ytre fare*. (Gullestad 1992, s. 815) Etter hvert utviklet Freud fortrenghingsbegrepet til å bety mer enn indre forsvar. I Freuds senere strukturteorier får *angst* en nøkkelrolle som motiv for fortrenghing: Angst innebærer antisipering av en faresituasjon, som er definert ved *hjelpeløshet* (Freud, 1923, 1926, referert i Gullestad (1992, s. 816)).

Selve prinsippet for fortrenghingen er hele tiden den samme: individet nekter å vedkjenne seg noe som er for angstfremkallende til å kunne tåles. Derfor blir det angstfremkallende fortrenghet, i det individet skjuler det ubehagelige for seg selv (Gullestad 1992, s. 816). Dette innebærer et paradoks, formulert slik av Sartre:

For å utøve sin virksomhet med nøyaktighet må sensoren *kjenne til* det som skal fortrenghes. Sensoren må *ta stilling* til hva som er akseptabelt og kan passere, og hva som er uantakelig og må fortrenghes. Men hvordan er det mulig at sensoren både kan vite om og ikke vite om en ting på samme tid? Hvordan kan den bestemme seg for ikke å vite noe den vet? (*Sartre, 1943, referert i Gullestad (1992, s. 818)*)

Dersom man ser på fortrenghing som *motivert*, dvs. ikke som et uttrykk for kausale prosesser er ikke dette lenger noe problem, hevder Gullestad (1992, s. 818). Det er ikke slik at man «bestemmer seg» for å fortrenghet det utillatelige. Ordet «bestemme» er i så fall ment metaforisk (*ibid.*). Men

hvordan kan selve fortrenningen være ubevisst? Gillet, 1987, referert i Gullestad (1992, s. 819) løser dette ved å postulere *to* ubevisste: i tillegg til det *dynamisk* ubevisste eksisterer bevisste strukturer som er ubevisste i *deskriptiv* forstand, som slett ikke kan oppleves og aldri gjøres tilgjengelige for bevisstheten. (*ibid.*)

Slik som dette er beskrevet i Gullestad (1992) forekommer det deskriptivt ubevisste som noe *ad hoc*. Det er klart at kognitive, biologiske prosesser ikke noensinne kan gjøres tilgjengelige for det bevisste, men hvorfor skulle fortrenningsmekanismene være der, dersom de ikke kan vises å ha en direkte biologisk oppgave?

Jeg vil i det følgende presentere teori og forskning som utvider synet på hvilke oppgaver det ubevisste kan ha, og hvor det ikke er usannsynlig at fortrenningsmekanismer kan ha en plass sammen med andre ikke-bevisste kognitive prosesser.

William James var muligens den av det 19. århundres akademiske psykologer som kraftigst argumenterte for det ubevisste mentale liv (Kihlstrom 1990, s. 448). James hevdet at mentale tilstander kan være ubevisste på minst to måter. For det første kan en mental begivenhet bli aktivt utelatt fra oppmerksomheten, for det andre argumenterte han for en inndeling av de bevisste i primære og sekundære osv. bevisstheter. Bare en av disse er tilgjengelig for fenomenologisk oppmerksomhet på et gitt tidspunkt. Han foretrakk å snakke om «sam-bevisste» eller «underbevisste»³ mentale tilstander heller enn ubevisste for å unngå selvmotsigelser⁴.

Broadbents filtermodell

Det er dagligdags erfaring at vi er nødt til å konsentrere oss om en liten del av alt vi *kan* være oppmerksom på. Vi får ikke med oss alt vi ser på det tidspunktet vi ser det. Men det vi ikke er oppmerksom på, kan det likevel påvirke oss?

En tidlig idé er at stimuli som oppmerksomheten rettes mot, blir filtrert bort på et relativt tidlig trinn i vårt oppmerksomhetssystem (Broadbent, 1958, referert i Atkinson, Atkinson, Smith & Hilgard (1987, s. 198)). Det har senere vist seg at en det må finne sted en god del prosessering i forkant av et utvalg ettersom et budskap kan forkastes eller avvises på grunn av sin betydning. Mye tyder på at vårt persepsjonssystem tar vare på materiale

³co-conscious, subconscious

⁴Jeg kan tenke meg at noe av motviljen mot å bruke ordet «unconscious» i et engelsktalende miljø er at det samme ordet «unconscious» betyr både «ubevisst» og «bevisstløs». At en kognitiv prosess kan være *bevisstløs* virker jo urimelig. Jeg tar for gitt at man er klar over at det er en forskjell, og vil ikke ta opp dette som en del av den videre diskusjonen.

som ikke er i oppmerksomheten i en kort periode og at det blir prosessert til en viss grad.

Ens eget navn kan godt fanges opp, selv om det blir ytret relativt svakt i en samtale som man ikke følger med på (det såkalte «cocktail party phenomenon»). Dette ville ikke kunne skje dersom hele samtalen går tapt på lavere nivå i persepsjonssystemet (Atkinson *et al.* 1987, s. 111).

Denne utvelgelsen er ikke bevisst, men det er ikke rimelig å anta at utvelgelsen er fortrenget. Det er derfor grunn til å anta at prosesseringen er ubevisst, men ikke ubevisst i dynamisk forstand.

The Hidden Observer

Hebb (1982, referert i Atkinson *et al.* (1987, s. 140)) beskriver en hypnosituasjon hvor den hypnotiserte ble gitt hypnotisk instruksjon om at han var døv så lenge Hilgard ikke hadde sin hånd på den hypnotisertes skulder. Denne personen var også blind, og med en hypnotisk induert døvhet var han nokså fullstendig avskåret fra det som foregikk rundt ham. Hilgard ble spurt fra salen om den hypnotiserte var så uresponsiv som han lot til å være. Hilgard ba da den hypnotiserte om å løfte fingeren sin dersom det faktisk var «noen del av ham» som var i stand til å høre. Til alles overraskelse, også den hypnotisertes, løftet fingeren seg.

Dette blir tatt som evidens for at det eksisterer en «skjult observator», en mental struktur som overvåker alt som skjer, også ting som blir oppfattet uten at den hypnotiserte er klar over det (Atkinson *et al.* 1987, s. 141).

Dette tar jeg til inntekt for at det eksisterer prosesser som vi ikke er bevisst, men som likevel kan påvirke hva vi gjør. Disse prosessene kan være mentale prosesser, de er i hvert fall i stand til å håndtere meningsbærende materiale. Men i hvor stor grad kan ubevisste prosesser bidra til hva vi gjør? Nyere teori og forskning tyder på at det kan være ganske mye.

Modeller fra «cognitive science»

«Artificial Intelligence» (AI), en viktig del av «cognitive science», har som sitt uttalte prosjekt er å lage en maskin som, med utgangspunkt i teorier om hvordan mennesker tenker, selv kan tenke og ha bevissthet (Penrose (1989, s. 3); Searle (1987, s. 18); Johnson-Laird (1988, s. 26)). Hjørnesteinen i AI er teorien om en relativt enkel matematisk modell som kan beregne enhver beregnbar matematisk oppgave, i kraft av at den kan håndtere relativt enkle symboler etter bestemte regler.

Hva er interessant med en maskin som kan oppnå *bevissthet* i en diskusjon om *det ubevisste*? Jo, det er blitt foreslått at den nyeste forskningen innen AI ikke dreier seg så mye om teoretisering rundt det bevisste som teoretisering rundt *det ubevisste*.

En relativt formell oppsummering av det teoretiske grunnlaget for denne tenkningen er lagt som vedlegg (se side 18). Utledningen er interessant, fordi den førte til en teori om en maskin som kunne beregne almenne prinsipper for enhver maskin som kan foreta beregninger (Nørretranders (1991, s. 78); Penrose (1989, s. 51); Johnson-Laird (1988, s. 51)). Den førte også til teorier om begrensninger for hva som i det hele tatt kan beregnes, og om det er slik at et faktisk, fysisk objekt (f.eks. en hjerne) kan bli beskrevet i form av beregnbare matematiske strukturer (Penrose 1989, s. 50).

Det som er beskrevet ovenfor er den såkalte *sekevensielle* Turing-maskin, som utfører sine beregninger i rekkefølge, én etter én. Den er heftet med to svakheter som kan formuleres som en rimelig kritikk mot sammenligningen av hjernen med en Turing-maskin: Den sekvensielle Turing-maskinen

- er langsom
- blir ødelagt av selv den minste skade

(Clark 1987, s. 5)

Konneksjonist-systemer (heretter PDP-nettverk etter Parallel Distributed Processing) ble lansert som en alternativ fremgangsmåte for å teste hypoteser om kognitive prosesser. Virkemåten deres kan illustreres ved å bruke det åpne marked som metafor: globale mønstre av tilbud og etterspørsel blir etablert i henhold til lokale interaksjoner mellom kjøp og salg. Overordnet kunnskap om etterspørsel blir på denne måten spredd mellom kjøperne, og overordnet kontroll av varer mellom produsentene (Clark 1987, s. 4).

I sitt verk «*The Emperor's New Mind*» hevder den britiske fysikeren Roger Penrose at et av hovedproblemene med AIs prosjekt er at vi ikke kjenner fysikkens lover godt nok til at vi kan forstå hvordan vi skal bære oss at med å fremstille en tenkende maskin (Penrose 1989, s. 16). I avsnittet «*Parallel computers and the 'oneness' of consciousness*» skriver han

Many people appear to be of the opinion that the development of *parallel* computers hold the key to building a machine with the capabilities of a human brain. [] A parallel computer, as opposed to a serial one, has a very great number of separate calculations carried out independently, and the results

of these largely autonomous operations are only intermittently combined together to give contributions to the overall calculation. The motivation for this type of computer architecture comes largely from an attempt to imitate the operations of the nervous system, since different parts of the brain indeed seem to carry out separate and independent calculational functions (e.g. with the processing of visual information in the visual cortex) (Penrose 1989, s. 389)

En karakteristisk ting ved bevisst tenkning er at den sentrerer rundt en ting av gangen. Det blir ofte sagt at «Jeg kan ikke tenke på mer en en ting av gangen!» Å tenke på flere ting av gangen, samtidig, uavhengig og uforstyrret av hverandre, er umulig, sier Penrose (1989, s. 399) Den parallelle datamaskin, hevder han, tjener bedre som et bilde på *det ubevisste!*

Et eksempel på et PDP-nettverk er Boltzmann-maskinen (Ackley *et al.*, referert i Clark (1987, s. 4)). Den er satt sammen av mange enkle enheter som opererer parallelt og som er koblet sammen med sine naboer ved lenker hvor informasjonsstrømmen kan gå begge veier. Clark (1987) foreslår at hver enkel enhet kan betraktes som et primitiv hypotese om et eller annet målområde. To enheter som står for motstridende hypoteser kan på denne måten ses på som en negativt vektende forbindelse. Hypoteser som støtter hverandre er en positiv forbindelse, osv. Disse enhetene vil påvirke og vil bli påvirket av andre enheter igjen, slik at alle enhetene tilsammen vil justeres mot en felles avgjørelse. Denne prosessen blir kalt *relaxation* (Clark 1987, s. 4).

PDP-nettverket har to fortrinn. Det første er at PDP-nettverkene er implementert på en slik måte at prosesseringen går mye raskere fordi systemet kan deles opp i små uavhengige prosesser som ikke behøver å vente på hverandre for å bli ferdig i den grad et sekvensiell maskin må. Dessuten er PDP-nettverk langt mer robuste. Dersom man fjerner en del av et godt konstruert PDP-nettverk så slutter det ikke å fungere, det blir bare svekket i sin funksjonalitet, kanskje på samme måten som hjernen blir svekket med alderen. Dette fenomenet er kjent som *graceful degradation* (Clark 1987, s. 6).

Broadbent, 1985, referert i Clark (1987, s. 9) hevder at PDP-nettverk bare er en av mange hensiktsmessige måter å beskrive kognitive prosesser på. Han hevder at PDP-nettverk ikke har psykologisk relevans per se. McClelland & Rumelhart, referert i Clark (1987, s. 9) går imot dette, og hevder at det nivået hvor beregningene foregår også er teoretisk relevant.

Utviklingsmessig, spekulerer Clark (1987, s. 9), kan det ha foregått slik at hjernen opprinnelig var et PDP-nettverk, med alle fordelene dette inne-

bærer: slutninger skjer raskt og «maskineriet» som behandler slutningene er relativt robust og gir god grobunn for videre utvikling. Men sekvensielle Turing-maskiner har også sine fordeler, f.eks. at de er mer nøyaktige. Det er lettere å beskrive formell resonering i en ren Turing-maskin. Kanskje er det slik, konkluderer Clark (1987, s. 12), at hjernen er et PDP-nettverk som bruker Turing-maskiner til de mer spesialiserte, høyere-ordens prosessene. Dette beror på om PDP-nettverk er i stand til å utføre sekvensielle prosesser på samme måte som mer tradisjonelle serielle Turing-maskiner.

Ubevisst kognitiv prosessering

En god del av mental aktivitet er ubevisst i den forstand at den ikke er gjenstand for fenomenologisk viten under noen omstendigheter. Visse ubevisste kunnskaper synes å være medfødt, annen kognitiv og motorisk kunnskap som blir tilegnet via erfaring ser ut til å bli gjort til rutine via øving. Denne prosessen kan beskrives som «kunnskapskompilering», noe som tilsier at det formatet som kunnskapen er representert i er blitt endret. På denne måten kan både medfødt og tillært kunnskap sies å være ubevisst (Anderson, 1982, Nisbett & Wilson, 1977, referert i Kihlstrom (1990, s. 448–449)).

Hva skal til for at adferd kan sies å være automatisert, i motsetning til intendert? Det viktigste er at den krever liten eller ingen kognitiv kapasitet, slik at den kan utføres samtidig med viljestyrte handlinger. Et eksempel på avansert automatisert kognitiv prosessering er gitt i «*Mærk Værden*»:

Moderne fotball bærer præg af mange og indøvede mønstre. De virkelige genier som Pelé, Cruyff, Netzer, Maradona, Landrup sprænger hele tiden mønstrene. Det er derfor de er så gode. (*Nørretranders 1991, s. 312*)

Dersom fotballspillerne måtte tenke seg om foran hver avgjørelse ville de ikke ha tid til å spille så godt som de gjør, hevder han. Slutningene skjer raskt, presist og — ubevisst.

Det har vært hevdet (Pierce & Jastrow, 1884, referert i Kihlstrom (1990, s. 450)) at en hendelse ikke kan bli analysert for mening uten at den har vært bevisst identifisert og betraktet. Det er blitt vist at dette ikke nødvendigvis er riktig. Marcel, 1983a, 1983b, referert i Kihlstrom (1990, s. 450), brukte en leksikalsk avgjørelsesoppgave hvor et stimulus-ord (prime) ble fulgt av et annet stimulus-ord (target). Forsøkspersonene skulle avgjøre om målet er et meningsfullt ord. Bedømmelsene ble fasilitert dersom prime også var et meningsfullt ord. Marcel fulgte sine primes med maskerende stimuli, slik at forsøkspersonene ikke pålitelig kunne oppdage primes. Likevel

hadde primingen som effekt at avgjørelsene ble fasiliterte i den leksikalske avgjørelsesoppgaven. Siden semantisk priming krever noen grad av semantisk prosessering, ser det ut til at analysering av mening kan finne sted på stimuli som er holdt utenfor bevisst viten.

Det ser ut til at effekten av semantisk priming kan reflektere tilegning av ny kunnskap, eller bare aktivering av eksisterende kunnskap. Når forsøkspersoner studerer lister av assosiasjonspår, som f.eks. «sour-grapes» og «small-potatoes», og så blir bedt om å gi det første ordet de kan komme på når cue er «sour» eller «small», så viser forsøkspersoner priming av targetrepsjonen uavhengig av i hvilken grad de er i stand til å gjenkalle ordparene de allerede er presentert for. Siden slike fraser som «sour grapes» eller «small potatoes» er vanlige engelske uttrykk, så ser primingen ut til å vise aktivering av kunnskap som alltid er lagret i semantisk hukommelse.

Det ser ut til, slutter Kihlstrom (1990, s. 451), at implisitt persepsjon er gjenstand for de samme føringene, dvs. at hendelser i de aktuelle omgivelser kan aktivt påvirke mental funksjonering utenfor bevisst viten, men bare dersom hendelsen aktiverer eksisterende kunnskap eller hvis personen tilegner aktiv oppmerksomhet mot den delen av stimulus-feltet hvor hendelsen finner sted. Disse antagelsene gjenstår å teste (Kihlstrom 1990, s. 451).

Det er altså grunn til å anta, hevder Kihlstrom (1990, s. 450), at det eksisterer implisitt eller ikke-bevisst hukommelse og persepsjon. Men det eksisterer også evidens for implisitt *tenking og læring*. Et eksperiment som utforsker implisitt læring gikk ut på å undersøke forsøkspersoners evne til å bruke regler som er tilegnet gjennom erfaring, uten kjennskap til reglene i seg selv. Implisitt læring er i og for seg allerede vist i og med at mennesker gjennom tilegning av språk lære seg å skjelne ugrammatiske setninger fra grammatiske (Kihlstrom 1990, s. 452). Reber, 1976, 1989, referert i Kihlstrom (1990, s. 452), har forsøkt å modellere denne prosessen i laboratoriet ved å utvikle kunstige grammatikker. Reglene i disse grammatikkene bestemmer konstruksjonen av velformede bokstavsekvenser.

I Reber's prosedyre skal forsøkspersonene lære seg utenat f.eks. 20 bokstavstrenger som er satt sammen etter et sett med ca. 20 regler. Eksempler på grammatisk korrekte strenger er PVPXVPS eller PTTTVPS. De blir så testet med et antall nye strenger, noen av disse er grammatisk korrekte, f.eks. PTTTTVPS, mens andre ikke er det, f.eks. PTVPXVSP. Selv om forsøkspersonene ikke er i stand til å gjøre fullstendig rede for de grammatiske reglene som de har induert fra de strengene de allerede har sett, så klarer de å skille ugrammatiske fra grammatiske strenger på bedre enn sjansenivå. Kihlstrom (1990, s. 453) hevder, med støtte i flere tilsvarende undersøkelser, at det synes å være rimelig å anta at disse eksperimentene viser tilegning

av ny kunnskap i fravær av både bevisst intensjon om å lære, og bevisst viten om hva som er lært.

Kihlstrom (1990, s. 456) postulerer to generelle typer av konklusjoner. For det første ser ikke det *bevisste* ut til å kunne identifiseres ved noen bestemte perseptuelle eller kognitive funksjoner — alle disse funksjonene kan foregå utenfor fenomenologisk viten. Bevissthet kan i stedet ses på som en opplevelsesmessig kvalitet som kan følge alle disse funksjonene. Bevisst viten kan ha bestemte konsekvenser for psykologisk funksjonering: de er nødvendige for viljestyrt kontroll, for å kommunisere sine egne mentale tilstander til andre, og for å instruere andre. Men bevissthet er ikke nødvendig for mange former for kompleks psykologisk funksjonering. For det andre antar Kihlstrom (1990, s. 457) en midlertidig klassifisering av ikke-bevisste mentale strukturer og prosesser som er grunnleggende for det kognitivt ubevisstes kontrollområde.

Innen prosedyral kunnskap, finnes det et antall komplekse prosesser som er ubevisste i egentlig forstand (Kihlstrom 1990, s. 457). Det vil si at de ikke er tilgjengelig for bevisstheten under noen omstendigheter. De opererer på deklarativ kunnskap uten verken bevisst vilje eller bevisst viten, enten fordi de er rutinerne eller fordi de er medfødt. Disse utgjør det *egentlig ubevisste*⁵

I prinsippet er deklarativ kunnskap tilgjengelig for fenomenologisk viten, og kan gjøres kjent enten gjennom introspeksjon eller retrospeksjon. Men det er nå klart at prosedyrell kunnskap kan inter reagere med, og nyttiggjøre seg deklarativ kunnskap som i seg selv ikke er tilgjengelig for bevisst viten. Implisitt kunnskap og hukommelse tyder på at det finnes en kategori av «før-bevisste» deklarative kunnskaps-strukturer. I motsetning til virkelig ubevisst prosedyral kunnskap er denne kunnskapen tilgjengelig for det bevisste under vanlige omstendigheter, men den krysser ikke selv terskelen for å bli representert i arbeidshukommelsen og derved bli gjort tilgjengelig for bevisst viten. Disse representasjonene, som underligger implisitt persepsjon og hukommelse, befinner seg ved bevissthetens yttergrenser. Dersom omstendighetene forandrer seg kan de i prinsippet bli tilgjengelige for det bevisste (Kihlstrom 1990, s. 457).

Hypnose og tilsvarende fenomener ser ut til å være eksempler på en type «underbevisst»⁶ deklarativ kunnskap. Denne deklarative kunnskapen kan nyttiggjøres av bevisst viten, og den omfatter kognitive prosesser med så høy aktivitetsgrad, så som aktiv utnyttelse av oppmerksomhet og kompleks mental prosessering, at den burde være tilstrekkelig for å kunne represen-

⁵ «Unconscious proper»

⁶ «subconscious»

teres i arbeidshukommelsen. Likevel ser det ut til at denne kunnskapen *i seg selv* ikke er tilgjengelig for bevisst viten (Kihlstrom 1990, s. 457).

Emosjoner og det ubevisste

Enkelte teoretikere ser på mentale fenomener som identisk med bevissthet — alt annet er nevrofysiologi (Johnson-Laird 1988, s. 354). Freud postulerte i tillegg et ubevisst gjemmeded for primitive instinktive drifter og annet som er for angstfremkallende til å få innpass i det bevisste (*ibid.*) Helmholtz syn er at det ubevisste i tillegg rommer godartede prosesser som kan underligge tenkning og persepsjon. Deres eksistens tyder på at mentale prosesser forekommer parallelt (*ibid.*).

Parallell prosessering kan foregå på mange måter, hevder Johnson-Laird (1988, s. 355). Den kan ikke beregne noe annet enn hva en enkelt sekvensiell maskin kan beregne, men den kan gjøre det mye raskere og med lavere risiko for at skade på systemet vil være ødeleggende. Flere forskere innen «cognitive science» hevder, skriver Johnson-Laird, at hjernen ikke kunne ha utviklet seg uten å deles opp i separate moduler. Men modulene kan komme i konflikt med hverandre dersom moduler som har samme virkeområde kommer med motstridende meldinger. Da er det naturlig å forestille seg at noen prosesser overvåker flere prosesser i et hierarki med prosesser som får et stadig større ansvarsområde. De øverste prosessene kan da overstyre motstridende meldinger slik at systemet ikke går i «vranglås» (Johnson-Laird 1988, s. 356).

Enkel bevissthet, hevder Johnson-Laird (1988, s. 356), som f.eks. ren opplevelse av hendelser som smerte, kan ha sin opprinnelse i utviklingen av høyere-nivåers monitorer fra nettverket av parallelle prosesser. Dette «operativsystemet» på toppen av hierarkiet setter mål for prosesser på lavere nivå og overvåker deres innsats. Siden det er på toppen, kan dets instruksjoner spesifiseres rent symbolsk, som f.eks. «reis deg og gå!» Disse instruksjonene vil tolkes og sendes videre som stadig mer finkornete beskjeder helt ned til sammentrekningen av muskelceller (Johnson-Laird 1988, s. 356)

Opplevelsen av virkeligheten er en triumf i sinnets arkitektur, hevder Johnson-Laird (1988, s. 356). Operativsystemet har ingen tilgang til de prosessene som det er basert på. Det finnes gode evolusjonsmessige argumenter for dette, hevder, han: hvis det var mulig å reflektere over hele persepsjonsprosessen, ville den måtte bli mye langsommere siden det ikke ville være mulig å stole på parallellprosessering. Det ville også være mulig å trekke sansningen i tvil, noe som kunne få dramatiske følger i en faresituasjon (*ibid.*).

Denne teorien impliserer at det eksisterer et skille mellom bevisstheten og ubevisste prosesser lavere i hierarkiet. Slike *dissosiasjoner* kan finnes, hevder Johnson-Laird (1988, s. 357) i mennesker.

En slående dissosiasjon, hevder Johnson-Laird (1988, s. 358), er fenomenet «blindsyn⁷». Etter skade på cortex striata i occipitallappen (Weiskrantz, 1980, Weiskrantz *et al.*, 1974, referert i Kihlstrom (1990, s. 453)) hevder enkelte pasienter at de er blitt blinde på et relativt stort område i sitt synsfelt. Men dersom pasientene blir tvunget til å gjette på hvor et flekk med lys som skinner innenfor det blinde området befinner seg, så klarer de det bemerkelsesverdig godt (Johnson-Laird 1988, s. 358).

En annen form for dissosiasjon, hevder Johnson-Laird (1988, s. 358), er *selvbedrag*. De mest bemerkelsesverdige eksemplene på dette er hysteriske paralyser, som ikke skyldes skade på sentralnervesystemet, men ubevisste motiver. Lord Adrian (Adrian & Yealland, 1917, referert i Johnson-Laird (1988, s. 358)) kurerter nevrotiske paralyser som følge av «shell shock» under første verdenskrig. Han antok at pasientene ikke simulerte, men var overbevist om at de var paralyserte. Det var ikke mulig å overbevise dem om noe annet. Han var henvist til å lure pasientenes *ubevisste* prosesser som var ansvarlige for tilstanden, og behandlet paralyseren som om den var ekte. Ved en enkel metode som ikke kunne ha noen helbredende virkning på en ekte paralyse kurerter han 90% av tilfellene. Johnson-Laird (1988, s. 358) hevder at kognitive terapiformer hviler på samme idé, nemlig å få i stand en forandring i ubevisste prosesser.

Dissosiasjon i kognitiv forstand gir et nokså forskjellig syn på ikke-bevisst mental funksjonering enn psykoanalytisk teori (Hilgard, 1977; Kihlstrom, 1984, referert i Kihlstrom (1990, s. 455)). Ikke-bevisst mentalt innhold er ikke begrenset til primitive seksuelle eller aggressive idéer og impulser, og de trenger ikke være irrasjonelle, alogiske, eller på noen måte vesentlig forskjellige fra bevisste kognitive prosesser. Det eneste reelle skille er at de ikke er tilgjengelig for bevisstheten. (Kihlstrom 1990, s. 455). Adskillelsen fra det bevisste trenger heller ikke være motivert av forsvar, eller å ha angstreduksjon som effekt, tvertimot kan adskillelsen fra det bevisste simpelthen være et resultat av bestemte psykologiske operasjoner.

Thus, nonconscious mental processes are not restricted to automatized procedural knowledge, and nonconscious mental contents are not limited to unattended or degraded percepts and memories. These differences suggest that dissociative processes deserve more attention from both cognitive and clinical psycho-

⁷ «blindsight»

logists than they have received in the recent past. (*Kihlstrom* 1990, s. 455)

Refleksjon over selvet, hevder Johnson-Laird (1988, s. 361), kan foregå ved at man bevisst kan forme *mentale modeller* over seg selv. På samme måte som en robot kan unngå å falle ned i huller, eller å kjøre på ting, fordi den har en intern representasjon av sine omgivelser som guide, så kan mennesker forme interne representasjoner av sine omgivelser, inkludert seg selv (Johnson-Laird 1988, s. 360). Dette forutsetter at rammeverket av refleksjoner og intensjoner, som vi bruker til å reflektere rundt vår adferd, og forklare den, ikke er noe epifenomen, men spiller en kausal rolle i våre tanker og handlinger. Denne teorien gir mening til en rekke fenomener, hevder Johnson-Laird (1988, s. 362).

Hvis man har lykket i å løse en rekke oppgaver av samme slag, og så plutselig ikke får til et, så kan man spørre seg: «Hva er det jeg gjør når jeg løser disse oppgavene?» Svaret er basert på evnen til å betrakte seg selv, å reise seg selv et nivå opp slik at man kan iaktta seg selv som tilskuer. Svaret kan hjelpe til å reformulere hvordan man skal fortsette på det nivået hvor den aktuelle handlingen finner sted (Johnson-Laird 1988, s. 362).

Det er ikke bare mulig å betrakte seg selv, men man kan gå et nivå høyere og betrakte seg selv som betrakter seg selv, betrakte seg selv som betrakter seg selv som betrakter seg selv, osv. Fri vilje, hevder Johnson-Laird (1988, s. 366), ligger i at man kan velge på hvilket nivå man vil betrakte seg selv, og handle i samsvar med det nivået.

Bevissthet, hevder Johnson-Laird (1988, s. 367), er en egenskap ved en klasse beregningsmessige prosedyrer. Resultatet av disse prosedyrene kan i prinsippet hentes fra serielle prosedyrer som ikke kan fremkalle bevissthet. Det er altså ikke noen måte å observere bevisst adferd for å være sikker på at den er bevisst. Ikke desto mindre er det i dagliglivet mulig å avgjøre om en person er bevisst: Ettersom bevissthet er det høyeste nivå av refleksjon, så går det an å finne ut i hvor stor grad en person er i stand til å reflektere over seg selv. Mennesker som kan rapportere at de har valgt en bestemt måte å handle på fordi de veide for og mot en rekke alternativer og kom fram til denne måten fordi de syntes den var den beste, og som kan si at de valgte denne måten å handle på som et resultat av refleksjon over et utvalg av metoder, kan regnes som ansvarlige, bevisst handlende mennesker med fri vilje. Pasienter med hjerneskade eller mentale sykdommer eller andre traumer kan tenkes å ikke ha tilgang til modeller av seg selv eller til en reflekterende prosedyre, og vil ikke være i stand til å ta del i en selv-reflekterende diskusjon. Stress, engstelse og andre faktorer øker vanlige menneskers tendens til å handle feil.

Selv om Johnson-Lairds diskusjon er åpen for moralsk kritikk, vil jeg gjøre bruk av hovedpoenget hans, nemlig at det bevisste er det øverste nivået i en dissosiativ struktur, i den videre diskusjonen.

Bevissthetens innhold faller inn i to kategorier: symbolsk og ikke-symbolsk. Det symbolske innholdet er persepsjoner, idéer, antagelser, hypoteser, osv. De kan beskrives i ord, som når man beskriver en oppfatning for en annen. (Johnson-Laird 1988, s. 368). Det ikke-symbolske innholdet i bevissheten er følelser og sansefornemmelser. Det er mulig å føle en følelse som frykt eller en fysisk fornemmelse som smerte. De er ikke symboler for noe, og har ikke noen intern proposisjonell struktur. Men det er et intimt forhold mellom bevissthetens symbolske og ikke-symbolske struktur.

Johnson-Laird (1988, s. 369) setter opp følgende tre «lover» for autonome organismer:

1. An organism must reproduce and ensure the survival of its genes
2. An organism must protect its own existence as long as such protection does not interfere with the First Law
3. An organism may cooperate with members of its own species and help them from coming to harm so long as such cooperation does not conflict with the First and Second Law

Evolusjonen har gitt levende vesener svært spesifikke mål, hevder Johnson-Laird (1988, s. 370), og spør: «hvordan blir de oversatt til atferd?»

Primitive dyr, som f.eks. insekter, har et medfødt sett av responser for å hankses med sine omgivelser: at en bestemt hendelse finner sted setter i gang en bestemt respons. Komplekse organismer, som f.eks. mennesker, rår over kraftige slutningsmekanismer og eksplisitte symbolske prinsipper for å bestemme hva som skal gjøres. Medfødte responser er ufleksible, slutninger koster tid og krefter. Finnes, det, spør Johnson-Laird (1988, s. 370), en mellomting?

Ja, svarer han, mellomtingen er *følelser*, som gjør organismen klar for ikke en spesifikk medfødt reaksjon, men en generell handlingsplan som passer til situasjonen.

Enkelte mål som organismer må hankses med har å gjøre med fysiske behov, men andre har å gjøre med eksterne handlinger og i særdeleshet forhold til andre medlemmer av den samme arten. Kroppslige behov har fysiske årsaker, f.eks. er det slik at et dyr som ikke får mat blir sultent. Kroppslige behov kan avsluttes av andre fysiske årsaker, som produserer andre kroppslige tilstander. Men relasjoner til andre medlemmer av arten, hevder Johnson-Laird (1988, s. 371), har psykologiske årsaker. De blir

skapt og avvirket av kognitive vurderinger. Oppfattelsen av et rovdyr forårsaker frykt, og denne frykten kan kommuniseres til andre medlemmer av arten ved ritualisert adferd som ikke har annen funksjon enn å kommunisere emosjonen til andre medlemmer av samme art. Johnson-Laird (1988, s. 371) hevder at emosjoner har sin opprinnelse som en kontrollmekanisme for interaksjon med andre medlemmer av arten.

De viktige situasjonene i en arts liv kan fordeles in i et relativt lite antall emosjoner, hevder Johnson-Laird (1988, s. 371). Det finnes bare en liten klasse av signifikante hendelser i et dyrs sosiale liv:

1. Å lage eller oppheve et bånd mellom omsorgspersoner og avkom
2. Aksept eller avvisning av bånd mellom et *par*
3. Aggresjon mot eller flukt fra rivaler, og tilsvarende reaksjoner mot byttedyr eller rovdyr

Denne ontologien, hevder Johnson-Laird (1988, s. 371), finnes også i mennesker og den kan f.eks. ytre seg i det følgende sett med evolusjonsmessig grunnleggende emosjoner:

1. Glede, som finner sted i vellykkede bånd, men kan ha forskjellige *modi* avhengig av om båndet er seksuelt eller mellom omsorgsperson og avkom
2. Sorg, som et resultat av separasjon fra et individ som man har et bånd til
3. Sinne, som en forløper til aggresjon, som kan ha forskjellige *modi* avhengig av om den er rettet mot rivaler, inntrengere, rovdyr eller byttedyr
4. Frykt, som en forløper til overgivelse til rivaler, flukt fra rovdyr, osv.
5. Avsky, som en forløper til avvisning

Dette tyder på at f.eks. dyr skulle være i stand til å føle tørst. Eksperimenter tyder på at de kan gjøre det. En tørst rotte vil slikke mot en stråle kald luft, noe som vil føre til at enda mer vann fordampes fra tungen (Johnson-Laird 1988, s. 373). Dette, hevder Johnson-Laird, er en villedning av kroppslige signaler. Slike følelser kunne tjene som et signal til hjernens operativsystem slik at adferden blir regulert til å oppnå et relevant biologisk mål. Elektrisk stimulering av visse punkter i en rottes hjerne, nær

hypotalamus, vil føre til at rotta vil presse en spak kontinuerlig hvis det fører til at den elektriske stimuleringen fortsetter (Olds & Milner, 1954, referert i Johnson-Laird (1988, s. 373)). I den senere tid er det vist at disse punktene i hjernen produserer neurotransmitteren *dopamin*.

Johnson-Laird (1988, s. 374) hevder at de forskjellige prosessorene i hierarkiet kan bruke neurotransmittere for å kommunisere med hverandre. Det behøver ikke være en bestemt neurotransmitter for hvert signal, men bestemte neurotransmittere kan brukes på bestemte punkter i systemet. I motsetning til setningene i et språk, har ikke signalene en eksplisitt symbolsk struktur. De er mer som alarmsignaler, hver med et komplekst mønster som ikke lett lar seg forveksle med andre.

Men kan det være slik at det ubevisste *bare* kan kommunisere med følelser? Dersom det kan finne sted såvidt avanserte prosesser i det ubevisste som tenkning og persepsjon, og det bevisste under visse omstendigheter kan gripe fatt i det ubevisstes deklarativer kunnskap, kan det tenkes at det ubevisste også kan bidra med informasjonsprosessering av svært avansert art, som Nørretranders (1991) hevder at den skotske fysikeren James Clerk Maxwell ga uttrykk for på dødsleiet, i en refleksjon over hvordan han var kommet frem til sine resultater:

Hvad der gøres af det, der kaldes mig selv, bliver, føler jeg, gjort af noget, der er større en mig selv i mig

Fortrengning og det kognitivt ubevisste

Flere forsøk er blitt gjort på å forene kognitiv psykologi og psykoanalyse, f.eks. Erdelyi, 1985, referert i Johnson-Laird (1988, s. 368) og i Lazarus (1991, s. 362). Gullestad (1992, s. 818) ser nærmest på kognitiv psykologi som en *konkurrent* i å forklare fortrenningsfenomenet: men kognitiv psykologi, hevder Gullestad (1992, s. 817), kommer til kort bl.a. fordi det ubevisste som kognitiv psykologi er en *snillere, mildere og mer rasjonelt* enn det freudske.

Johnson-Lairds beskrivelse av det ubevisste som generator for emosjoner tegner et bilde som slett ikke behøver å bli oppfattet som snilt: han inkluderer både sinne og aggresjon i sin forståelse av det ubevisste. Gullestad (1992, s. 818) hevder videre at forskningen har begrenset seg til implisitt kognisjon, og ikke beskjeftiget seg med implisitt emosjon og motivasjon, og konkluderer med at

På denne bakgrunn kan det hevdes at den eksperimentelt orienterte forskningen ennå ikke er kommet i posisjon til å nærme

seg den type psykologiske fenomener som det psykoanalytiske fortrenningsbegrepet referer til (*Gullestad 1992, s. 817*)

Selv om den kognitive psykologien ikke har nærmet seg fortrenningsbegrepet per se, mener jeg likevel det er grunn til å hevde at emosjoner i den senere tid *er* blitt tatt alvorlig i kognitiv psykologi. Kognitiv psykologi kan muligens også komme med en alternativ vinkling på fortrenningsbegrepet: selv om kognitiv psykologi ikke beskjeftiger seg med fortrenning, oppfatter jeg dens postuleringer av det ubevisste som delt inn i avdelinger som en mulig forklaring på *hvordan* angstfremkallende kognisjoner blir fortrent. Jeg anser det ikke som urimelig å anta at det *kognitivt ubevisste* kan ha lært hva som er ubehagelig for det bevisste, og kan opptre som en sensurinstans i forhold til det bevisste. På den måten kan både fortrenningsmekanismene og det fortrente forbli utilgjengelig for bevisstheten — hele tiden.

Gullestad (1992, s. 817) kritiserer dissosiasjonsteorien for at adskillelsen av avdelingene ikke er *motivert* ut fra et forsvarshensyn. Det kan slås fast, skriver hun, at med en slik teori gir det ingen mening å snakke om at personen forsvarer *seg*, for det er ikke noe *seg* å forsvare. Men Johnson-Laird hevder jo nettopp at det *er* et *selv* å forsvare, bl.a. fordi det er en eksistens å opprettholde.

Konklusjon

I det foregående har jeg forsøkt å belyse visse problemstillinger i forhold til det ubevisste i lys av nyere teori og forskning. Dette er et stort og vanskelig område: det er vanskelig, føler jeg, så si noe bevisst om noe som er ubevisst.

PDP-modeller har fått forholdsvis stor plass, og jeg har ikke brukt litteratur som er grunnleggende *uenig* eller stiller seg kritisk til idéen om at det er mulig å fremstille maskiner som kan modellere bevissthet eller det ubevisste. Selv en steil kritiker som Searle (1987, s. 40) hevder at hjernen er en digital datamaskin. I artikkelsamlingen *Artificial intelligence: The Case Against* er det en artikkel av Hubert L. Dreyfus som skrev boken «*What Computers Can't Do*». Her argumenterer han for at AI er basert på feilaktige grunnleggende antagelser om menneskelig kunnskap og forståelse (Dreyfus 1987, s. 41). Men han går ikke mot den grunnleggende ideen at datamaskiner en gang vil bli intelligente, og hevder at PDP-modeller kan være den rette veien å gå (Dreyfus 1987, s. 54).

Mye tyder på at det er rimelig å anta at det ubevisste er ansvarlig for en ikke uvesentlig del av høyere ordens informasjonsprosessering, både i området fortrenning, men også hva angår kognitiv prosessering på et relativt

høyt nivå. Hvis ikke mye av den tenkningen og informasjonsbehandlingen vi foretar var ubevisst, ville vi ikke hatt tid til annet. Dette kan dreie seg om alt fra å kjøre bil (en automatisert prosess når det først er lært) til kanskje langt mer komplekse ting, kanskje så komplekse som å foreta moralske valg, eller å planlegge framtida i grove trekk.

Nyere teori og forskning tyder på at det språket som det ubevisste gjør bruk av i kommunikasjon med det bevisste er *emosjoner*. Emosjoner varsler oss om ting i omgivelsene, fra nytelsen av et godt måltid eller gleden ved å slappe av og nyte en vakker solnedgang, til smerter når vi brenner oss eller trækker på noe skarpt. Heller ikke da er vi bevisst hva som faktisk forårsaker de følelsene som oppstår.

Det er konvensjonell visdom at man skal stole på den førtste innskytelsen, «den er som oftest den riktige.» Dette kan kanskje forklares ved at den første innskytelsen er levert av det ubevisste, med all den informasjonsprosesseringskapasitet som det ubevisste rår over. Dersom man avveier konklusjonen, så har man kanskje ikke tid eller «ork» til å vurdere alle trinn fram mot en løsning, og så blir «magefølelsen» tuftet på et mer solid grunnlag enn den rasjonelle, bevisste konklusjonen!

Det er ikke dermed sagt at man først og fremst skal stole på følelsene. Jeg tror at det er rimelig å anta at man tross alt har mer kontroll over det bevisstes utledninger fram mot en konklusjon. Det er også disse utledningene man i første rekke kan dele med andre. Men dersom den konklusjonen man kommer fram til strider mot det man føler, så kan det tenkes at det er klokt å prøve å finne ut *hvorfor* man har motstridende følelser. Mye tyder på at det ubevisste kan betraktes som en «vis mann» eller en «klok kone.» Å aktivt overse sine følelser kan være det samme som å kaste vrak på gode råd.

Tillegg: Church-Turing tesen

Hovedpoenget som *Church-Turing tesen* hviler på er at alle spesifikasjoner for beregning av matematiske oppgaver kan reduseres til instruksjoner for en Turing-maskin. Church-Turing-tesen er at

Turing-maskinen definerer, matematisk, hva vi mener med en algoritmisk (eller effektiv eller rekursiv eller mekanisk) prosedyre (Penrose 1989, s. 50).

Hvorfor er det i det hele tatt spennende å finne en maskin som kan beregne en hvilken som helst beregnbar oppgave? Jo, for det første fordi det ikke bare er snakk om *algebra*, dvs. regneoppgaver av ren matematisk natur av typen « $2+2=?$,» eller for den saks skyld mer avanserte oppgaver hvor svaret er et tall. Det kan også dreie seg om såpass avanserte ting som algebraiske eller trigonometriske uttrykk, dvs. beregning av funksjoner hvor svaret også er en funksjon (Penrose 1989, s. 50). For det andre har dette en rekke erkjennelsesteoretiske implikasjoner, som postulert i *Church-Turing tesen*.

Beregnbarhet

Penrose (1989, s. 49) skriver at en Turing-maskin kan beregne et hvilket som helst beregnbart problem dersom maskinen blir gitt gode instruksjoner for hvordan beregningene skal gjøres. Alle beregningene behøver ikke være spesifisert ved starten, maskinen kan skape og løse sine egne beregninger underveis mot målet. Turing-maskiner kan bli beskrevet på forskjellige måter, slik at en beskrivelse kan være mer effektiv enn en annen for et gitt problem, eller man kan kjøre flere maskiner samtidig for at beregningene skal kunne utføres raskere, dette er et av prinsippene bak parallellprosessering.

Kompleksitet

Nørretranders (1991) hevder at Turing samtidig som Church-Turing-tesen ble formulert, ble klar over at det finnes beregningsoppgaver som en Turing-maskin ikke kan regne seg gjennom på en overskuelig måte, dvs. regneoppgaver som er så komplekse at man ikke kan vite om de er ferdige før maskinen kommer med svaret. Nørretranders går så langt som til å hevde at man

ikke kan vite på forhånd når en datamaskin blir ferdig med en oppgave før den har forsøkt⁸. (Nørretranders 1991, s. 79)

Videre: alt som kan beregnes av en digital datamaskin kan bli beregnet av en Turing-maskin. Det følger ikke dermed nødvendigvis at bevissthet er et beregnbart fenomen, det er mulig at bevissthet også beror på fenomener som ingen datamaskin kan mestre (Johnson-Laird (1988, s. 51); Searle (1987, s. 40)) Men hvis bevissthet er et beregnbart fenomen, da går det i teorien an å lage datamaskiner som er bevisste.

Det er viktig å være klar over at PDP-nettverk ikke bare er store parallellprosessorer. Det å behandle data parallelt er ikke nok. Det som teller er programmet av gruppeavgjørelser i samarbeid (Clark 1987, s. 5). En sekvensiell Turing-maskin kan gjøre de samme beregningene som et PDP-nettverk kan, men PDP-nettverket er mye raskere og, dets virkemåte er *kanskje* mer likt måten mennesker tenker Clark (1987, s. 14).

PDP-modellene ble møtt med stor entusiasme da de kom, men de har ennå ikke produsert de resultatene som var forventet. Det har vært snakk om «mye skrik og lite ull.» Jeg vil foreslå tre grunner for dette:

1. PDP-modeller er i seg selv lite forklarende. Den diffuse måten de opererer på gjør at ting som er simulert på PDP-maskiner i seg selv kan være vanskelig å forklare. Denne ulempen kan imidlertid delvis oppveies av at PDP-modellen er åpen for inspeksjon.
2. Det er ikke lett å forutsi hvilket resultat en PDP-modell vil komme med. Man må ha datamaskiner til disposisjon, eller annet utstyr som kan settes opp som en PDP-modell. Det krever med andre ord et visst materielt utstyr som ikke nødvendigvis er lett tilgjengelig.
3. Å beskrive ting som PDP-modeller krever mennesker med høyt spesialisert kunnskap

⁸Dette er ikke helt riktig. Det finnes lange tradisjoner innen matematikk og databehandling for beregning av hvor lang tid en beregning vil ta, hvor man ikke tar utgangspunkt i hva som skal beregnes, men i hvilken type beregning det er snakk om. I enkelte tilfelle kan man vite nøyaktig hvor mange operasjoner som trengs og derfor også hvor lang tid det vil ta (de såkalte $O(N^a)$ -type problemene). Problemet oppstår først med de oppgavene som ikke er av en slik art at man kan vise hvilken klasse de tilhører. Her gjelder det at man ikke kan bevise om programmet en gang vil stoppe, eller om det vil fortsette i det uendelige. Det vil forstette i det uendelige dersom det ikke finnes noe svar, men maskinen kan ikke selv vite om det finnes noe svar før den eventuelt har funnet svaret. Og ettersom det ikke er mulig å vite hvor mange beregninger som trengs for å finne svaret, så finnes det ikke noe tidspunkt hvor man vet at nå har maskinen holdt på så lenge at det er sikkert at den aldri vil finne noe svar. Men dette gjelder altså langt fra alle beregninger

Alt i alt kan den uteblivende utbredelsen av PDP-modeller ha deler av sin forklaring i at det er tvil om utbyttet kan være verdt innsatsen.

Referanser

- Atkinson, Rita L., Atkinson, Richard C., Smith, Edward H. & Hilgard, Ernest R., (1987) *Introduction to Psychology*. San Diego, California: Harcourt Brace Jovanivich, 9. utgave
- Clark, Andy, (1987) Connectionism and Cognitive Science. I: John Hallam & Chris Mellish (red.), *Proceedings of the AiSB 1987: Advances in AI*, Chichester. John Wiley & Sons
- Dreyfus, Hubert L., (1987) Misrepresenting Human Intelligence. I: Rainer Born (red.), *Artificial Intelligence: The Case Against*. London: Routedledge
- Gullestad, Siri Erika, (1992) Om begrepet fortrenning. *Tidsskrift for Norsk Psykologforening* 29 (9)
- Hogan, Robert, (1985) What every student should know about personality psychology. I: Vivian Parker Makovsky (red.), *The G. Stanley Hall Lecture Series*, bind 6. Washington, D. C.: The American Psychological Association
- Johnson-Laird, Philip N., (1988) *The Computer and the Mind*. Cambridge, Massachusetts: Harvard University Press
- Kihlstrom, John F., (1990) The Psychological Unconscious. I: L. A. Pervin (red.), *Handbook of Personality: Theory and Research*. New York, New York: Guilford
- Lazarus, R. S., (1991) Cognition and motivation in emotion. *American Psychologist* 46: 352–367
- Nørretranders, Tor, (1991) *Mærk Verden*. Viborg: Gyldendal
- Penrose, Roger, (1989) *The Emperor's New Mind*. New York, New York: The Oxford University Press
- Pervin, Lawrence A., (1989) *Personality: Theory and Research*. New York, New York: John Wiley & Sons, 5. utgave
- Searle, John R., (1987) Minds, Brains, and Programs. I: Rainer Born (red.), *Artificial Intelligence: The Case Against*. London: Routedledge