



ROSKILDE UNIVERSITET
KANDIDATUDDANNELSEN
FILOSOFI OG VIDENSKABSTEORI &
PSYKOLOGI

SPECIALEAFHANDLING

THE COMPUTATIONAL THEORY OF MIND:
TEORIENS BEGRÆNSNINGER

TOBIAS ROSSEN
STUDIENR. 57469
6/4-2021
ANTAL ANSLAG: 164991

VEJLEDER: PATRICK BLACKBURN
MEDVEJLEDER: ERNST SCHRAUBE

Abstract

This thesis explores five objections that Susan Schneider raises against Jerry Fodor's claim that the computational theory of mind very likely isn't able to explain more than a very limited part of our cognitive processes. In *Modularity of Mind* (1983), *The Mind Doesn't Work That Way* (2001) and *LOT 2: The Language of Thought Revisited* (2008) Fodor provides two different but closely related arguments for the limitations of the computational theory of mind. The overall claim is that it seems impossible for the computational theory of mind to explain how our daily cognitive processes fixate our beliefs. First he claims that for the computational theory of mind to be true our cognitive system must somehow only choose a few relevant background beliefs to update given a new belief. It must also only choose a few relevant background beliefs from which the estimation of confirmation of our daily hypotheses are computed. The problem for the computational theory of mind is that we have no idea what makes background beliefs relevant. This is the problem of relevance. Second he claims that it's very likely that the computational theory of mind can't explain the fact that we somehow estimate certain properties of beliefs and of theories which the beliefs and theories have in virtue of their relation to the whole system of our background beliefs. This is the problem of globality. In *The Language of Thought: A New Philosophical Direction* (2011) Schneider raises two objections against the relevance problem and three objections against the globality problem. I claim that all Schneider's objections fail but that she raises a serious question to one part of Fodor's argumentation for the globality problem. She also does provide a refutation of that same part of Fodor's argumentation but she fails to refute a reformulation of the argumentation.

Indhold

Indledning	3
Mit undersøgelsesspørgsmål	6
Afhandlingens opbygning	9
Det grundlægende i CTM.....	11
Common-sense psykologi som videnskab	11
The Language of Thought	13
CTM.....	16
Fodors relevans- og globalitetsargument	24
Relevansargumentet	24
Globalitetsargumentet	30
Schneiders indvendinger mod Fodors relevans- og globalitetsargument	39
Schneiders indvendinger mod relevansargumentet	39
1. indvending - Et skakprogram.....	39
Mit svar	40
2. indvending - Global Workspace Theory	41
Mit svar	43
Schneiders indvendinger mod globalitetsargumentet	46
1. indvending - Et skakprogram.....	46
Mit svar	48
2. indvending - CTMs ressourcer	49
Mit svar	51
3. indvending - En problematisk definition og usandsynligheden for globale egenskabers eksistens	52
Mit svar	57
Konkluderende bemærkninger	62
Litteraturliste	64

Indledning

Psykologien er en videnskab der har haft og stadig har svært ved at afgrænse sin genstand: Psyken. Dette har samtidigt gjort det svært for psykologien at forene sig om en metode til at undersøge denne genstand videnskabeligt. Et hurtigt blik ned gennem, bare en lille del, af psykologiens (meget overordnede) historie, viser mange forskellige opfattelser af hvordan psyken skal betragtes og hvordan den kan undersøges. Wilhelm Wundt (1832-1920), der grundlagde det første institut for eksperimentel psykologi i Leipzig, argumenterede i 18- og 1900-tallet for at psykologiens genstand er *den bevidste oplevelse*, og at en videnskabelig undersøgelse af psyken består i introspektive undersøgelser af den indre sammenhæng mellem elementerne i vores oplevelser og deres sammenhæng med ydre stimuli (Christensen, 2013: 45-46). Sigmund Freud (1865-1939), en mand der formentlig mere end nogen anden satte sit præg på 1900-tallets psykologi, argumenterede i slutningen af 1800-tallet og starten af 1900-tallet for, at psykologiens genstand ikke kun er bevidstheden, men også - og især - underbevidstheden. Han mente man fik adgang til underbevidstheden gennem menneskers fortællinger om deres drømme og frie associationer under kontrollerede forhold (Christensen, 2013: 47-49). Stort set samtidig med at Freud argumenterede for underbevidstheden som en genstand for psykologien, blev behaviorismen en dominerende tilgang til psykologien i USA. Behaviorismen hævdede at man umuligt kunne studere bevidstheden videnskabeligt, hvis bevidstheden overhovedet var et meningsfuldt begreb (Christensen, 2013: 54). Ifølge behaviorismen kunne psykologiens genstandsfelt kun være adfærd, og videnskaben måtte således bestå i undersøgelser af sammenhænge mellem stimuli og den adfærdsmæssige respons (Christensen, 2013: 55-56). Behaviorismen adskilte sig således radikalt fra både Wundt og Freud. I 1950'erne og 1960'erne blev der, blandt andet som modreaktion til behaviorismen, argumenteret for flere andre tilgange til psykologien, der efterfølgende vandt udbredelse (Christensen, 2013: 57, Eysenck & Keane, 2013: 2). Den eksistentiale og humanistiske psykologi argumenterede for at bevidstheden er det egentlige menneskelige, og anklagede dermed behaviorismen for at forbigå en essentiel del af mennesket og psykologien. Den eksistentiale og humanistiske psykologi mente således, ligesom Wundt, at *den bevidste oplevelse* er psykologiens genstand. Fokus var imidlertid på menneskets eksistentielle vilkår, især dets frihed og de problemer, der følger med denne frihed. Derudover var fokus på menneskets oplevelse som helhed, og ikke de atomistiske dele, som Wundt hævdede var psykologiens genstand (Christensen, 2013: 57-58). Samtidig med at den eksistentiale og humanistiske psykologi vandt udbredelse, etablerede kognitionspsykologien sig

også som en videnskab. Kognitionspsykologiens genstand er de (indre) processer, der er involveret i vores forsøg på at forstå verden omkring os og i vores beslutninger. Problemløsning, perception, læring, sprogforståelse og beslutninger er således typiske processer kognitionspsykologien undersøger (Eysenck & Keane, 2013: 1-2). Kognitionspsykologiske forskere kritiserede behaviorismen for ikke at kunne give en tilstrækkelig forståelse af vores kognitive processer. Menneskers sproglige adfærd kunne for eksempel umuligt forstås gennem stimuli-respons-forhold mellem omverden og menneskers adfærd. Der måtte også ligge indre processer til grund (Eysenck & Keane, 2013: 2, Fodor, 1978: 100-101).

Allerede ud fra denne meget overfladiske og ufærdige gennemgang af psykologiens historie, er det tydeligt at psykologiens genstand og hvordan den skal undersøges, har været et omstridt emne. I dag er psykologien ikke meget mere forenet. Der er stadig psykologer der følger i Freuds fodspor (Christensen, 2013: 51). Den eksistentiale og humanistiske psykologi har stadig sin plads i psykologien og kognitionspsykologien har om nogen også stadig sin plads i psykologien (Christensen, 2013: 54, 59). Psykologiens splittelse viser sig derudover ved, at det i dag er normalt at inddele psykologiens forskellige tilgange til dets genstand og undersøgelsesmetoder, i en naturvidenskabelig, en humanistisk og en samfundsvidenskabelig tilgang, som alle adskiller sig markant fra hinanden på flere områder (Christensen, 2013: 82-83). De adskiller sig for eksempel radikalt i svarene på spørgsmål som: 'Er psyken noget der kun kan forstås i sin helhed eller kan den deles op i elementer, der kan studeres isoleret?', 'Hænger psyken uadskilleligt sammen med omverden og de samfundsmæssige betingelser for menneskers handlinger?', 'Kan man studere psyken gennem statistiske undersøgelser eller kan man udelukkende opnå indsigt i psyken gennem kvalitative undersøgelser?', 'Er det muligt at finde generelle love om psyken?' 'Kan psykiske tilstande reduceres til hjernetilstande?' (Christensen, 2013: 87-88). Nogen vil mene at svarene på disse spørgsmål kommer an på hvilken del af psyken man er interesseret i, og at de forskellige tilgange således belyser forskellige dele af den samme genstand (Christensen, 2013: 7, 85). Det er jeg umiddelbart ikke uenig i, men det er ikke klart hvordan de forskellige dele hænger sammen, og der er langt fra enighed blandt forskere om hvorvidt der med de forskellige tilgange belyses forskellige dele eller om der er tale om modstridende opfattelser af psykologiens genstand (Christensen, 2013: 113, Dreier, 2007: 11, 14, 16-17, Sterelny, 1990: 199). Psykologien er ikke en samlet videnskab (Flyvbjerg og Sampson, 2011: 1-2)

Jeg er mere end almindeligt interesseret i alle disse spørgsmål om psykologiens genstand og hvordan den kan undersøges, og jeg er af den umiddelbare overbevisning, at psykologien *kan* forenes

og blive en mere samlet videnskab. Med grundige filosofiske undersøgelser af de grundlag der er for de forskellige tilgange til psykologien, deres forskelle og ligheder, og de (fejl)slutninger de hviler på, er jeg (foreløbigt) overbevist om, at psykologiens genstand kan afgrænses og psykologien kan blive forenet i en samlet tilgang. Eller, i hvert fald, i en (mere) samlet forståelse af hvordan de mange forskellige tilgange til psyken hænger sammen. Der er selvfølgelig lang vej igen, og psykologiens historie taler måske imod at det nogensinde skulle kunne lade sig gøre. Det mener jeg imidlertid ikke man skal lade sig slå ud af. At komme et lille skridt nærmere en afgrænsning af psyken og hvordan den lader sig undersøge, hvilket *må* være muligt, er et skridt nærmere en forenet psykologi.

Denne specialeafhandling udgør en forsøg på at komme et skridt nærmere en afgrænsning af psyken og hvordan den kan undersøges. I løbet af min studietid på Roskilde Universitet er jeg hovedsageligt blevet præsenteret for humanistiske og samfundsvidenskabelige tilgange til psykologien. Det vil sige, tilgange, der ofte mener, at vi ikke kan undersøge psyken¹ gennem dets delelementer og isoleret fra de sociale og samfundsmæssigt betingede omgivelser. Det hævdes også at man dermed ikke kan finde årsagsforbindelser mellem psykiske elementer og mellem omgivelser og psyken (Se for eksempel Holzkamp et al., 2013: 20-22, 70, 289-290). Jeg er imidlertid ikke blevet overbevist om, at disse humanistiske og samfundsmæssige tilgange har helt ret i dette. Jeg mener at vores common-sense forståelse af vores psykiske liv, med tanker og følelser som årsagsbærende elementer, er et godt udgangspunkt for en videnskab. Jeg mener derudover at en videnskab med fordel kan inddele vores tanker og følelser i forskellige elementer, præcisere disse elementer og undersøge deres virkninger på hinanden. Kognitionspsykologien tager netop udgangspunkt i en del af vores common-sense forståelse af vores psykiske liv, og den gør det (hovedsageligt) med en naturvidenskabelig tilgang, hvor psyken indeles i elementer og årsagssammenhænge undersøges (Christensen, 2013: 96). Der er tilmed udformet præciserende filosofiske grundlag for kognitionspsykologien. Jeg har derfor, og fordi jeg kun sporadisk er blevet introduceret for naturvidenskabelige tilgange til psyken, valgt at undersøge kognitionspsykologien og et af dens filosofiske grundlag. Det grundlag jeg vil undersøge, består overordnet i en påstand om at vores kognitive system fungerer som en form for computer. En undersøgelse af denne påstand, vil forhåbentlig resultere i at vi kommer bare et lille skridt nærmere en forståelse af hvordan vi kan afgrænse psyken og således forenes om måder at undersøge psyken på.

¹ Eller *subjektet*, som er det foretrukne begreb

Mit undersøgelsesspørgsmål

Kognitionspsykologien etablerede sig for alvor som en videnskab i 1950'erne. Mere eller mindre fra start af var ideen om at vores kognitive processer er en form for computer-processer en grundlæggende ide i kognitionspsykologien (Eysenck & Keane, 2013: 2, Stanford 1). Hvad der mere præcist ligger i, og kan ligge i, at vores kognitive processer er en form for computer-processer, og om dette er rigtigt, er blevet diskuteret lige siden. En af de formentlig vigtigste og mest fremtrædende fortalere for computer-forståelsen var filosofen Jerry Fodor (1935-2017) (Bailey, 2013: 125). Fodor argumenterede gennem flere bøger og artikler for bestemte grundlæggende forståelser af vores kognitive system og vores kognitive processer og hvordan disse kan forstås som computer-processer. Han udformede med andre ord et filosofisk grundlag for kognitionspsykologien. En af de overordnede påstande han fremførte var, at vi mentalt repræsenterer verden omkring os (og disse repræsentationer selv), og at vores kognitive processer består af bestemte former for kausale kæder mellem de mentale repræsentationer. Det er disse kausale kæder, der udgøres af computer-processer (Fodor, 1978: 31, 73-75, 165, Fodor, 1987: 17). Computer-processer er her, groft sagt, forstået som fysiske processer, hvor de fysiske tilstande der forårsager hinanden, kan tolkes som symboler hvis semantiske indhold udgør et bevis for en gyldig slutning. Vores mentale repræsentationer er altså "tolkningen" af de fysiske processer og kausalkæderne mellem repræsentationerne udgør slutningsprocesser (Fodor, 1987: 18-19, Fodor, 1978: 73-75, 198). Med denne forståelse består vores kognitive processer således i *symboler, der gennem deres fysiske struktur forårsager hinanden i overensstemmelse med (rationelle) slutningsprocesser*, og vores kognitive system er dermed et *symbolmanipulerende system* (Fodor & Pylyshyn, 1988: 12-13, Fodor, 1987: 18-19).

Ideen om at vores kognitive system er et slags symbolmanipulerende system, var, udover at udgøre et grundlag for kognitionpsykologien, en af de vigtigste ideer i bevidsthedsfilosofien i 1900-tallet. Det var den fordi den udgør en fysisk teori om vores mentale tilstande, der ikke reducerer mentale tilstande til typer af fysiske (hjerne-)tilstande eller dispositioner for adfærd (Bailey, 2013: 125). Den forener altså den mere eller mindre uimodsigelige kendsgerning, at der er noget, der er anderledes end det rent fysiske, men som hænger sammen med det fysiske, samtidig med den ikke påstår en 'magisk' sjælelig substans. Ifølge Fodor var computer-forståelsen også et stort fremskridt i forhold til den associationistiske forståelse af vores mentale liv, der havde præget 1800-tallets psykologi og som ifølge Fodor udgør det eneste alternativ til en filosofisk teori om vores mentale processer (Christensen, 2013: 39, Fodor, 1998: 10). Associationismen hævdede ligesom den

symbolmanipulerende forståelse, at vi har mentale repræsentationer, der forårsager hinanden. Associationismen betragtede imidlertid de kausale forbindelser som *associationistiske* forbindelser frem for *regelstyrede* sekvenser (som i den symbolmanipulerende forståelse). Det vil sige, det blev hævdet at forbindelserne mellem mentale repræsentationer er styret af hvor tit de forskellige repræsentationer (eller de forskellige dele af repræsentationerne) er optrådt sammen eller efter hinanden, eller hvor meget de ligner hinanden (Fodor, 1998: 10). Problemet med denne associationistiske forståelse er, ifølge Fodor, at den ikke kan redegøre for vores rationelle mentale liv. Rent associationistiske love kan ikke forklare hvordan en (formodet) sand tanke generelt fører til den næste sande tanke, som tanker gør i vores daglige slutningsprocesser (Fodor, 1998: 10, Fodor, 1987: 18). Dette kan de symbolmanipulerende love. De kan netop give en mekanistisk forklaring på vores rationelle tænkning (Fodor, 2001: 52-53, Fodor, 1987: 18).

I dag spiller den symbolmanipulerende forståelse stadig en afgørende rolle i kognitionspsykologien, selvom den er blevet kritiseret fra flere sider og selvom en anden form for computer-forståelse, konnektionisme, har vundet udbredelse (Eysenck & Keane, 2013: 20-22, Web 1). Fodor har blandt andet forsvaret den symbolmanipulerende forståelse overfor konnektionisme, ved at argumentere for, at konnektionisme er en form for associationistisk teori, og den derfor ikke kan redegøre for flere kendsgerninger ved vores mentale liv, heriblandt det at vores mentale liv er rationelt (Fodor, 2001: 16, 18, Fodor & Pylyshyn, 1988). Ifølge Fodor er den symbolmanipulerende forståelse i kognitionspsykologien, på trods af kritikken af den og på trods af konnektionismens udbredelse, således reelt "*the only game in town*" (Fodor, 1998: 23).

Den symbolmanipulerende forståelse af vores kognitive system kaldes for *The Computational Theory of Mind* (CTM) eller *The Classical Computational Theory of Mind* for at adskille den fra den konnektionistiske computer-forståelse (Web 1, Fodor & Pylyshyn, 1988). Jeg kalder den CTM fremover, og lader det være op til læseren at huske at adskille den fra den konnektionistiske forståelse.

På trods af at Fodor var en af de mest fremtrædende fortalere for CTM, var han imidlertid samtidig en der gjorde meget ud af argumentere for at CTM sandsynligvis kun kan være en underlæggende teori for en meget begrænset del af vores kognitive processer. Ifølge Fodor er der meget der tyder på, at vi kun kan betragte vores demonstrative² slutningsprocesser og de indledende

² Jeg bruger betegnelsen *demonstrative slutninger* til at referere til slutninger der er logisk gyldige eller nødvendige. 'Hvis P er tilfældet, er Q tilfældet. P er tilfældet, altså er Q tilfældet' er for eksempel en demonstrativ slutning. Ikke-demonstrative slutninger er slutninger der ikke er logisk nødvendige: 'Jeg har set 100 svaner. De har alle været hvide, altså er alle eksisterende svaner er hvide'.

processer i vores perceptionsprocesser og sprogforståelsesprocesser som computer-processer i CTMs forstand (Fodor, 2001: 38-39, Fodor, 1983: 107). I *The Mind Doesn't Work That Way* (2001) udtrykker han, at problemerne for CTM er så alvorlige, at han anbefaler at kognitionsforskere udelukkende fokuserer på den lille del af vores kognitive processer, der kan betragtes som computer-processer og ellers venter på, at der er en der kommer på en god ide til hvordan vi skal forstå resten (Fodor, 2001: 52-53).

Han fremførte første gang argumentationen for problemerne ved CTM i *The Modularity of Mind* (1983), og gav en uddybning flere gange efterfølgende. Især i *The Mind Doesn't Work That Way* fokuserer han på problemerne ved CTM, men også i *LOT 2: The Language of Thought Revisited* (2008) fremlægger han en argumentation. I "Reply to Steven Pinker 'So How Does the Mind Work?'" (2005) forsvarer han desuden kort sin argumentation.

Da Fodor ikke mente at konnektionisme eller den fysikalistisk tilgang, der reducerer mentale tilstande til typer af fysiske tilstande, kan udgøre et grundlag for kognitionspsykologien, er det ifølge Fodor et problem for kognitionspsykologien og bevidsthedsfilosofien i det hele taget, at CTM kun gælder en meget begrænset del af vores kognitive processer (Fodor, 2001: 46-52, Fodor, 1978: 1-26). For hvordan skal vi ellers forklare vores rationelle mentale liv og sammenhængen mellem det mentale og det fysiske (hjernen)? (Fodor, 2008: 125, Fodor, 2001: 38-39).

Susan Schneider, som er professor i filosofi ved University of Connecticut, mener imidlertid, at Fodor er for pessimistisk omkring CTM. I *The Language of Thought: A New Philosophical Direction* (2011) formulerer hun en gentænkning og videreudvikling af Fodors filosofiske grundlag for kognitionspsykologien. Her fremfører hun blandt andet flere indvendinger mod Fodors argumentation for at CTM kun kan gælde for en begrænset del af vores kognitive processer og konkluderer, at der ikke er grund til at være så pessimistisk omkring CTM som Fodor er (Schneider, 2011: 63). Dette på trods af at den empiriske forskning stort set kun har produceret fremskridt i forståelsen af vores perceptionsprocesser, sprogforståelsesprocesser og demonstrative slutningsprocesser (Schneider, 2011: 61-62, 89-90, se også Eysenck & Keane, 2013). Jeg vil i denne afhandling undersøge, om Schneiders indvendinger mod Fodors argumentation for CTMs begrænsede gyldighedsområde er holdbare og om vi således har mindre grund til at være pessimistiske omkring CTM. Jeg vil med andre ord svare på spørgsmålet:

Er Susan Schneiders indvendinger mod Jerry Fodors argumentation for at CTM kun kan gælde en begrænset del af vores kognitive processer holdbare?

Spørgsmålet er kun rettet mod den argumentation som Fodor har fremlagt i *Modularity of Mind (MOM)*, *The Mind Doesn't Work That Way (TMDWTW)*, *LOT 2: The Language of Thought Revisited (LOT 2)* og i "Reply to Steven Pinker 'So How Does the Mind Work?'" og den argumentation som Schneider har fremlagt i *The Language of Thought: A New Philosophical Direction*. Har Schneider eller Fodor fremlagt anderledes argumentation andre steder, er dette altså ikke taget i betragtning i min afhandling.

Det skal derudover understreges at jeg udelukkende undersøger Fodors kritik af CTM og Schneiders forsvar for CTM. Jeg undersøger således ikke hvorvidt konnektionisme eller andre filosofiske grundlag for en videnskab om vores kognitive processer, kan løse de problemer, eller slet ikke har de problemer, som CTM har ifølge Fodor. Fodor har imidlertid fremført grundig argumentation for at de to store alternativer til CTM, en reduktionistisk fysikalisme og konnektionisme, på hver deres måde er meget mere utilstrækkelige tilgange til en forståelse af vores kognitive processer end CTM (se for eksempel Fodor (1968), Fodor og Pylyshyn (1988) og Fodor (2001, s. 46-52)).

Afhandlingens opbygning

For at svare på om Schneiders indvendinger mod Fodor er holdbare, vil jeg først redegøre for det grundlæggende i CTM, så det er klart, hvad der diskuteres. Redegørelsen vil hovedsageligt tage udgangspunkt i Fodors fremlægning af CTM, men den vil også tage udgangspunkt i dels Pinkers beskrivelse i *How the Mind Works* (1997) og dels Kim Sterelnys beskrivelse i *The Representational Theory of Mind*. Derudover vil jeg i redegørelsen give en beskrivelse af en såkaldt turingmaskine. Denne beskrivelse hviler til dels på Alan M. Turings beskrivelse i "On Computable Numbers, With and Application to the Entscheidungsproblem" (Cooper and Leeuwen, 2013).

Efter redegørelsen for CTM, vil jeg redegøre for Fodors argumentation for, at CTM sandsynligvis ikke kan udgøre et grundlag for hvordan vores kognitive system generelt fungerer. Hans argumentation består overordnet i to forskellige argumenter. Disse argumenter har imidlertid flere ting til fælles, og Fodor skelner kun tydeligt mellem dem i *LOT 2*. Schneider mener, i modsætning til Fodor, at de er radikalt forskellige argumenter og forholder sig derfor til dem adskilt. Jeg gengiver dem derfor også som adskilte argumenter. I *LOT 2* fremfører Fodor det ene argument under overskriften "Relevance" og det andet under overskriften "Globality". Det ene argument kalder

jeg således for relevansargumentet og det andet for globalitetsargumentet. Dette er i overensstemmelse med Schneider.

Jeg redegør først for relevansargumentet og derefter for globalitetsargumentet. Globalitetsargumentet er delt i to gengivelser. Det skyldes for det første, at Fodor fremlægger argumentationen på én måde i *MOM* og på en anden i *TMDWTW*, mens argumentationen i *LOT 2* er en mindre detaljeret form for blanding af argumentationen i *MOM* og *TMDWTW*. For det andet, er der tvetydigheder i de forskellige formuleringer, som jeg mener kan tydeliggøres og afklares gennem en opdelt gengivelse af globalitetsargumentet. Tvetydighederne i globalitetsargumentet gør desuden, at min gengivelse i højere grad er en fortolkning af hans argumentation, end relevansargumentet er.

Efter redegørelsen og fortolkningen af Fodors argumenter, redegør jeg for Schneiders indvendinger og giver mine svar på dem. Schneider fremfører to former for indvendinger mod relevansargumentet og tre former for indvendinger mod globalitetsargumentet. Jeg redegør for hver indvending og giver efter hver redegørelse mit svar til Schneiders indvendinger. Det ene argument mod relevansargumentet, baserer sig på den kognitionspsykologiske teori Global Workspace Theory. For at præcisere det grundlæggende i denne teori, vil min gengivelse af teorien til dels basere sig på Bernard Baars og Murray Shanahans fremstilling af teorien i "Applying Global Workspace Theory to the Frame Problem" (2005).

Jeg argumenterer for at begge indvendinger mod relevansargumentet er forfejlede. Begge indvendinger baserer sig, som jeg ser det, på en misforståelse af argumentet, hvilket fører til, at indvendingerne ikke er indvendinger mod Fodors argument. Den ene indvending kan imidlertid omformuleres, så den udgør et angreb på en præmis i min forståelse af relevansargumentet. Her argumenterer jeg for, at den omformulerede indvending ikke svækker relevansargumentet.

Efterfølgende argumenterer jeg for, at alle Schneiders tre argumenter mod globalitetsargumentet er forfejlede, men at hun trods alt får stillet alvorlige spørgsmålstejn ved en del af Fodors argumentation. Den første indvending mener jeg ikke angriber globalitetsargumentet, sådan som jeg forstår det. Den anden indvending baserer sig på en misforståelse af en afgrænset del af Fodors argumentation og angriber dermed en præmis jeg ikke mener, der er i Fodors argument. Den tredje indvending er mere kompliceret. Her går Schneider først i dybden med en del af Fodors argumentation i *TMDWTW* og fremfører indvendinger mod denne del af argumentationen. Derefter omformulerer hun samme del af Fodors argumentation i tilfældet af, at hendes forståelse af Fodors argumentation baserer sig på formuleringer, der fremlægger argumentet anderledes end Fodor havde

intenderet det (det vil sige, Schneider mener det er muligt at Fodor har været uopmærksom og formuleret sig forkert). Hun fremfører derpå en indvending mod denne omformulering af Fodors argumentation. Til sidst fremfører hun indvendinger mod resten af Fodors argumentation i globalitetsargumentet. Jeg mener, at hun får gendrevet den del af Fodors argumentation som hun går i dybden med, uden at omformulere Fodors argumentation. Jeg mener imidlertid, at hendes omformulering af argumentationen ikke bliver gendrevet. Hun forholder sig ikke til en af hovedpointerne i resten af Fodors argumentation, hvilket gør, at hendes gendrivelse af en præmis i det omformulerede argument, ikke er tilstrækkeligt til at gendrive globalitetsargumentet. Schneider får til gengæld stillet alvorlige spørgsmålstejn ved denne omformulering af argumentet.

I forhold til hendes indvendinger mod resten af Fodors globalitetsargument, mener jeg for det første, at disse indvendinger er svage og for det andet viser hendes indvending mod den omformulerede del af globalitetsargumentet, at hendes indvendinger mod resten af globalitetsargumentet ikke er holdbare.

Jeg vil til sidst komme med konkluderende bemærkninger, der relaterer min afhandling til min indledende interesse i at forstå psyken og hvordan den kan undesøges.

Det grundlægende i CTM

CTM er som sagt en teori, der hævder at menneskers kognitive processer kan forstås som en form for computer-processer. Jeg vil i det næste redegøre for det grundlæggende i denne teori. Jeg vil først redegøre for hvordan CTM udgør et grundlag for en videnskab, der tager udgangspunkt i vores common-sense psykologi. Derefter vil jeg redegøre for teorien om at vi er i besiddelse af et language of thought (LOT), et tankesprog. CTM hænger uløseligt sammen med denne teori. Ud fra dette vil jeg præcisere hvad der ligger i CTMs påstand om at vores kognitive system fungerer som en computer.

Common-sense psykologi som videnskab

CTM udgør som sagt et grundlag for en videnskab, der tager udgangspunkt i vores common sense-psykologi. Med common-sense psykologi refereres der til det, at vi til dagligt forklarer og forudsiger menneskers mentale liv og deres handlinger gennem deres epistemiske indstillinger (propositional

attitudes) (Fodor, 1987: 9-10). Epistemiske indstillinger er mentale tilstande som det at antage noget, det at ønske noget, det at frygte noget. Det er med andre ord mentale tilstande, der består i en bestemt indstilling til et udsagn (Fodor, 1987: 11). For eksempel kan man antage (indstilling) at jorden er flad (udsagn), ønske (indstilling) at AC Milan vinder Champions League (udsagn), og frygte (indstilling) at bjørnen vil løbe efter en (udsagn). Fodor hævder at forklaringerne og forudsigelserne i vores common-sense psykologi baserer sig på tre forskellige slags årsagsforbindelser til epistemiske indstillinger: 1) Årsagsforbindelser mellem vores ydre omgivelser og epistemiske indstillinger, 2) årsagsforbindelser mellem epistemiske indstillinger og andre epistemiske indstillinger og 3) årsagsforbindelser mellem epistemiske indstillinger og vores handlinger (Fodor, 1987: 12). En typisk forklaring med udgangspunkt i epistemiske indstillinger kunne for eksempel være, at forklare det, at jeg forleden gik over til min ven og at jeg gjorde det med en paraply i hånden, på følgende måde:

Jeg *ønskede* at besøge min ven.

Jeg *antog* at jeg kunne gå derover på 15 minutter og at jeg ikke havde andre transportmuligheder til rådighed.

Jeg *sansede* at skyerne på himlen udenfor mit vindue var mørke.

Jeg *antog derfor* at der var stor sandsynlighed for at det ville regne inden længe.

Jeg *antog derfor* at jeg med stor sandsynlighed ville gå i regnvej, hvis jeg inden længe gik ud ad døren mod min ven.

Jeg *antog* derfor at der var stor sandsynlighed for at jeg ville blive våd, hvis jeg valgte at gå over til min ven inden længe.

Jeg *ønskede* ikke at blive våd.

Jeg *antog* at en paraply kunne skærme for regnen.

Jeg *ønskede* derfor at medbringe en paraply.

Et argument for at vores common sense-psykologi er et godt udgangspunkt for en videnskab, er ifølge Fodor, at vi mennesker med udgangspunkt i antagelser om andres epistemiske indstillinger, hele tiden forudsiger og forklarer menneskers handlinger med forholdsvis stor succes. En ven fortæller mig, at han vil være på Nørreport St. kl. 15 og jeg konkluderer, at der vil han formentlig være omkring kl. 15. Det gør jeg fordi mennesker som regel ønsker at udtrykke deres (reelle) tanker og intentioner når de kommunikerer sprogligt (og derfor som regel rent faktisk kommunikerer deres tanker og intentioner), og fordi jeg ud fra en persons tanker og ønsker ofte med stor succes kan udlede deres

handlinger (Fodor, 1987: 3). Når min ven fortæller mig at han vil være på Nørreport kl. 15, antager jeg, at han ønsker at være på Nørreport kl. 15 og ønsker han at være der, vil han formentlig forsøge at være der kl. 15. Selv hvis det ikke var en ven, ville jeg tænke det samme og kunne forudsige personens handlinger med samme succes (Fodor, 1987: 3)

Selvom vores common sense-psykologi er forholdsvis succesfuld i dets forudsigelser, er den imidlertid (som med alle common sense-antagelser) behæftet med mange fejl og overgeneraliseringer. Det er for eksempel, ifølge Fodor, common sense at betragte de fleste af vores mentale tilstande og processer som bevidste, og at betragte meget af det, der er i psyken som noget, der er tillært. Dette er der meget psykologisk forskning og teori, der tyder på ikke er sandt.

Med common sense-psykologiens mangler og fejl, er der således brug for en videnskab, der empirisk kan undersøge de egentlige sammenhænge mellem vores epistemiske indstillinger (Fodor, 1987: 15). Det er altså denne videnskab CTM udgør et filosofisk grundlag for. CTM giver et grundlag for hvordan vi skal forstå epistemiske indstillinger og for hvordan vi skal forstå de kausale forbindelser imellem dem.

The Language of Thought

CTM bygger på teorien om at vi har et tankesprog. At vi har et tankesprog, betyder at vi har et mentalt repræsentationssystem, der kan generere uendeligt mange repræsentationer (Fodor, 1987: 16-17). Når vi for eksempel beslutter os for at tage en paraply med i hånden inden vi går ud ad døren og over til en ven, så har vores mentale repræsentationssystem forud for dette genereret mentale repræsentationer af verden omkring os på en bestemt måde. Vi har for eksempel haft mentale repræsentationer af mørke skyer og af hvordan vi måske bliver våde hvis vi går over til vores ven uden en paraply. Vores mentale repræsentationer kan sidestilles med empirismens beskrivelse af ideer, men Fodor pointerer at der med mentale repræsentationer ikke menes indre *billeder*, som empirismen hævdede (Fodor, 1998: 7). Ifølge Fodor ligner vores mentale repræsentationer i stedet generelt ord og sætninger. Det der er karakteristisk for vores mentale repræsentationer er nemlig, at de udgør en form for udsagn om verden. Vi repræsenterer verden omkring os mentalt gennem ideer som 'skyerne er mørke', 'det vil komme til at regne', 'jeg kan gå over til min ven på 10 minutter' og så videre. Disse ideer har udsagnsmæssig karakter i og med, at de udgør eller repræsenterer noget, der kan være enten sandt eller falsk (Fodor, 1978: 31-32). Det er for eksempel sandt eller falsk at skyerne er mørke. For at en ide kan udgøre noget der er sandt eller falsk, må den specificere en

egenskab eller tilstand og noget der tilskrives denne egenskab eller tilstand ('er mørke' er tilstanden og 'skyerne' er det der tilskrives egenskaben i mit ene eksempel). Hvis man for eksempel udelukkende havde et indre billede af det mentale udsagn 'der sad en hund i en bil' ville objektet og tilstanden ikke være specificeret. Et indre billede viser nemlig altid et uendeligt antal forskellige tilstande eller egenskaber. Man kan for eksempel ikke have et billede af 'en bil med en hund i' uden bilen og hunden har en form og størrelse og forskellige andre egenskaber. Ved at 'kigge' på et indre billede af en hund i en bil ville man derfor ikke være i stand til at afgøre hvilket udsagn dette billede skulle udtrykke. Man ville ikke vide om det var 'den lille hund sad på forsædet' eller 'hunden var i bilen' eller 'hunden var i bilen alene' eller et af de uendelig mange andre udsagn, der kan udtrykkes med det billede. Hvis et billede af en hund i en bil bruges som et udsagn, er det med andre ord ikke klart hvad der skal til for at udsagnet er sandt (Fodor, 1978: 180-182). Udsagnet 'der sad en hund i en bil' er imidlertid ikke tvetydigt på denne måde. Hvis der er en hund i en bil, er udsagnet sandt (Fodor, 1978: 181). Vores mentale repræsentationer kan altså ikke udgøres af indre billeder alene. Mere specifikt kan de mentale repræsentationer, der har *udsagnsmæssig* karakter, ikke udgøres af indre billeder. Til gengæld mener Fodor ikke at der er noget problem i at vores ideer der ikke udgør udsagn, men udelukkende refererer til et eller andet, sådan som ideen 'bil' refererer til en bestemt slags genstand, *til tider* kan udgøres af en form for indre billeder, så længe der er udsagnsmæssige ideer der specificerer hvad billedet refererer til (Fodor, 1978: 182).

At vores mentale repræsentationer har udsagnsmæssig karakter, implicerer at vores repræsentationssystem på mange måder ligner vores rigtige sprog. For det første må vores indre, mentale repræsentationssystem bestå af et stort, men afgrænset, vokabular af repræsentationer, der minder om ordene i vores sprog (Fodor, 1978: 31-32, 123). Vi har for eksempel mentale repræsentationer, der svarer til ordene 'bil', 'hund', 'er', 'i' og så videre. Disse mentale ord kan givetvis kombineres så de udgør udsagn. Deres kombinatoriske karakter gør, for det andet, at vi kan producere uendeligt mange nye mentale sætninger, som vi aldrig har tænkt før, ligesom vi kan producere uendeligt mange sætninger vi aldrig har hørt eller læst før i vores rigtige sprog. Vores mentale repræsentationssystem er altså, ligesom vores rigtige sprogsystem, produktivt (Fodor, 1978: 31). For eksempel kan jeg uden problemer skrive, sige og tænke sætningen 'I går kørte brandbilen ind i bjerget af basketball-bolde, der lå midt i lyskrydset hvor Jagtvej og Tagensvej krydser'. Det er en sætning jeg aldrig har skrevet, sagt eller tænkt før og jeg kan producere nye sætninger i det uendelige (i princippet, ikke i praksis selvfølgelig), som jeg kan forbinde med denne sætning. Jeg kan for eksempel forlænge sætningen på følgende måde: 'I går kørte brandbilen ind i bjerget af basketball-

bolde, der lå midt i lyskrydset hvor Jagtvej og Tagensvej krydser og hvor en klovn stod ved siden af bjerget og jonglerede med vaffelis, der var købt i Hellerup af klovnenes kæreste, der arbejdede som læge på Gentofte sygehus, der...'

På trods af at vores mentale repræsentationssystem på mange måder ligner et rigtigt sprog, er det vigtigt at pointere at det ifølge Fodor, ikke *er* et rigtigt sprog. Vi tænker ikke i det (eller de) sprog vi taler. Vores repræsentationssystem er et form for *tankesprog* bag vores rigtige sprog. Fodor påpeger for eksempel, som argumentation for at vi ikke tænker i det sprog vi taler, at det synes mere end almindeligt sandsynligt at visse dyr og børn også repræsenterer verden mentalt (Fodor, 1978: 56).

Teorien om at vi på denne måde er i besiddelse af et mentalt repræsentationssystem eller et tankesprog, et *language of thought* (LOT), udgør et grundlag for hvordan CTM forstår epistemiske indstillinger. Kort sagt består en epistemisk indstilling ifølge CTM i en relation mellem en organisme og en mental repræsentation. Det vil sige for hver en organisme O og for hver en epistemisk indstilling E mod et udsagn U, er der en relation R og en mental repræsentation MR, hvor MR betyder P, og O har E hvis O har R til MR (Fodor, 1987: 17). En organisme der antager at AC Milan har vundet Champions League, er således i en bestemt relation (en 'antagende' relation) til den mentale repræsentation 'AC Milan har vundet Champions League'. En organisme der ønsker at AC Milan vinder Champions League, er således i en anden relation (en 'ønskende' relation) til den mentale repræsentation 'AC Milan vinder Champions League'. Man kan også beskrive det som at en person har en repræsentation i sin *antagelsesboks* eller sin *ønskeboks* eller sin frygtboks etc. (Fodor, 1998: 8).

Fodor bemærker i *The Language of Thought*, at denne påstand om at epistemiske indstillinger består i en relation mellem en organisme og en mental repræsentation, imidlertid ikke kan gælde helt generelt. Den epistemiske indstilling *at vide noget*, er for eksempel afhængig af hvordan verden rent faktisk er, så det at vide noget består ikke udelukkende i en relation mellem en organisme og en mental repræsentation. Verdens tilstand har også betydning for hvorvidt man rent faktisk *ved* noget (Fodor, 1978: 75-76, note 16). At beskrive epistemiske indstillinger som noget, der består i en relation mellem en organisme og en mental repræsentation, er imidlertid en tilstrækkelig beskrivelse af den grundlæggende forståelse af epistemiske indstillinger i CTM.

Fodor bemærker desuden at CTM er neutral omkring hvorvidt epistemiske indstillinger *er forårsaget af* relationer mellem organismer og mentale repræsentationer eller om epistemiske indstillinger *korresponderer med* relationer mellem organismer og mentale repræsentationer. CTM

har altså ikke et standpunkt omkring epistemiske indstillinger (selvstændige) metafysiske karakter (Fodor, 1978: 77).

Det skal derudover pointeres at når CTM hævder at vi har et tankesprog eller et mentalt repræsentationssystem, så består disse repræsentationer ikke nødvendigvis i *bevidste* ideer. CTM og LOT er neutral omkring repræsentationers fænomenologiske karakter. De kan lige så godt være bevidste, som ubevidste (Fodor, 1978: 52-53). En organisme kan således godt antage en masse ting uden at have fænomenologiske oplevelser, der svarer til disse antagelser. Fodor formoder for eksempel at langt de fleste af vores kognitive processer, det vil sige kausale forbindelser mellem epistemiske indstillinger, foregår ubevidst (Fodor, 1983: 104).

CTM

Mens LOT for sig selv giver en teori om hvad vores epistemiske indstillinger er og hvordan vi skal forstå vores repræsentationer af verden omkring os, giver CTM, sammen med LOT, en teori om hvordan epistemiske indstillinger *forårsager* hinanden. Mere præcist udgør CTM sammen med LOT en teori om de kausale mekanismer, der ligger bag vores *tænkning* (Fodor, 1998: 9-10). Her er tænkning forstået meget bredt. Både perceptionsprocesser, læringsprocesser, problemløsningsprocesser og almindelige slutningsprocesser betragtes blandt andet som tænkning (Fodor, 2008: 113). Vores tænkning dækker altså over alle de bevidste og ubevidste processer, der udgør vores eller vores kognitive systems forsøg på at finde ud af hvordan verden omkring os (og 'i' os) tager sig ud og hænger sammen. Med andre ord er tænkning de processer, der fører til fikseringen af vores antagelser (Fodor, 1978: 52-53, Fodor, 2008: 113).

Det er ifølge Fodor karakteristisk for vores tænkning at sande antagelser og slutninger fører til *nye* sande antagelser og slutninger. Vores tænkning er med andre ord rationel (Fodor, 1987: 14). De mentale processer der gik forud for min beslutning om at gå over til min ven med paraply i hånden, udgør et eksempel på sådanne sandhedsbevarende, rationelle kæder af antagelser:

Jeg sansede at skyerne på himlen udenfor mit vindue var mørke.

Jeg antog *derfor* at der var stor sandsynlighed for at det ville regne inden længe

Jeg antog *derfor* at jeg med stor sandsynlighed ville gå i regnvej, hvis jeg inden længe gik ud ad døren mod min ven.

Jeg antog *derfor* at der var stor sandsynlighed for at jeg ville blive våd, hvis jeg valgte at gå over til min ven inden længe.

Det der gør dette til en sandhedsbevarende tankeproces er, at det som regel er sandt, at hvis skyerne er mørke, vil det regne inden længe, og det er nødvendigvis sandt, at hvis det sandsynligvis begynder at regne inden længe, så vil man sandsynligvis gå i regnvejr hvis man går ud inden længe - og så videre. Den ene (formodentligt) sande antagelse fører til den næste. Grunden til 'derfor' er med kursiv, er for at understrege at sandhedsbevarende mentale processer hviler på *indholdet* af antagelserne. De hviler på betydningen af repræsentationerne som mine antagelser retter sig mod. Det er *fordi* jeg antager 'at jeg ville gå i regnvejr', at jeg antager 'at jeg ville blive våd'. Der er altså en bestemt semantisk relation mellem repræsentationerne i mine antagelser. Det er mere præcist dette, at der er en bestemt - en rationel - semantisk relation mellem repræsentationerne i vores tankeprocesser, Fodor mener er karakteristisk for vores tænkning. Det er de kausale mekanismer bag disse rationelle tanker CTM udgør en teori om.

Turingmaskinen

CTM forklarer de grundlæggende mekanismer bag vores rationelle tænkning ved at påstå, at vores kognitive system mere eller mindre fungerer som en computer. Fodor beskriver overordnet en computer som et fysisk system, der er indrettet således at vi kan tilskrive dets fysiske tilstande en betydning, der gør, at de fysisk-kausale processer i systemet kan tolkes som symbolmanipulationer, der respekterer de semantiske relationer mellem symbolerne. Det vil sige, de rationelle forbindelser mellem symbolerne bevares. Når CTM påstår at vores kognitive system fungerer som en computer, påstår den således ifølge Fodor, groft sagt, at vores kognitive system består af et fysisk system, hvis processer kan tolkes som transformationer af symboler, der bevarer den (rationelle/sande) semantiske relation mellem symbolerne (Fodor, 1978: 73, Fodor, 1987: 19). Denne forståelse af vores kognitive system baserer sig hovedsageligt på Turing og hans model over en symbolmanipulerende maskine (Fodor, 2001: 12-14, 18, 24, Fodor, 1983: 39). Vi kan således få en dybere forståelse af hvad der mere præcist ligger i, at vores kognitive system er et fysisk system, der transformerer symboler i overensstemmelse med deres semantiske relation, ved at forstå Turings maskine.

Turing udformede en abstrakt model af en maskine der, som sagt, kan manipulere symboler, og han viste at denne maskine kan evaluere eller beregne sig frem til konklusionen i enhver *formel* gyldig slutning. Det vil sige, den kan beregne enhver slutning hvis gyldighed groft sagt er givet

gennem de *formelle egenskaber* ved udsagnene i slutningen (Fodor, 2001: 12-13, note 4). Slutter man for eksempel fra udsagnene 'Hvis Tobias går i regnvejr, bliver han våd' og 'Tobias går i regnvejr' til 'Tobias bliver våd', er dette en formel gyldig slutning. Slutningen er et eksempel på en modus ponens slutning: P medfører Q, P er tilfældet, altså er Q tilfældet. Det er en formel gyldig slutning i og med at slutningen ville være gyldig ligeegyldig hvilke udsagn vi erstattede P og Q med, så længe de passer i formlen.

At en turingmaskine kan beregne sig frem til konklusionen på en formel gyldig slutning vil med andre ord sige, at den kan manipulere symboler, der bevarer sandhedsrelationen mellem symbolerne. En turingmaskine er altså en computer som beskrevet før.

Maskinen består 1) af et uendeligt langt stykke bånd eller papir inddelt i felter, der kan indeholde symboler og 2) af en enhed eller et hoved, der kan rykke et felt til højre eller venstre og som kan registrere ('læse'), printe og slette symboler på båndet. Registreringen af et symbol og maskinens tilstand, afgør om hovedet skal rykke et trin til højre eller venstre, printe et symbol på båndet eller slette et symbol på båndet og om maskinen skal ændre tilstand. Maskinen modtager symboler skrevet på båndet som input ved at 'læse' disse symboler et felt ad gangen. Den udfører operationer *på baggrund af* symbolerne (det vil sige på baggrund af formen på det læste. Hvis det er en fysisk realiseret maskine, vil det altså være på baggrund af symbolets fysiske form eller struktur (Fodor & Pylyshyn, 1988: 12-13, Fodor, 1987: 18)). Operationerne den udfører på baggrund af symbolerne, er derudover afhængig den tilstand maskinen er i, når den læser symbolerne. Maskinen læser et symbol og trigges derpå til at udføre en operation. Operationerne består i at rykke et felt til højre eller venstre, printe eller slette ét symbol. Maskinen kan læse allerede registrerede symboler flere gange, hvis det er nødvendigt (hvorvidt den gør det, er som sagt afhængigt af maskinens tilstand og de symboler det læser). Båndet med symboler udgør således også en form for informationslager, som maskinen kan 'konsultere' i sine beregninger (Cooper & Leeuwen, 2013: 17). En turingmaskine er altså overordnet en maskine, der kan være i forskellige tilstande og som kan registrere forskellige symboler og som transformerer de symboler maskinen læser afhængigt af den tilstand maskinen er i og afhængig af de symboler den registrerer.

En turingmaskine kan givetvis være indrettet på mange måder. Det vil sige én maskine kan være indrettet til at kunne være i én gruppe af tilstande, at kunne registrere én gruppe af symboler og at kunne udføre én gruppe af operationer, mens en anden kan være indrettet til at kunne være i en anden gruppe af tilstande, registrere en anden gruppe af symboler og så videre. De operationer en given maskine er indrettet til at udføre, kan beskrives med en maskine-tavle. Maskine-tavlen angiver

hvilke symboler maskinen kan modtage som input og hvad maskinen gør, hvis den er i en bestemt tilstand og læser et bestemt symbol.

Et eksempel på en maskine-tavle kunne være:

Maskinens tilstand (til højre) Input (nedenfor)	1	2	3	4	5	6	7
Y	_L1	_L1	YL4	YR5	YR6	YR6	_R7
_	_L1	YR2	halt	YR5	YL3	AL3	YR6
1	1L2	AR2	AL3	1L7	AR5	AR6	1R6
A	1L1	YR6	1L4	1L4	1R5	1R6	_R2

En maskine, der beskrives med denne maskine-tavle, kan modtage fire forskellige symboler som input, og være i syv forskellige tilstande. De to første symboler i hver celle (hvoraf maskinen kun kan læse det første) angiver hovedets operation/output og det sidste symbol angiver den tilstand maskinen skal skifte til. AR2 betyder således Print A, ryk ét felt til højre (R = right) og skift til tilstand 2. (Sterelny, 1990: 7).

Det Turing viste, da han viste at en turingmaskine kan evaluere formelt valide slutninger, var således, at der kan indrettes en maskine, der følger en maskine-tavle som 'sikrer', at hvis maskinen læser nogle (læselige) symboler, der repræsenterer præmisser i en formel valid slutning, vil maskinen transformere disse til symboler, der repræsenterer konklusionen i slutningen. Ifølge Fodor kunne Turing vise dette, fordi symbolernes logiske form kan repræsenteres fysisk, og fordi der kan udformes algoritmer eller 'beviser' for slutninger, der hviler på symbolers logiske form (Fodor, 2001: 12, 13 18, Fodor, 1978: 73).

At en turingmaskines operationer følger en maskine-tavle eller et sæt algoritmer betyder desuden at hvis maskinen blev realiseret fysisk, ville dens operationer således ikke kun kunne beskrives fysisk, men også gennem maskine-tavlens symboler. Det betyder, at en turingmaskine - eller de grundlæggende elementer at læse symboler og transformere disse i overensstemmelse med algoritmer - i princippet kan realiseres fysisk på mange forskellige måder (Fodor, 1978: 15, 74) .

Vores elektroniske computere, der kan være bygget på forskellige måder, men udfører de samme algoritmiske operationer, udgør et eksempel på dette (Fodor & Pylyshyn, 1988: 63).

Med dette, og med en uddybning af hvad CTM forstår ved mentale repræsentationer, kan det nu præciseres hvad der ligger i at vores kognitive system ligner en computer og hvordan computer-analogien udgør et grundlag for de kausale mekanismer bag vores tænkning. For det første kan vi beskrive de grundlæggende elementer ved en (fysisk realiseret) turingmaskine som:

Et (fysisk) system

- der, til et givet tidspunkt, er i en bestemt tilstand
- og på den ene eller den anden måde registrerer bestemte symboler (det vil sige fysiske entiteter med semantiske egenskaber, der er repræsenteret gennem symbolets fysiske struktur og logiske form)
- (som “gemmes” og)
- som udløser nogle fysiske processer, der korrelerer med bestemte algoritmer
- som (er algoritmer, der) pålideligt transformerer de registrerede symbolerne til nye symboler hvor den semantiske relation mellem symbolerne bevares.
- Dette system kan i princippet realiseres fysisk på mange forskellige måder.

Vores kognitive system kan altså ifølge CTM beskrives med disse grundlæggende elementer. Dette betyder, for det andet, at vores kognitive system må indeholde fysiske symboler, der manipuleres. CTMs påstand er, at vores mentale repræsentationer i vores tankesprog, *er* symboler. De er fysisk, afgrænsede entiteter med semantiske egenskaber, og vores kognitive system, er et system, der manipulerer disse entiteter (Fodor, 1998: 10). Mere præcist udgøres hvert slags mentale ‘ord’, i vores mentale ordforråd, af en fysisk entitet med en unik fysisk struktur, som repræsenterer det i verden ordene refererer til (eller en operator som ‘og’, ‘eller’, ‘plus’ etc.). Ethvert mentalt ords unikke fysiske struktur repræsenterer det ordet refererer til ved at den unikke struktur korrelerer med de formelle egenskaber ved det en repræsentation refererer til (det vil sige, de kategorier, det der referes til, hører under) og med dets grammatiske egenskaber (er det for eksempel et substantiv, et verbum etcetera) (Schneider, 2011: 11-12, 67, Fodor & Pylyshyn, 1988: 12-13). Disse mentale repræsentationer, og deres fysiske struktur, kan sættes sammen på forskellige måder, og udtrykker, gennem bestemte sammensætninger, udsagn der har en logisk form (Schneider, 2011: 11-12, Fodor & Pylyshyn, 1988:

12-13). Det er disse udsagns logiske form som vores kognitive system kan 'læse' og manipulere gennem den fysiske struktur, der repræsenterer den logiske form (Fodor, 1998: 9). Vores kognitive system er altså en form for turingmaskine, der læser vores mentale sætninger og transformerer dem til nye sætninger. Den fysiske og formelle struktur som disse sætninger udgøres af, og som er det, der trigger vores kognitive maskineri, kaldes symbolernes *syntaks* (Fodor, 1987: 18). Vores kognitive system er således en syntaks-drevet maskine.

Ved at påstå at vores kognitive system på denne måde fungerer som en computer, påstår CTM således at de kausale mekanismer bag vores sandhedsbevarende tanker, er *fysisk kausale mekanismer mellem (afgrænsede) symboler (mentale sætninger) i et symbolmanipulerende system, der kan beskrives med algoritmer*. Algoritmerne består i en beskrivelse af de mentale repræsentationer maskinen skal producere og de tilstande den skal skifte til, når den, i en bestemt tilstand, 'læser' en mental repræsentation. Med andre ord svarer vores epistemiske indstillinger til forskellige tilstande i et symbolmanipulerende system. En persons (organismes) ønsker kan for eksempel beskrives som måltilstande eller instruktive tilstande i personens symbolmanipulerende system. En persons (baggrunds)antagelser kan beskrives som de mentale repræsentationer personens symbolmanipulerende system har indskrevet i sit informationslager. En persons perceptioner, kan beskrives som indskrevne mentale repræsentationer, der er registreret gennem sensorer (Pinker, 1997: 78, Fodor, 1978: 75).

Et eksempel kan måske illustrere hvordan det hele hænger sammen.

Et eksempel

En person, Pernille, der forundret ser sin hund i sin bil på trods af hun ikke har lukket hunden derind, kan for eksempel ønske at finde ud af om hunden selv har åbnet en bildør og hoppet ind. Ud fra CTM kan vi give hendes kognitive situation følgende hypotetiske beskrivelse: Pernille, eller hendes kognitive system, er i tilstanden 'find svar på om O foretog handlingen H'. Hendes tankeproces fra 'Min hund kan ikke åbne en bildør' til 'Min hund har ikke åbnet bildøren selv' kan derudover beskrives som en meget simpel algoritme, hvis vi forestiller os at hendes kognitive system er indrettet således, at det kun kan læse tre forskellige mentale repræsentationer: 1) 'O kan ikke foretage handlingen H', 2) 'O kan foretage handlingen H' og 3) 'Mangler viden om hvorvidt O kan foretage handlingen H'. Ved at registrere en mental repræsentation af formen 'O kan ikke foretage handlingen H', i det her tilfælde 'Min hund kan ikke foretage handlingen at åbne en bildør', transformerer hendes

kognitive system denne mentale repræsentation til output-repræsentationen 'O foretog ikke handlingen H', det vil sige 'Min hund foretog ikke handlingen at åbne en bildør'. Ved at registrere en af de to andre repræsentationer, transformerer hendes kognitive system repræsentationen til den mentale repræsentation 'måske foretog O handlingen H'. Hendes tankeproces kan således beskrives med følgende meget simple algoritme: Hvis tilstanden er 'Find svar på om O foretog handlingen H', og hvis 1) registreres som input, produceres som output: 'O foretog ikke handlingen H'. Hvis 2) eller 3) registreres, produceres som output: 'Måske'.

Eksemplet her illustrerer (forhåbentligt) for det første hvordan tænkning ifølge CTM kan beskrives gennem symbolmanipulerende algoritmer, som et fysisk system kan udføre. Det er forholdsvis let at forestille sig at et fysisk system kan være indrettet sådan, at registreringen af en fysisk afgrænset struktur, der repræsenterer 'O kan ikke foretage handlingen H' automatisk producerer en anderledes fysisk struktur der repræsenterer 'O foretog ikke handlingen H', og ved registreringen af en fysisk struktur der repræsenterer 'O kan foretage handlingen H' automatisk producerer en fysisk struktur der repræsenterer 'måske foretog O handlingen H'. Et symbolmanipulerende system består af en masse af sådanne automatiske symboltransformerende processer (Pinker, 1997: 71-76).

For det andet kan eksemplet være med til at illustrere et bestemt element ved en symbolmanipulerende maskine, der er afgørende for en del af Fodors argumenter for at CTM formentlig ikke er sand. Det er det element at *symbolernes syntaks (deres sammensatte form og dermed deres fysiske struktur) er essentielt for at maskinen overhovedet kan manipulere symbolerne*. Det der gør at 'O kan ikke foretage handlingen H' fører til produktionen af symbolet 'O foretog ikke handlingen H', er at det fysiske system, er i en mål-tilstand, og at det registrerer et symbol, der har den fysiske struktur, som lige præcis repræsenterer syntaksen 'O kan ikke foretage handlingen H'. Det er symbolets syntaks der *trigger* maskinen til at udfører operationen 'O kan ikke foretage handlingen H'. Symbolernes syntaks fungerer altså i et symbolmanipulerende system som nøgler, der låser op til den næste operation (Fodor, 1987: 18-19).

Ikke-demonstrative slutninger

Denne teori om vores kognitive system (at det er et symbolmanipulerende system, der transformerer mentale repræsentationer gennem deres syntaks) hviler som sagt på at Turing har vist at der kan konstrueres en syntaks-drevet maskine som beregner formelt gyldige slutninger, hvilket med andre

ord vil sige demonstrative slutninger. En stor del af vores tænkning består imidlertid ikke i demonstrative slutninger (Fodor, 2008: 113). Selve Pernilles antagelse 'min hund kan ikke åbne en bildør' er formentlig en antagelse, der er blevet fikseret gennem ikke-demonstrative slutninger. At Turing kun viste at en maskine kan beregne demonstrative slutninger, betyder imidlertid ikke i sig selv at der ikke findes symbolmanipulerende systemer, der kan beregne ikke-demonstrative slutninger, som dem vi dagligt foretager os. Hvis man kan opstille algoritmer for ikke-demonstrative slutninger på baggrund af de mentale repræsentationers logiske form, så er det muligt for et symbolmanipulerende system at beregne ikke-demonstrative slutninger. Ifølge Fodor er det en udbredt opfattelse blandt kognitionsforskere at det vil vise sig at vores tænkning - den demonstrative og ikke-demonstrative - generelt *kan* beskrives gennem algoritmer eller regler, på baggrund af mentale repræsentationers logiske form. Med CTM har vi således ifølge disse forskere et grundlag for hele kognitionspsykologien. Et grundlag der overordnet indebærer at kognitionspsykologiens opgave er at finde frem til konstitutionen af vores mentale repræsentationer og at finde algoritmerne for vores kognitive processer (Fodor, 2008: 106). Fodor selv mener, som sagt, at det kun er en begrænset del af vores ikke demonstrative processer, der kan beskrives gennem algoritmer på denne måde (Fodor, 1983: 107).

Opsamling

CTM udgør således et filosofisk grundlag for en videnskab om vores epistemiske indstillinger og den måde de forbindes med hinanden på. En epistemisk indstilling består af en relation mellem en organisme (et fysisk system) og en mental repræsentation. En mental repræsentation i en epistemisk indstilling er et *sammensat* symbol, hvilket vil sige det er en fysisk entitet med semantiske egenskaber, der er repræsenteret gennem symbolets fysisk/formelle konstitution. Sammensætningen kaldes repræsentationens syntaks. De mentale processer - vores tænkning - består i en transformation af de repræsentationer der er indeholdt i vores epistemiske indstillinger. Vores kognitive system er et fysisk system, der er indrettet således at det følger algoritmer, der sikrer, at transformationerne bevarer de semantiske relationer mellem repræsentationerne. Det er repræsentationernes syntaks - den formelle sammensætning af symboler - der 'styrer' algoritmerne og dermed sørger for at de semantiske relationer bevares. Vores kognitive processer kan altså beskrives gennem algoritmer, der beskriver et symbolmanipulerende systems processer.

Det er således psykologiens opgave at finde frem til hvordan vores mentale repræsentationers konstitution og hvilke algoritmer der beskriver transformationerne af vores mentale repræsentationer.

Fodors relevans- og globalitetsargument

Fodor fremfører, som sagt, to forskellige argumenter for, at det kun er en begrænset del af vores kognitive processer, der kan beskrives som computer-processer i CTMs forstand. I *LOT 2* præsenterer han det ene argument under overskriften 'Relevance' og det andet argument under 'Globality'. Jeg kalder, som sagt, derfor det ene for relevans-argumentet og det andet for globalitets-argumentet.

Jeg vil i det næste først redegøre for relevans-argumentet og dernæst globalitetsargumentet. Under redegørelsen for relevansargumentet vil jeg også kort redegøre for to typiske svar eller løsninger på relevansproblemet som Fodor forholder sig til. Dette til dels fordi det giver en uddybende forståelse af argumentet og til dels fordi en del af Schneiders argumentation refererer til en af disse typiske løsninger.

Min gengivelse af globalitetsargumentet er delt i to. Det er det fordi Fodor formulerer argumentet på forskellige måder i *MOM*, *TMDWTW* og *LOT 2* og fordi der fokuseres på forskellige komponenter af argumentet i de forskellige formuleringer.

Relevansargumentet

Til dagligt svarer vores kognitive system hele tiden på spørgsmål som 'hvilke genstande er der omkring mig? (hvad er det firkantede, sorte objekt, der er sat fast på væggen?)', 'hvad var skyld i den lyd? (hvad var skyld i det brag, der kom ude fra køkkenet?)', 'hvordan gør jeg det (hvordan åbner jeg låget på en flaske med toilettrens?)', 'hvordan skete det (hvordan har jeg fået lus)?' og så videre. Slutningerne eller processerne bag disse daglige perceptioner, læringsprocesser og ræsonnementer, der fikserer vores antagelser, er, som nævnt tidligere, typisk ikke-demonstrative. De er derudover i høj grad ubevidste (Fodor, 1983:104). Det er derfor svært at sige noget om deres form. Ifølge Fodor har vi imidlertid god grund til at antage at processerne formentlig har mere eller mindre samme form som vores bevidste processer. Fodor tænker her mere præcist på de bevidste processer, der foregår når vi bedriver videnskab (Fodor, 1983: 104). I videnskaben afkræfter og bekræfter (underbygger)

forskere hypoteser primært gennem ikke-demonstrative slutninger. Givet at en konsekvens af de daglige perceptioner, læringsprocesser og tænkning, er, at vi indoptager nye antagelser og forkaster tidligere antagelser i vores system af baggrundsantagelser (vores antagelsessystem), og givet at det er svært at forestille sig at dette ikke sker gennem ikke-demonstrative slutninger, er der en god grund til at antage at disse processer også foregår gennem af- og bekræftelsesprocesser af hypoteser på nogenlunde samme måde som i videnskaben (Fodor, 2008: 113, 121). Hvad der (derudover) er karakteristisk for videnskabens eksplicite slutningsprocesser, er dermed formentlig også karakteristisk for vores typiske kognitive processer (Fodor, 1983: 104, Fodor, 2008: 113). Ét karaktertræk ved videnskabelige bekræftelser af hypoteser er, ifølge Fodor, at de er det han kalder for isotropiske (Fodor, 1983: 105). At videnskabens bekræftelser er isotropiske betyder, at enhver tidligere videnskabeligt etableret kendsgerning og teori en forsker kender til, kan have betydning for bekræftelsen eller afkræftelsen af en hypotese forskeren undersøger (Fodor, 2008: 115). I princippet kan botanikken begrænse astronomien, hvis bare forskeren kunne finde på en måde at forbinde botanikken med astronomien (Fodor, 1983: 105). Det er for eksempel i princippet muligt, gennem etablerede kendsgerninger, at finde en så præcis forbindelse mellem solens energiudladning og plantelivet på jorden, at man ville kunne teste hypoteser om solens energiudladning ved at observere hvordan plantelivet opfører sig. Grunden til at det i princippet er muligt, er ifølge Fodor, at enhver forsker implicit tilskriver sig den grundlæggende overbevisning, at hele verden hænger sammen kausalt og at ingen endnu ved *hvordan* hele verden er kausalt forbundet (Fodor, 1983: 105). Vi *ved* altså ikke at botanikken ikke på nogen måde begrænser astronomien.

I og med at videnskabens bekræftelser af hypoteser mere eller mindre svarer til mange af de kognitive processer der foregår i dagligdagen, har vi ifølge Fodor god grund til at formode at vores daglige non-demonstrative slutningsprocesser typisk også er isotropiske. I princippet kan alle antagelser i vores system af baggrundsantagelser have betydning for vores kognitive systems bekræftelser af de daglige hypoteser, som systemet bearbejder når det giver os svar på hvordan verden omkring os er. Her opstår problemet for CTM. CTM betragter vores kognitive processer som *symbol*manipulerende processer, og, som det fremgik tidligere, er det et grundlæggende karaktertræk ved disse processer, at de baserer sig på en registrering af de mentale repræsentationers *syntaks*. Vores system af baggrundsantagelser må således ifølge CTM være repræsenteret gennem en masse repræsentationers syntaks. Problemet for CTM er, ifølge Fodor, at vores system af antagelser er *enormt* stort (Fodor, 2001: 31). I *LOT 2* hævder han at det er *uendeligt* stort (Fodor, 2008: 115). Givet at alle antagelser i vores system af antagelser i princippet kan have betydning for vores kognitive

systems bekræftelser af de daglige hypoteser, kan et symbolmanipulerende system umiddelbart kun beregne i hvor høj grad hver mental repræsentation i vores antagelsessystem af- eller bekræfter en hypotese, ved at læse alle mentale repræsentationer i vores antagelsessystem. I og med at vores antagelsessystem er enormt stort eller uendeligt stort, er dette imidlertid ikke muligt “(...) *in, as one says, ‘real time’*” (Fodor, 2008: 115). Det vil sige, et symbolmanipulerende system med et antagelsessystem der har samme omfang som menneskers, kan ikke beregne i hvor høj grad hver af dets antagelser af- eller bekræfter en hypotese, indenfor det tidsinterval vi normalt fikserer vores antagelser (hvis vores system af baggrundsantagelser er uendeligt, vil det slet ikke kunne lade sig gøre inden for noget tidsinterval). Når vores kognitive system skal vurdere hvorvidt en hypotese er sand eller ej (eller hvor sandsynligt det er at den er sand), må der, hvis vores kognitive system er et symbolmanipulerende system, dermed være et eller andet, der på en eller anden måde filtrerer vores antagelser, således at de kognitive processer kun består i beregninger på *bestemte* antagelsers af- eller bekræftelse af en hypotese. Problemet er at vi ikke har nogen ide om hvordan denne filtreringsproces skulle foregå eller hvad der afgør filtreringen (Fodor, 2008: 115). Man kunne hævde at filtreringen sker mere eller mindre tilfældigt (og dermed også hævde at vi ikke kan formulere noget der ligner love omkring vores antagelsesfikseringer), men dette stemmer ikke overens med den grad af succes vi alle har med vores antagelser: “(...) *witness our not all being dead*”, som Fodor bemærker (Fodor, 2008: 116). Vi må altså på en eller anden måde forholdsvis ofte beregne på de antagelser, der er de (mest eller tilstrækkeligt) *relevante* for af- eller bekræftelsen af en hypotese, for at CTM kan være sand. Vi har imidlertid ingen ide om hvad der generelt gør antagelser relevante for bekræftelsen af en hypotese. Der synes ikke at være nogle generelle egenskaber ved mentale repræsentationer, der kan indikere at de er relevante for vurderingen af en hypotese (Fodor, 2008: 116, 124, Fodor, 1983: 114).

Problemet gælder desuden ikke kun for vurderingen af hypoteser (*potentielle* antagelser). Det gælder også for vurderingen af hvorvidt vi skal forkaste eller korrigere antagelser, der allerede er i vores antagelsessystem, som følge af etableringen af en ny antagelse (Fodor, 1983: 112-113). Da vores antagelsesfiksering er isotropisk kan etableringen af en ny antagelse i princippet påvirke enhver antagelse i vores antagelsessystem. Det er, ligesom ved bekræftelsen af en potentiel antagelse, imidlertid givet, at hvis vores kognitive system er et symbolmanipulerende system, så gennemgår vi ikke hele vores antagelsessystem hver gang vi har etableret en ny antagelse. Spørgsmålet er så hvordan vi formår at opdatere vores antagelsessystem så succesfuldt, som vi gør? Hvordan beslutter vi eller vores kognitive system hvilke antagelser, der formentlig skal opdateres givet etableringen af en eller flere nye antagelser? (Fodor, 1983: 114-115).

Problemet kan illustreres hvis man forestiller sig, at man skal designe en robot, der kan udføre almindelige dagligdags handlinger, som for eksempel at ringe til en person. Fodor formulerer dette eksempel i *MOM* (1983). Jeg giver her en opdateret og omskrevet version af eksemplet.

Lad os sige at den robot man designer, blandt andet skal kunne instrueres til at ringe til en person, Mary, for at høre om hun kommer for sent til middag. Man koder derfor robotten således at den 'ved' at den kan få fat i Marys nummer ved at konsultere telefonbogen. Når den instrueres til at ringe til Mary, finder den således først Marys nummer og begynder at taste hendes nummer på telefonen. Lad os sige vi har konstrueret sådan en robot og instruerer den til at ringe til Mary. Den finder således Marys nummer og begynder at taste. I og med robotten taster nummeret, ændrer verden sig. Verdenstilstanden går fra at robotten ikke har trykket på en tast, til at den har trykket på en tast. Denne ændring har formentlig medført andre ændringer. Telefonen er for eksempel formentlig lukket for udefrakommende opkald, telefonen producerede formentlig en lyd lige efter robotten trykkede på en tast, skærmen på telefonen ændrede sig formentlig også og så videre. Når robotten har trykket på en tast, er den altså nødt til at repræsentere antagelsen 'jeg har trykket på en tast på telefonen' og den er dermed også nødt til at vurdere om dens 'antagelsessystem' skal opdateres yderligere, som følge af denne antagelse. Det vil sige, den må vurdere om der er nogle antagelser i systemet, der skal forkastes eller korrigeres. Problemet er at det i princippet kan være alle antagelser, der har ændret sandhedsværdi som følge af, at robotten har trykket på en tast. Dette gør det svært at designe robotten så den relativt ofte netop, eller tilnærmelsesvis, korrigerer den gruppe af antagelser, hvis sandhedsværdi ændres som følge af dens handling. For robotten kan givetvis ikke gennemgå alle sine antagelser for hver en handling, den foretager. I hvert fald ikke hvis den skal foretage handlinger i et tempo, der ligner menneskers. Men hvordan skal den så være designet? Der er givetvis ikke en fast gruppe af antagelser, der kræver revurdering hver gang man foretager en eller anden handling. Den gruppe af antagelser der kræver vurdering, ændres afhængigt af de handlinger der udføres og i hvilke kontekster de udføres. Det er ifølge Fodor også åbenlyst at nye antagelser ikke kommer pakket med information omkring hvilke antagelser, der har en sandhedsværdi som formentlig ændres. Det er heller ikke sådan at den gruppe af antagelser, der formentlig skal ændres som følge af en handling (erhvervelsen af en ny antagelse), kan afgrænses gennem hvor nye de er, gennem hvor generelle de er eller gennem de semantiske relationer mellem antagelserne og beskrivelsen af handlingen (Fodor, 1983: 113-115).

Den heuristiske løsning

Ifølge Fodor er mange kognitionsforskere af den overbevisning at relevansproblemet kan løses ved hjælp af heuristikker (Fodor, 2008: 117). Heuristikker har ifølge disse forskere potentiale til at udgøre en løsning da det kun er *i princippet* at enhver antagelse i vores antagelsessystem kan have relevans for bekræftelsen af en hypotese. Det er således muligt at vores kognitive system, ud fra egenskaber ved en given hypotese (og ved andre repræsentationer, der er fremtrædende i en given situation), kan tilnærme sig de reelt relevante antagelser (Fodor, 1983: 115, Fodor, 2001: 42).

I *LOT 2* anerkender Fodor at heuristikker i princippet er en mulig løsning, men han har imidlertid ikke stor tiltro til at vi finder passende heuristikker (Fodor, 2008: 125). Problemet er at der ifølge Fodor endnu ikke er nogle heuristikker, der har nærmet sig en løsning, på trods af at flere har forsøgt sig. Alle forslag har være indholdsløse. Indtil videre er de heuristikker, der er blevet foreslået, så vidt Fodor ved, alle en version af ‘gør det, der virkede sidst’ (Fodor, 2008: 117). Fodor mener at sådanne heuristikker er indholdsløse, da det ikke er klart hvordan *det der virkede sidst* afgrænses. Hvis man på et tidspunkt listede forbi en hund der lå og sov i køkkenet, uden at vække den, var det så det at man *listede*, der gjorde den ikke vågnede? Eller var det det at man listede forbi *den hund*? Eller var det det at man listede forbi en hund der lå og sov *i køkkenet*? Og så videre. En ‘gør det, der virkede sidst’-heuristik forudsætter altså en løsning på en form for relevansproblem: Hvilken afgrænsning af det man har erfaret, er den mest relevante eller brugbare afgrænsning for en heuristik? (Fodor, 2008: 119).

Den modulære løsning

I *MOM* argumenterer Fodor for at vores kognitive system er delt op i et centralt system og nogle undersystemer, der leverer repræsentationer til dette centrale system. Undersystemerne kalder Fodor for inputsystemer. Det er systemer, der mere eller mindre af sig selv leverer forskellige informationer om verden omkring os (Fodor, 1983: 39, 52-53). For eksempel foreslår Fodor at kandidater til inputsystemer, blandt flere andre, kunne være den del af vores visuelle system, der beregner genstandes former, og den del af vores auditive system, der beregner grammatiske beskrivelser af sproglige ytringer (Fodor, 1983: 47). En hovedpointe er at disse automatiske, domænespecifikke undersystemer udgør moduler. Det essentielle ved det Fodor forstår som et modul er, at de informationer systemerne konsulterer, er *begrænset* (Fodor, 1983: 71). De har således ikke adgang til hele vores antagelsessystem. Man kunne for eksempel forestille sig at inputsystemet, der beregner genstandes tredimensionelle former, kun har adgang til informationer om bestemte fysiske

karakteristika (som for eksempel de konkave linjer der opstår mellem forskellige farver) og deres sammenhæng med genstandes former. Systemet har altså ikke adgang til informationer som 'fodbolde er runde og for det meste hvide' eller 'Lionel Messi er historiens bedste fodboldspiller'. Det *centrale system* har til gengæld adgang til disse og alle andre antagelser i hele vores antagelsessystem. Det centrale systems opgave er nemlig at integrere alle de informationer inputsystemerne leverer og fikser vores antagelser om det, der er omkring os på baggrund af inputsystemernes informationer og vores baggrundsviden (Fodor, 1983: 39).

Det er imidlertid ikke i princippet givet Fodor har ret i dette. Det er muligt at forestille sig at det centrale system består af en masse moduler, der er designet til at behandle bestemte former for spørgsmål. Med sådan et system af moduler ville det muligvis ikke udgøre et problem for vores antagelsesfikseringsprocesser, at de er isotropiske, for de antagelser eller informationer de forskellige moduler konsulterer er per definition begrænset. Selvom enhver antagelse i vores antagelsessystem i princippet kan sige noget om sandhedsværdien af vores hypoteser, så vil vores modulære system sørge for at det kun er en begrænset del af vores antagelser, der bliver konsulteret. Alle de antagelser modulerne konsulterer, ville med sådan en modulær forståelse af vores centrale system, formodentlig kunne læses 'in real time' og problemet med hvilke antagelser, der er relevante at beregne på, opstår ikke (Fodor, 2001: 55).

Denne mulighed afviser Fodor imidlertid. Det grundlæggende argument han fremfører er, at et modulært system, med domænespecifikke og informationsmæssigt begrænsede moduler, er nødt til på en eller anden måde at fordele de forskellige informationer der kommer fra inputsystemerne ud til de rigtige moduler. Denne proces *kan* ikke være modulær, medmindre fordelingen sker direkte gennem inputsystemerne (der ifølge Fodor er modulære) (Fodor, 2001: 71-74). For hvis det ikke sker gennem inputsystemerne, så er informationerne ikke fordelt når de leveres til det centrale system, og ligegyldigt hvor mange eller hvilke slags moduler, der skulle fordele disse informationer, så ville informationerne ikke være fordelt når de leveres. Hvis fordelingen ikke sker gennem inputsystemerne, må fordelingen altså hvile på en ('indre') systemenhed, der har adgang til *alle* informationerne (Fodor, 2001: 72-73). Hvis fordelingen *sker* gennem inputsystemerne må det vise sig empirisk ved at vi automatisk opfatter noget som hørende til det ene og det andet domæne. Ifølge Fodor er der imidlertid ikke nogen empirisk evidens for at det er tilfældet, med undtagelse af nogle eksperimenter der har vist, at vi automatisk opfatter prikker der bevæger sig på en bestemt måde, som organismer. Fodor pointerer at denne undtagelse næppe kan redegøre for alle de situationer vi opfatter som sociale. Hvordan vil denne undtagelse for eksempel redegøre for at vi opfatter det at snakke i

telefon med en anden person som en social situation? Eller for de utallige andre situationer vi opfatter som sociale, men som ikke lader sig afsløre som sociale gennem bestemte bevægelser, der kan registreres visuelt? (Fodor, 2005: 28). Undtagelsen kan givetvis heller ikke redegøre for alle de andre domænefordelinger, der skulle være. Der er altså nærmest ingen empirisk evidens for at input-systemerne skulle fordele informationer ud til moduler.

Globalitetsargumentet

Globalitetsargumentet er et argument Fodor fremfører i både *MOM*, *TMDWTW* og *LOT 2*. Fodor formulerer imidlertid argumentet på forskellige måder i disse tre bøger og han er heller ikke lige udførlig i de tre bøger. Overordnet fokuserer han på én komponent af argumentet i *MOM*, men uden at være særlig udførlig, mens han fokuserer på en anden komponent i *TMDWTW*, hvor han er forholdsvis udførlig. I *LOT 2* fokuserer han på begge komponenter, men uden at gå i dybden. Jeg vil her gengive argumentationen udførligt ud fra først den ene komponent og så den anden. Jeg kalder den første gengivelse for Glob 1. Den vil tage udgangspunkt i den komponent, der fokuseres på i *MOM*. Jeg kalder den anden gengivelse for Glob 2. Den tager udgangspunkt i den komponent, der fokuseres på i *TMDWTW*. Det er, igen, ikke forskellige argumenter. De fokuserer udelukkende på forskellige komponenter. Grunden til at jeg deler min gengivelse af globalitetsargumentet op i to på denne måde, er for det første at det gør (min forståelse af) argumentet mere klar og tydelig og for det andet at nogle af Fodors formuleringer i de forskellige bøger, umiddelbart kan opfattes som udtryk for at han har forskellige forståelser af de mentale processer, han mener er et problem for CTM. Jeg vil således fremlægge begge formuleringer og tydeliggøre at de formentlig ikke er udtryk for forskellige forståelser.

Glob 1

Globalitetsargumentet hviler på en præmis, der minder om den relevansargumentet hviler på. Relevansargumentet hviler på den præmis at videnskabens bekræftelsesprocesser er isotropiske og at karaktertræk som videnskabens bekræftelsesprocesser har, formentlig også er karaktertræk vores kognitive systems ubevidste bekræftelsesprocesser har. Globalitetsargumentet hviler på et andet karaktertræk, der ifølge Fodor kan tilskrives videnskabens bekræftelsesprocesser. Globalitetsargumentet hviler på den antagelse at videnskabens bekræftelsesprocesser er det Fodor kalder Quineianske (Quineian). At videnskabens bekræftelsesprocesser er Quineianske betyder, at

den grad af bekræftelse en hypotese tildeles, til tider er følsom overfor egenskaber ved *hele* videnskabens system af teorier og hypoteser, som den givne hypotese vurderes sammen med. Det vil sige, bekræftelsen er ikke følsom overfor egenskaber, som hver antagelse i teorisystemet har for sig, men ud fra egenskaber der kan tilskrives hele systemet som sådan (Fodor, 1983: 107-108). Fodor skriver i *MOM*:

“(...) the point about being Quineian is that we might have two astrophysical theories, both of which makes the same predictions about algae and about everything else that we can think of to test, but such that one of the theories is better confirmed than the other - e.g., on grounds of such considerations as simplicity, plausibility, or conservatism. The point is that simplicity, plausibility and conservatism are properties that theories have in virtue of their relation to the whole structure of scientific beliefs *taken collectively*” (Fodor, 1983: 108)

I og med videnskabens bekræftelsesprocesser er Quineianske, er vores kognitive processer det også, som det fremgik af relevansargumentet. Vores kognitive processer er således karakteriseret ved, at graden af bekræftelse en hypotese tildeles, til tider er følsom over for egenskaber, der hører til hele vores antagelsessystem som hypotesen vurderes sammen med (Fodor, 1983: 104, Fodor, 2008: 113). Som det fremgår af citatet, vil dette nærmere sige, at der 1) er nogle bestemte egenskaber som til tider er afgørende for bekræftelsen af vores hypoteser og 2) disse egenskaber er egenskaber, som hypoteserne har *i og med* at de er inkluderet i hele vores antagelsessystem. Han kalder sådanne egenskaber for globale egenskaber. Som det også fremgår af citatet er det egenskaber som simplicitet, plausibilitet og konservatisme Fodor tænker på. I *LOT 2* uddyber han i hvilke generelle situationer vi beregner simplicitet og konservatisme. Han skriver:

“Suppose I have a set of beliefs I’m considering altering in some way under (e.g.) the pressure of experience. Clearly I would prefer the alteration I settle on is the simplest of the available ways to accommodate the recalcitrant data (...). Suppose that my sole belief is that P, but that I am now considering also endorsing either the belief that Q or the belief that R. What I therefore want to evaluate, if I’m to maximize simplicity overall, is whether P&Q is simpler than [*sic*] P&R (...).

Likewise, *mutatis mutandis*, for other parameters of belief systems that one would like to maximize all else equal; for example, the relative conservatism of such commitments.

Nobody wants to change his mind unless he has to; and if one has to, one prefers to opt for the bare minimum of change. The trouble is, once again, that conservatism is a global property of belief systems. On the face of it, you can't estimate how much adding P would alter the set of commitments C by considering P and C severally; on the face of it, conservatism isn't a property of beliefs that they have taken severally" (Fodor, 2008: 121-122).

Problemet med simplicitets og konservatismes globale og holistiske karakter er, at for at beregningen af disse egenskaber kan hvile på syntaksen af vores repræsentationer, er vores kognitive system nødt til at beregne på hele vores antagelsessystem med hver af de to teorier inkluderet i antagelsessystemet. Som det fremgik i relevansargumentet kan det imidlertid umuligt være det vores symbolmanipulerende system gør. Antagelsessystemet er for stort til at et symbolmanipulerende system kan beregne på hele antagelsessystemet 'in real time'. Globalitetsargumentet ender således med at påstå en form for relevansproblem: Hvordan udvælger vi den gruppe af antagelser vi beregner simpliciteten og konservatismen af en teori ud fra? Og hvordan skulle denne udvælgelse hvile på syntaksen af vores repræsentationer? Her hævder Fodor igen at vi ikke har nogen ide. Der er for eksempel igen ingen af de hidtidige heuristiske løsningsforslag, der ikke er indholdsløse eller ikke forudsætter en løsning på problemet (Fodor, 2008: 114, Fodor, 2001: 28, 42, Fodor, 1983: 115-116).

For at gøre argumentet så tydeligt som muligt vil jeg her give et konkret, hypotetisk eksempel på en situation hvor en person kunne finde på at beregne simplicitet. Jeg vil derudfra repetere hvorfor denne beregning ikke kan hvile på syntaksen af vores repræsentationer. Eksemplet hviler blandt andet på Fodors beskrivelser af hvilke situationer vi beregner simplicitet. Det vil sige beskrivelser som:

"(...) the point about being Quineian is that we might have two astrophysical theories, both of which makes the same predictions about algae and about everything else that we can think of to test, but such that one of the theories is better confirmed than the other - e.g., on grounds of such considerations as simplicity, plausibility, or conservatism" (Fodor, 1983: 108).

og:

"Suppose I have a set of beliefs I'm considering altering in some way under (e.g.) the pressure of experience. Clearly I would prefer the alteration I settle on is the simplest of the available ways to accommodate the recalcitrant data" (Fodor, 2008: 121).

I Glob 2 vil det vise sig at disse formuleringer adskiller sig fra hans formuleringer i *TMDWTW*, og dette muligvis kunne lede til en misforståelse.

Eksemplet tager udgangspunkt i en person, Pernille, der har et antagelsessystem, der blandt andet indeholder antagelserne: 'Planeterne, solen og månen bevæger sig konstant og i samme retning rundt om jorden i en cirkel' (Pernille er for eksempel en der ikke er særlig belæst og baserer sin antagelser på sine sporadiske observationer af planeterne, solen og månen). Vi antager at Pernille endnu ikke har observeret noget der skulle modsige hendes antagelser. Fra jorden observerer hun nu, i løbet af en eller anden tidsperiode, at fire af planeterne til tider ændrer deres bevægelsesretning. De bevæger sig i den modsatte retning af den de hidtil har bevæget sig. Dette stemmer givetvis ikke overens med hendes antagelse om at planeterne, solen og månen bevæger sig rundt om jorden i en fast cirkel. Hendes erfaring presser hende altså til at ændre sine antagelser om planeterne, månens og solens bevægelser. Hun spørger derfor sine to bedste venner hvorfor de tror at nogle af planeterne tilsyneladende til tider bevæger sig den modsatte vej. De to venner har straks hver deres teori. Den ene, Alberte, foreslår at planeterne, solen og månen bevæger sig rundt om jorden i cirkler, men at de fire planeter der til tider bevæger sig 'baglæns', tilmed bevæger sig i tilpas mange epicykler til at det går op. Dette er teori A. Den anden ven, Barbara, foreslår at planeterne og solen slet ikke bevæger sig rundt om jorden, men at jorden og planeterne bevæger sig rundt om solen og kun månen bevæger sig rundt om jorden. Dette er teori B. Begge teorier forklarer de lejlighedsvis 'baglæns' bevægelser og giver (umiddelbart) de samme forudsigelser om hvad Pernille mener hun kan observere fra jorden. Hun har altså to teorier til rådighed, der (som tingene er nu) ikke adskiller sig i forklarings- og forudsigelseskraft. Hun ønsker derfor at finde ud af hvilken af teorierne, der er den mest simple af de to teorier.

Ifølge Fodor er simpliciteten af teori A og B afhængig af simpliciteten af hele Pernilles antagelsessystem med henholdsvis teori A og B indlejret. Et symbolmanipulerende system kan således kun beregne sig frem til simpliciteten af de to teorier ved først at beregne simpliciteten af hele antagelsessystemet med teori A indlejret i systemet og derefter hele antagelsessystemet med teori B indlejret i systemet. Det er imidlertid givet at det ikke er det vi gør, da det er en alt for omfattende proces. Beregningen af simplicitet kan altså ikke umiddelbart hvile på syntaksen af repræsentationerne i de teorier hvis simplicitet vi vurderer. I og med at vi ikke beregner simpliciteten af hele vores antagelsessystem, må vi beregne simpliciteten af teori A og B på en anden måde. Men hvordan beregner vi simpliciteten af teori A og B? Måske kan vi godt gengive hvordan vi gjorde det

med teori A og B, men kan vi generalisere denne metode til de andre gange vi beregner simpliciteten af teorier? Kan vi formulere en algoritme? Formentlig ikke, mener Fodor. Der er i hvert fald ikke nogen der har formuleret en endnu (Fodor, 2001: 31, 37-38, 41-46).

Glob 2

Glob 2 hviler hovedsageligt på Fodors formulering af globalitetsargumentet i *TMDWTW* og på bestemte formuleringer i LOT2. Her er der ikke fokus på hvordan beregninger af simplicitet og konservatisme er globale eller holistiske egenskaber og derfor kræver en gennemgang af hele ens antagelsessystem. Fokus er i stedet på simplicitets og konservatismes kontekstsensitive karakter og hvordan den kontekstsensitive karakter gør at beregninger af simplicitet og konservatisme formentlig ikke kan hvile på syntaks. I *TMDWTW* gør Fodor først og fremmest meget ud af gøre det klart hvad syntaks er, at mentale processer ifølge CTM hviler på syntaks og at syntaksen af mentale repræsentationer ikke kan ændre sig. Det sidste betyder med andre ord, at syntaksen er en essentiel del af en repræsentation i CTM og at syntaksen individuerer en mental repræsentation i CTM (Fodor, 2001: 12-13, 20-25, 28). Derfra fremfører han argumentation for at simplicitet og konservatisme er eksempler på egenskaber ved mentale repræsentationer, som repræsentationerne ikke har essentielt eller intrinsisk og dermed er egenskaber, der ikke kan repræsenteres i syntaksen af en repræsentation. Han skriver først om simplicitet:

“Simplicity is, I think, a convincing example of a context-dependent property of mental representations to which cognitive processes are responsive (...). Think of the simplicity of a thought as just whatever determines, for any given theory that you add it to, how much it complicates(/simplifies) that theory. Then simplicity is an intrinsic (i.e., context invariant) property of thoughts if and only if each contributes a constant increment (/decrement) to the overall simplicity of whatever theory you conjoin it to. Pretty clearly, however, the contribution of a thought to determining the simplicity of a theory is not context invariant by this criterion. Rather, what effect adding a new belief has on the overall simplicity of one’s prior epistemic commitments depends on *what one’s prior epistemic commitments are*. Accommodating a planetary regression or two need hardly phase your astronomy if it’s of the geocentric persuasion; but it would complicate our heliocentric astronomy pretty much to extinction.

Likewise for the role of simplicity in practical reasoning. The thought that there will be no wind tomorrow significantly complicates your arrangements if you had intended to sail to Chicago, but not if your plan was to fly, drive, or walk there. But, of course, the syntax of the mental representation that expresses the thought *no wind tomorrow* is the same whichever plan you add it to. The long and short is: The complexity of a thought is not intrinsic, it depends on the context” (Fodor, 2001: 25, 26).

Med dette mener Fodor altså at have vist at simplicitet (/kompleksitet) ikke er en essentiel del af en antagelse (a thought), og simpliciteten af en antagelse dermed ikke kan hvile på syntaksen. Argumentationen er at antagelser som “no wind tomorrow” og “one or two planetary regressions occurred” har én simplicitet når de er forbundet med én gruppe af repræsentationer, der udgør en plan eller teori og en anden simplicitet når de er forbundet med en anden gruppe af repræsentationer, der udgør en plan eller teori (i og med at antagelserne forårsager forskellige grader af ændringer i kompleksiteten hos de forskellige teorier). Det kan altså ikke være syntaksen i “no wind tomorrow” og “one or two planetary regressions occurred” for sig selv, der repræsenterer deres simplicitet.

Fodor mener at egenskaben konservatisme har samme kontekstsensitivitet. Konservatisme refererer i denne sammenhæng, som det fremgik i et citat tidligere, til det at vi foretrækker at ændre så lille en del af vores antagelsessystem som muligt. Det vil sige jo mere konservativ en hypotese er, jo færre ændringer i vores antagelsessystem kræver godtagelsen af hypotesen. Konservatisme er ligesom ved simplicitet, en egenskab vi beregner hvis vi er nødsaget til at ændre i vores antagelsessystem ‘under the pressure of experience’, og de ændringer vi har til rådighed ikke adskiller sig i simplicitet eller den mængde data de kan redegøre for (Fodor, 2001: 33, note 13). Konservatisme handler imidlertid ikke kun om at opgive så få antagelser som muligt, det handler også om at opgive så få *centrale* antagelser som muligt. Centralitet er et parameter for hvor afgørende en antagelse er for en teori. Meget centrale antagelser, er antagelser som hele teorier mere eller mindre hviler på, mens meget lidt centrale antagelser, er antagelser, der følger af teorierne, men som godt kan opgives uden at teorien falder fra hinanden (Fodor, 2001: 34). Det er mere præcist denne egenskab ved antagelser, centraliteten, som Fodor mener ikke er en intrinsisk egenskab, og som derfor ikke kan hvile på syntaks.

Han viser at centraliteten af antagelser ikke er intrinsisk ved at give et eksempel på en antagelse, der har forskellig centralitet afhængig af hvilket teorisystem antagelsen er indlejret i. Antagelsen han bruger som eksempel er noget i retning af ‘alle objekter, der har været i frit fald, har

accelereret i overensstemmelse med deres vægt'. Før i tiden var dette en meget central antagelse, som man byggede sin videnskab på. I dag, med vores masse-centrerede (fremfor vægt-centrerede) mekaniske love, har denne antagelse ikke samme centralitet. Antagelsen følger ikke direkte af den masse-centrerede fysik. Det vil med andre ord sige, før i tiden ville hele videnskabens vægt-centrerede teorisystem falde sammen hvis man opgav antagelsen 'alle objekter, der har været i frit fald, har accelereret i overensstemmelse med deres vægt', mens hvis vi opgav antagelsen i dag, ville det ikke have nogen betydning for videnskabens masse-centrerede teorisystem (Fodor, 2001: 35). Centraliteten af antagelsen 'alle objekter, der har været i frit fald, har accelereret i overensstemmelse med deres vægt' er altså afhængig af hvilket teorisystem antagelsen er inkluderet i. Det er således ikke en intrinsisk egenskab ved antagelsen.

Simplicitet og centralitet er altså ikke intrinsiske eller kontekstinvariante egenskaber ved vores antagelser, og de kan dermed heller ikke hvile på syntaksen af disse antagelser. Ud fra dette konkluderer Fodor at der dermed er visse mentale processer hvor det, der bestemmer den kausale rolle som de mentale repræsentationer (i processerne) har, skifter fra kontekst til kontekst (Fodor, 2001: 28). Det vil sige, der er mentale processer, hvor det der afgør processernes forløb eller 'outputs', ikke er (en af) repræsentationernes intrinsiske, individuerende egenskaber. Det kan således ikke være syntaksen af hver af de mentale repræsentationer for sig, der afgør deres rolle i vores beregninger af disse egenskaber (Fodor, 2001: 28). Det er et problem for CTM, for ifølge CTM hviler beregninger netop på syntaksen af vores repræsentationer.

Så langt adskiller argumentationen i Glob 2 sig tydeligt fra Glob 1. Her i Glob 2 var fokus på at vise, at simplicitet og centralitet er kontekstsensitive egenskaber og de dermed ikke kan hvile på syntaksen af repræsentationerne. I Glob 1 var fokus på at påpege at simplicitet og centralitet er egenskaber som hypoteser og teorier har gennem hele det antagelsessystem de er indlejret i, og at en syntaksdrevet-maskine derfor kun kan beregne simplicitet og centralitet ved at læse hele vores antagelsessystem, hvilket vores kognitive system givetvis ikke gør.

Efter at have argumenteret som fremlagt i Glob 2, vender Fodor imidlertid tilbage til simplicitetens og centralitetens globale karakter i *TMDWTW*. Han pointerer først at det strengt taget ikke er udelukket at simplicitet og centralitet *har* en syntaks. Hvis man forbandt alle antagelser i ens antagelsessystem med hinanden og med den givne antagelse til én repræsentation, så er det muligt at denne samlede repræsentation har en syntaks, der repræsenterer den samlede repræsentations simplicitet eller centraliteten af de forskellige dele af den samlede repræsentation, og dermed den givne antagelses simplicitet eller centralitet (Fodor, 2001: 29, 36). En beregning af denne

repræsentation er imidlertid en beregning af hele ens antagelsessystem. Her er vi altså tilbage ved Glob 1: Et symbolmanipulerende system kan ikke beregne på hele ens antagelsessystem 'in real time'. "Indeed, the totality of one's epistemic commitments is vastly too large a space to have to search whatever it is that one is trying to figure out" som han formulerer det i *TMDWTW* (Fodor, 2001: 31).

Fodor hævder i øvrigt i *TMDWTW* at det er en *kendsgerning*, at vi beregner simpliciteten og centraliteten af teorier gennem ideer der ikke indeholder *hele* teorier. Vi begriber kun dele af teorierne når vi beregner deres globale egenskaber. Vi udvælger på en eller anden måde en gruppe antagelser vi beregner egenskaberne ud fra. Han skriver:

"Whole theories can't be the units of computation (...). It isn't simply that whole theories are generally too big to get one's head around - too big to think about all at once. It's also that assessments of confirmation can be, should be, and generally are called for in respect of objects much less elaborate than the totality of one's cognitive commitments" (Fodor, 2001: 31-32)

Problemet er altså ikke bare at symbolmanipulerende systemer ikke kan beregne på hele antagelsessystemer, men også at beregninger af egenskaber, der hænger sammen med hele vores antagelsessystem, generelt beregnes gennem ideer, der ikke indeholder hele vores antagelsessystem. Fodor giver ikke noget eksempel på hvordan vi beregner på et udvalg af antagelser og han argumenterer heller ikke yderligere for at det er tilfældet. Han mener tilsyneladende, at det er givet.

Ligesom ved Glob 1, ender Glob 2 således også i en form for relevansproblem: Hvordan udvælger vi antagelser at beregne globale egenskaber ud fra? Strengt taget adskiller argumentationen i Glob 2 sig altså ikke fra Glob 1. Ikke udover at fokus er rettet mod forskellige komponenter af argumentationen, og udover at problemet med globale egenskaber bliver uddybet i Glob 2.

Glob 2 adskiller sig imidlertid tilsyneladende fra Glob 1 på en anden front. Fodors beskrivelse af de situationer hvor vi beregner simplicitet er forskellige. I Glob 1 beskriver Fodor det direkte som situationer hvor vi skal vurdere hvilken af to teorier vi skal indoptage i vores antagelsessystem og hvor disse to teorier ikke adskiller sig i forklarings- og forudsigelseskraft. Mit eksempel med Pernille, der skal vurdere om hun skal indoptage teori A eller B, illustrerer sådan en situation. I *TMDWTW* beskriver Fodor imidlertid to anderledes situationer, hvor han mener vi beregner simplicitet. Jeg tænker her på Fodors eksempler med "one or two planetary regressions occurred" og "no wind tomorrow" (Fodor, 2001: 25, 26). Ud fra disse eksempler på beregning af simplicitet mener Fodor

tilsyneladende, at vi beregner simplicitet når vi skal vurdere om accepten af en antagelse vil komplicere én persons antagelsessystem mere end en anden persons (anderledes) antagelsessystem og når vi skal vurdere om accepten af en antagelse, vil komplicere én persons plan mere end en anden persons (anderledes) plan. Det er umiddelbart ikke den samme situation, som den Pernille er i. De forskellige beskrivelser af hvornår vi beregner simplicitet er imidlertid, som jeg ser det, ikke et udtryk for at Fodor mener vi beregner simplicitet i forskellige slags situationer, selvom det måske umiddelbart kan virke sådan. Fodor mener nemlig, som det fremgik i Glob 1, at vi beregner simpliciteten af en teori ved at beregne simpliciteten af det antagelsessystem teorien er en del af (*“What I therefore want to evaluate, if I’m to maximize simplicity overall, is whether P&Q is simpler than [sic] P&R”* (Fodor, 2008: 121)). Det vil sige, når vi er en i en situation, som Pernille var, hvor vi skal vælge mellem to teorier, så beregner vi de to teoriers simplicitet *ved* at beregne den simplicitet inklusionen af den ene teori medfører for antagelsessystemet og den simplicitet inklusionen af den anden teori (for sig) medfører. Lidt karikeret kan Fodors forståelse af beregning af simplicitet beskrives med følgende algoritme:

Hvis du er i tilstand E: ‘Beregn simpliciteten af teori A og B. Sammenlign de to beregninger og vælg den mest simple’

Så skift til tilstand K: ‘Beregn den simplicitet accepten af teori A medfører for hele dit antagelsessystem. Beregn derefter den simplicitet accepten af teori B medfører. Sammenlign og vælg den der medfører det mest simple antagelsessystem’

(Hvis du er tilstand K: ‘har vi ingen ide om hvad du skal gøre. Vores kognitive system holder det skjult’)

Det vil sige, vi beregner simplicitet når vi skal vælge mellem to teorier og det implicerer at vi vurderer i hvilken grad accepten af hver af de to teori for sig komplicerer vores antagelsessystem.

Schneiders indvendinger mod Fodors relevans- og globalitetsargument

I *The Language of Thought: A New Philosophical Direction* forsvarer Schneider CTM overfor Fodors relevans- og globalitetsargument. Hun fremfører som sagt to indvendinger mod relevansargumentet og tre indvendinger mod globalitetsargumentet. Jeg vil her først redegøre for hendes indvendinger mod relevansargumentet og derefter for hendes indvendinger mod globalitetsargumentet. Efter redegørelsen for hver indvending, giver jeg mit svar til hendes indvendinger.

Schneiders indvendinger mod relevansargumentet

I den første indvending Schnider fremfører mod relevansargumentet, hævder hun at kunne vise, at relevansargumentet intuitivt er ugyldigt. I den anden indvending hævder hun at en forholdsvis bredt accepteret teori om vores kognitive system og bevidstheden, Global Workspace Theory (GW), kan udgøre et grundlag for en løsning på det problem relevansargumentet ifølge Schneider trods alt fremlægger. Jeg behandler her først hendes argumentation for at relevansargumentet er ugyldigt og derefter hendes argumentation for at GW kan udgøre et grundlag for en løsning på relevansproblemet.

1. indvending - Et skakprogram

Konklusionen på Fodors relevansargument, er at det er meget sandsynligt at store dele af vores kognitive processer ikke er symbolmanipulerende processer i CTMs forstand. Schneider mener at kunne vise at konklusionen ikke intuitivt følger af disse præmisser. Hun hævder at hvis relevansproblemet opstår i et ukontroversielt symbolmanipulerende system, så indikerer det, at det ikke sandsynligt at store dele af vores kognitive processer ikke er symbolmanipulerende proceser. Hun mener at skakprogrammer udgør eksempler på sådan et symbolmanipulerende system. Ifølge Schneider er det et normalt problem for skakprogrammører at få skakprogrammer til at beregne sig frem til kompetente træk under de tidsbegrænsninger, der normalt er ved skak. Problemet opstår fordi beregninger af de bedst mulige træk består i at beregne et kolossalt antal mulige træk ud i fremtiden (Schneider, 2011: 35). Dette løses imidlertid ved hjælp af begavede heuristikker. Det vil sige i stedet

for at mulige træk og modtræk beregnes, så bruges der heuristikker til at vælge nogle træk ud, som der beregnes på. Ifølge Schneider illustrerer dette en løsning på et relevansproblem. Hun skriver:

“This is precisely the issue of locating algorithms that best allow for the quick selection of a future move from the greatest possible projection of potential future configurations of the board (...). And this is just the relevance problem, as Fodor and others have articulated it” (Schneider, 2011: 35).

Skakprogrammets evne til at foretage kompetente træk i nogenlunde samme tid som mennesker, uden at beregne sig frem til alle mulige træk og modtræk, skal altså vise at der findes computere, der kan løse en form for relevansproblem, og dette skal vise at relevansargumentet ikke viser at det er usandsynligt at alle vores kognitive processer mere eller mindre er symbolmanipulerende som CTM hævder (Schneider, 2011: 36).

Mit svar

Jeg mener at Schneiders argumentation er forfejlet. En betingelse for at Schneiders argumentation er gyldig er, at det er et relevansproblem i Fodors forstand, som skakprogrammet løser. Jeg er enig i at skakprogrammet løser en *form* for relevansproblem. Programmet skal vælge hvilke træk, der bedst kan betale sig at regne på, da det ikke har tid til at beregne på alle træk. Med andre ord skal det altså finde de mest *relevante* træk at beregne på. Spørgsmålet er om det valg skakprogrammet skal foretage, udgør et problem *på samme måde*, som det udgør et problem for vores kognitive system. Det mener jeg ikke det gør. Der er forskel på 1) at beregne sig frem til det bedst mulige træk og 2) at beregne sig frem til bekræftelsen af en hypotese eller beregne sig frem til hvorvidt en antagelse i ens antagelsessystem skal forkastes eller korrigeres. Der er dermed også forskel på hvad der er relevant at beregne på. Der er forskel på ‘hvilke træk er relevante at beregne på?’ og ‘hvilke antagelser er relevante at beregne på?’. Det der gør bestemte *træk* relevante at beregne på, er ikke det, der gør bestemte *antagelser* relevante. Det er således ikke den samme heuristik, der skal bruges. Skakprogrammet har altså ikke løst det relevansproblem som vores kognitive system står overfor. Det har udelukkende løst det som skakprogrammet står overfor. At et skakprogram beregner på de mest relevante træk viser altså, ikke i sig selv, at relevansargumentet intuitivt ikke er gyldigt.

I og med at skakprogrammet løser en form for relevansproblem, er det imidlertid muligt at den løsning skakprogrammet bruger - det vil sige den heuristik den bruger - alligevel kan bruges til at løse relevansproblemet. Måske kan vores kognitive systems relevansproblem løses af en heuristik, der er af samme slags, som den der bruges i skakprogrammet. Schneiders argumentation kunne altså reddes hvis heuristikken eller heuristikkerne hun refererer til, kan bruges til at løse vores kognitive systems relevansproblem. Problemet er bare at Schneider ikke uddyber hvad det er for en heuristik skakprogrammer bruger. Den eller de bestemte heuristikkers *udformning* er ikke en del af hendes argumentation. Hendes argumentation er altså ikke holdbar.

Det kan derudover nævnes at skakprogrammer indeholder andre betingelser for udformningen af heuristikker end vores kognitive system, hvilket gør det usandsynligt at skak-heuristikker skulle kunne være af den samme slags som dem vores kognitive system eventuelt bruger. De heuristikker kognitionsforskere ifølge Fodor har foreslået indtil nu, er alle en form for 'gør det, der virkede sidst'-heuristikker. Som Fodor argumenterede, forudsætter disse heuristikker imidlertid en løsning på et relevansproblem, da der er uendelig mange elementer af det vi oplevede sidst, som kan have relevans for den situation vi er i nu. Udformningen af heuristikker, der i grove træk har formen 'gør det, der virkede sidst' forudsætter imidlertid ikke løsningen på et relevansproblem, hvis de er heuristikker til et skakprogram. I et skakprogram er der et *begrænset* antal elementer, der kan have relevans for om en situation er en, der ligner en situation som virkede sidst. Enten er en bestemt prik på et felt, som den samme brik var på i en anden situation, hvor et bestemt træk virkede, eller også er den bestemte brik på et andet felt. Skakprogrammet skal altså 'bare' afgøre hvilke skakbrikker der var hvor da et træk 'virkede' og huske dette til næste træk. Skakprogrammet - eller skakprogrammørerne - skal ikke afgøre om det var det at skakbrikken var i den ene side af feltet eller den anden side af feltet, der gjorde at et træk virkede. Skak-programmer har altså netop ikke det relevansproblem som vores kognitive system har.

2. indvending - Global Workspace Theory

Udover at hævde at skakprogrammer viser at Fodors relevansargumentet er ugyldigt, hævder Schneider også at der findes en forholdsvis ny, men veletableret teori om vores kognitive system og bevidstheden, der udgør et grundlag for en løsning på relevansproblemet. En teori der tilmed vinder accept blandt flere og flere indenfor kognitionsforskningen (Schneider, 2011: 43). Teorien hun refererer til er Global Workspace Theory. Den grundlæggende ide i GW er at informationsflowet i

vores kognitive system består af en kombination af serielle og parallelle processer. Der er for det første en masse specialiserede processer-enheder, som har ansvaret for kognitive processer som (aspekter af) perception, langtidspaner, problemløsning, sprogforståelse og så videre. Disse specialiserede enheder konkurrerer parallelt om at få adgang til det, der kaldes et globalt arbejdsrum, hvor information bearbejdes serielt (Shanahan & Baars, 2005: 165). På et neurologisk plan består det globale arbejdsrum af globale hjerneprocesser (pan-cortical), mens de specialiserede, parallelle processer foregår i lokalt afgrænsede områder i hjernen (Schneider, 2011: 43-44, note 9). Når en information fra de parallelle, specialiserede processer vinder adgang til det globale arbejdsrum, sendes denne information således til hele systemet. Det vil sige, den fordeles ud til alle de forskellige specialiserede processer. Processerne i det globale arbejdsrum, er desuden bevidste processer, mens de specialiserede, parallelle processer er ubevidste processer. Det der afgør om en information vinder adgang til det globale arbejdsrum kan beskrives som en form for opmærksomhedsmekanismer (Shanahan & Baars, 2005: 165, Schneider, 2011: 43-44). Det er altså vores opmærksomhed, der på en eller anden måde afgør hvilke af de ubevidste processers repræsentationer, der når vores bevidsthed. Schneider beskriver ikke mere præcist hvordan udvælgelsen foregår. Dette er imidlertid heller ikke vigtigt for hendes argumentation for at GW kan udgøre et grundlag for en løsning på relevansproblemet.

I sin argumentation for at GW kan udgøre et grundlag for en løsning på relevansproblemet, skriver hun indledningsvis:

“To appreciate the GW approach to the relevance problem, let us first consider the predicament that the standard LOT faces. Suppose that a classical system needs to make a relevance determination. The standard concern is that it must engage in serial computations that examine each member of a long list of alternatives, one by one (...). Assuming that the database is large, and knowing what we do about the speed of neural processing, it is difficult to grasp how a relevance determination can be made in real time” (Schneider, 2011: 44).

Ifølge Schneider består relevansargumentet altså i at hævde, at et symbolmanipulerende system ikke kan beregne relevansen af hver antagelse i et helt antagelsessystem ‘in real time’. Hendes argument for at GW kan udgøre en løsning på dette problem, består således i at vise at GW’s model over vores kognitive system, giver et grundlag for at det *er* muligt for et symbolmanipulerende system at beregne relevansen ‘in real time’. Dette gør hun først ved at påpege at grunden til at Fodor mener et

symbolmanipulerende system ikke kan beregne på et helt antagelsessystem 'in real time', er at han forstår symbolmanipulerende processer, som processer der kun kan foregå serielt. Ifølge Schneider er dette en antikvarisk forståelse. CTM kan sagtens harmonerer med parallelle processer. Hun argumenterer ikke uddybende for dette, men påpeger bare at der ikke er nogen forskere, som har den symbolmanipulerende forståelse af vores kognitive system, der mener at vores kognitive processer foregår serielt (Schneider, 2011: 45). I og med at CTM kan harmonerer med et system fyldt med parallelle processer, er der ifølge Schneider håb for CTM. Mange parallelle processer medfører givetvis en kraftig tidsbesparelse for et symbolmanipulerende system. Det er således ikke så åbenlyst at et symbolmanipulerende system ikke kan beregne på hele vores antagelsessystem 'in real time'. Spørgsmålet er bare om der er en plausibel teori om vores kognitive system, der harmonerer med et systembolmanipulerende system og som indeholder en masse parallelle processer. GW, der ifølge Schneider fortsat møder større og større accept blandt kognitionsforskere, er netop sådan en teori. GW er altså en teori, der, med sine mange parallelle processer, kan udgøre et grundlag for relevansproblemet. Ifølge Schneider er der således ikke grund til at være så pessimistisk omkring CTM (Schneider, 2011: 46).

Mit svar

Jeg mener at Schneiders argumentation for at GW kan udgøre et grundlag for en løsning på relevansproblemet er forfejlet. Hendes argumentation baserer sig for det første på en misforståelse af relevansargumentet. Hendes opfattelse af problemet er at et symbolmanipulerende system ikke 'in real time' kan beregne sig frem til hvilke antagelser i ens antagelsessystem, der er relevante. Det er ikke det, der er problemet. Problemet er at vi ikke har nogen ide om *hvordan* relevansen overhovedet beregnes. Det store spørgsmål er: Hvad er det der afgør, om en antagelse er relevant for en bekræftelsesproces eller en revurderingsproces og således bliver udvalgt til at blive beregnet på, når vi fikserer vores antagelser? At GW udgør en empirisk underbygget teori, der viser at symbolmanipulerende processer ikke behøver være serielle, er altså ikke noget fremskridt i forhold til relevansproblemet.

Man kunne indvende at relevansproblemet (som jeg forstår det), imidlertid hviler på en anden form for antagelse om at et symbolmanipulerende system ikke kan beregne på hele vores antagelsessystem 'in real time'. Det der ikke kan beregnes in real time' er ikke relevansen, men i hvor høj grad hver antagelse (af/)bekræfter en given hypotese eller hvorvidt hver antagelse skal korrigeres

givet etableringen af en ny antagelse. Det er altså muligt at GW rent faktisk udgør et grundlag for en løsning på relevansproblemet. Hvis Schneider ville hævde dette og hun havde ret, ville der ikke være noget grundlag for et *relevans*problem, for vi ville således ikke udvælge en gruppe af antagelser at beregne på. Vores kognitive system ville i stedet beregne på (stort set) alle repræsentationer i vores antagelsessystem.

Hvorvidt denne reformulering af Schneiders argumentation er holdbar, er imidlertid for det første afhængig af hvorvidt vores antagelsessystem er *uendeligt stort*. Fodor hævder som sagt i *LOT 2* at vi har et uendeligt stort korpus af antagelser eller 'cognitive commitments'. Hvis det virkelig er uendeligt stort, så hjælper det ikke på relevansproblemet at et symbolmanipulerende system kan beregne på en meget stor, men afgrænset gruppe af antagelser. Med et uendeligt stort antagelsessystem, er vi *nødsaget* til at vælge en relevant gruppe ud, da beregningen af hele vores antagelsessystem ville være en uendelig beregning. GW udgør altså ikke et grundlag for en løsning på relevansproblemet hvis vores antagelsessystem virkelig er uendeligt stort.

Fodor giver imidlertid ikke nogen argumentation for at vores antagelsessystem skulle være uendeligt stort. Han mener tilsyneladende at det er åbenlyst. Jeg er tilbøjelig til at give ham ret, men jeg mener ikke det er åbenlyst. Det mener jeg ikke det er, hovedsageligt fordi det ikke er åbenlyst hvad der præcis ligger i 'vores antagelsessystem' eller 'our corpus of epistemic/cognitive commitments', som Fodor også formulerer det. Det er ikke noget Fodor uddyber. Jeg mener imidlertid at kunne vise at beregninger af hvorvidt *enhver* antagelse i antagelsessystem afkræfter eller underbygger en given hypotese, ender ud i et relevansproblem. Det er således ligegyldig hvilken kapacitet et symbolmanipulerende system har. Et relevansproblem opstår.

Lad os sige at vores system er et symbolmanipulerende system med helt enorm kapacitet og lad os forestille os at en person, Pernille, står på et hustag i en by og overvejer om hun er i stand til at springe over på et andet hustag, der er omkring 2 meter fra det tag hun står på. Hun undersøger derpå hypotesen 'jeg kan godt springe over til det andet tag'. Hendes kognitive system gennemgår hendes antagelsessystem. Det vil sige det beregner på hver antagelse én efter én i parallelle processer. Lad os sige en af de parallelle processer er nået til antagelsen 'jeg har engang set en person springe fra et hustag til et andet hustag'. Hendes system skal nu beregne om den antagelse be- eller afkræfter hendes hypotese, og i hvor høj grad den gør det. Lad os sige at det hurtigt beregner sig frem til at den ikke modsiger systemets hypotese. Spørgsmålet er så om den *underbygger* en bekræftelse af systemets hypotese. Hvordan finder systemet ud af det? Jeg kan ikke komme på nogen bedre metode, end at det undersøger hvor meget antagelsen ligner beskrivelsen af ens situation. Hvordan skulle antagelsen

ellers underbygge Pernilles hypotese? Jo mere situationen i antagelsen ligner Pernilles nuværende situation, jo mere underbygger antagelsen altså hendes hypotese. I så fald skal systemet altså beregne hvor meget den 'genkaldte' situation ligner Pernilles nuværende situation og hvor meget personen i den 'genkaldte' situation ligner Pernilles for at finde ud af hvor meget antagelsen underbygger hendes hypotese. Problemet er imidlertid at der er uendeligt mange elementer af situationerne og af en selv og personen, der kan ligne hinanden. Tagenes afstande mellem hinanden, tagenes farver, materialerne tagene er bygget af, vejret, tøjet, kropsstørrelserne, kropsdelestørrelserne, alderen og så videre³. Der opstår således et relevansproblem. Hvilke elementer er relevante for sammenligningen? GWs parallelle processer hjælper ikke på dette problem.

Jeg mener altså ikke GW udgør et grundlag for en løsning på relevansproblemet. Teorien nærmer sig ikke nogen løsning på hvordan vi *udvælger* antagelser eller elementer af antagelser i vores beregninger.

Jeg vil derudover påpege at Schneider ikke har ret i at Fodor betragter symbolmanipulerende systemer, som nogen der ikke kan foretage parallelle beregninger. I "Connectionism and cognitive architecture: A critical analysis" (1988) skriver han klart og tydeligt at den symbolmanipulerende forståelse af vores kognitive system er neutral omkring hvorvidt der kan foregå parallelle beregningsprocesser (Fodor & Pylyshyn, 1988: 15). Det er altså ikke fordi Fodor betragter symbolmanipulerende systemer, som systemer der kun kan foretage serielle beregninger, at han mener at vores antagelsessystem er for stort til at et symbolmanipulerende system kan beregne på hele dette antagelsessystem 'in real time'.

Opsamling

Jeg mener således ikke Schneiders indvendinger mod relevansargumentet er holdbare. Hendes eksempel på et symbolmanipulerende system, der ifølge hende løser et relevansproblem er *ikke* et eksempel på et system, der løser et relevansproblem i Fodors forstand. Skakprogrammet løser en *form* for relevansproblem, men ikke det relevansproblem vores kognitive system står overfor. At skakprogrammet 'løser' en form for relevansproblem, er ikke tilstrækkeligt til at vise at relevansargumentet er ugyldigt.

³ Denne uendelighed af antagelser i vores antagelser er måske det der gør at Fodor mener vores antagelsessystem er uendeligt stort

Schneiders argumentation for at GW udgør et grundlag for en løsning på relevansproblemet, hviler på en misforståelse af relevansargumentet. Det at GW indeholder en model af vores kognitive system, der har stor beregningskapacitet, indebærer ikke at GW kan udgøre et grundlag for en løsning på relevansproblemet. Et system med stor kapacitet hjælper ikke på at finde ud af hvilke antagelser der er relevante, hvilket *er* relevansproblemet. Uden en regel for relevans vil beregninger af hvorvidt enhver antagelse i ens antagelsessystem afkræfter eller underbygger en hypotese eller kræver korrigerende givet en ny antagelse, være en uendelig proces.

Schneiders indvendinger mod globalitetsargumentet

Schneider giver tre forskellige indvendinger mod globalitetsargumentet. Jeg vil i de næste afsnit redegøre for de tre indvendinger og give mit svar til hver indvending efter hver redegørelse. For at forstå Schneiders indvendinger, er man imidlertid nødt til forstå hvad hun mener globalitetsargumentet består i. Hun har nemlig ikke den helt samme forståelse af argumentet, som den jeg har fremlagt. Hun gengiver globalitetsargumenget forud for hver indvending hun fremfører, og disse gengivelser adskiller sig på forskellige måder. I den første gengivelse redegør hun kun kort for det grundlæggende problem for CTM, som hun mener globalitetsargumentet hævder. I den næste gengivelse citerer hun en passage fra *LOT 2*, hvor Fodor formulerer argumentet. Hun mener argumentet i *LOT 2* er anderledes end i *TMDWTW*, og derfor kræver et andet svar (Schneider, 2011: 29). I hendes sidste gengivelse uddyber hun sin forståelse af argumentet, som hun fremlagde den i den første gengivelse. Hun ekspliciterer her det hun forstår som præmisserne i argumentet. Hendes tre indvendinger mod globalitetsargumentet hænger sammen med disse tre forskellige gengivelser. Jeg vil derfor kort redegøre for eller citere hendes forskellige gengivelser af argumentet forud for hver redegørelse for hendes indvendinger.

1. indvending - Et skakprogram

Forud for hendes første indvending mod globalitetsargumentet opsummerer hun globalitetsargumentet på følgende måde:

“(a) *The globality argument*. Global properties are properties of a LOT sentence that are supposed to be determined by the nature of the larger group of sentences that the sentence is

entertained with. Consider, for instance, the LOT sentence the *espresso machine broke*. On the one hand, if you consider it in tandem with sentences about your plan to vacation in London next year, it will likely not alter the plan. On the other hand, if the sentence is entertained together with mental sentences that concern a plan to make espresso tomorrow morning, it will ruin the plan. The role the sentence plays in computation is *context sensitive* then, depending upon what other mental sentences are under consideration. And the problem with LOT and CTM is that cognition, being computational, is sensitive only to the “syntax” of the particular mental representation. Syntactic properties are *context insensitive* properties of a mental representation. That is, what a mental representation’s syntactic properties are does not depend on what the other mental representations in the plan happen to be. But whether a given mental representation has the global properties it has will depend on the nature of the other sentences in the relevant group. If this is correct, the global properties do not supervene on syntactic properties. Hence, LOT and CTM are untenable as a general account of how the cognitive mind works (...)” (Schneider, 2011: 31-32).

Hun bemærker derudover senere at globalitetsargumentet er et *in principle*-argument. Det vil sige, hun mener at argumentet skal vise at det *principielt* er umuligt at vores kognitive system er et symbolmanipulerende system, som CTM hævder. Hvis argumentet har ret, kan det således ikke lade sig gøre at empirisk forskning på et tidspunkt vil vise, at vores kognitive system rent faktisk er et symbolmanipulerende system (Schneider, 2011: 40-41).

Indvendingen hun fremfører mod globalitetsargumentet efter denne gengivelse af argumentet, ligner den ene indvending hun fremfører mod relevansargumentet. Ligesom med relevansproblemet mener Schneider, at hvis globalitetsproblemet opstår i ukontroversielle symbolmanipulerende systemer, er Fodors globalitetsargument ikke gyldigt. Hendes (første) indvending mod globalitetsargumentet består således i at give et eksempel på et globalitetsproblem der opstår i et ukontroversielt symbolmanipulerende system (Schneider, 2011: 34). Det ukontroversielle symbolmanipulerende system hun bruger som eksempel er igen et skakprogram. Der findes ifølge Schneider skakprogrammer, der har to strategier til rådighed, når det skal vælge et træk. Den strategi programmet vælger, er styret af det træk modstanderen har foretaget. Mere præcist er valget af strategi afhængigt af hvilken strategi programmet vurderer er mest simpelt givet modstanderens træk. Ifølge Schneider er det, at programmet skal vælge den mest simple strategi ud fra modstanderens

forudgående træk, et globalitetsproblem. Det mener hun fordi simpliciteten af strategierne er afhængig af det træk, der foretages af modstanderen. Hun skriver:

”(...) the impact that the addition of the information about what the opponent’s first move was on the simplicity of each of the two plans fails to supervene on the type identity of the string of symbols encoding the information about the opponent’s first move. Instead, the impact of the addition of the string of symbols to the simplicity of each plan depends on the way that the string interacts with the other sentences. (...) It thereby appears that a globality problem emerges (...)“ (Schneider, 2011: 35)

I og med skakprogrammet ikke har nogen problemer med at beregne simpliciteten af de to strategier, viser skakprogrammet at symbolmanipulerende systemer *kan* beregne globale egenskaber. Hun konkluderer således, på baggrund af dette skakprogram-eksempel, at Fodors globalitetsargument ikke er gyldigt (Schneider, 2011: 35).

Mit svar

Jeg mener at der er flere ting galt med denne indvending. For det første har hun efter min mening misforstået argumentet, hvilket fører til at hun fejlagtigt kategoriserer argumentet som et *in principle*-argument. Argumentet er *ikke* et *in principle*-argument. Fodor hævder både at globale egenskaber strengt taget *kan* have en syntaks og at beregninger af globale egenskaber som simplicitet og centralitet *i princippet* kan foretages af heuristikker, som kan hvile på syntaks. Det at vise, at et symbolmanipulerende system ”løser” et globalitetsproblem, viser således ikke *i sig selv*, at Fodors argumentation er ugyldig. Hun er nødt til at vise at systemet ikke beregner sig frem til graden af den givne globale egenskab (graden af simplicitet), ved at beregne på alle de ‘strings of symbols’ der er afgørende for hvilken grad en given strategi har den globale egenskab. Schneider uddyber imidlertid ikke hvordan skakprogrammer vælger den mest simple strategi. Hendes argumentation viser altså ikke (i sig selv) at Fodors globalitetsargument er ugyldigt.

For det andet, er en betingelse for at Schneiders argumentation er gyldig, at det globalitetsproblem, der ‘opstår’ i skakprogrammet, er et globalitetsproblem i Fodors forstand. Schneider fremlægger globalitetsproblemet som om det består i, at nogle mentale sætninger har bestemte egenskaber gennem den gruppe af sætninger de er ‘entertained with’. Så vidt Fodor mener

at man ikke kan adskille teorier fra det antagelsessystem de er indlejret i, er dette ikke globalitetsproblemet. Globalitetsproblemet er at nogle mentale sætninger har deres egenskaber 'in virtue of their relation to the *whole structure* of scientific beliefs *taken collectively*' eller gennem *hele vores antagelsessystem*. For at et globalitetsproblem kan siges at opstå hos et symbolmanipulerende system, må det symbolmanipulerende system altså skulle beregne en egenskab, som kun et enormt antal 'strings of symbols' repræsenterer. Hvis det symbolmanipulerende system beregner denne egenskab ved at 'læse' alle 'strings of symbols', så er det desuden, som sagt, ikke et symbolmanipulerende system, som vores kognitive system kan udgøres af. For at systemet kan siges at møde et egentligt globalitetsproblem, må det altså beregne disse globale egenskaber ved hjælp af heuristikker. Derudover skal det også være heuristikker, som vores kognitive system kan bruge. Schneider uddyber imidlertid ikke hvordan skakprogrammer vælger deres strategier, så det fremgår ikke om skakprogrammer anvender heuristikker, som vores kognitive system kan anvende. Derudover, som jeg argumenterede i mit svar til hendes første indvending mod relevansargumentet (der også hvilede på et skakprogram), indeholder skakprogrammer andre betingelser for udformningen af heuristikker end vores kognitive system. Det er ikke et problem for udformningen af heuristikker til skakprogrammer at skulle afgrænse hvilke elementer, der afgør om en heuristik kan anvendes. Dette gør det altså mindre sandsynligt at de heuristikker et skakprogram bruger, er heuristikker vores kognitive system kan bruge.

Schneiders argumentation er altså ikke gyldig. At der findes et skakprogram som vælger mellem to planer ud fra hvor simple de er, givet et bestemt træk fra modstanderen, fører ikke til at Fodors globalitetsargument er ugyldigt.

Jeg vil imidlertid til Schneiders forsvar bemærke at Fodor ikke klart og tydeligt afgrænser hvad han forstår ved simplicitet og at han med sine eksempler i *TMDWTW* udtrykker en meget tvetydig forståelse af simplicitet og dets globale karakter. Jeg tænker her især på eksemplet med 'no wind tomorrow'. Jeg vil vende tilbage til dette med uddybende bemærkninger i mit svar til Schneiders tredje indvending.

2. indvending - CTMs ressourcer

Schneiders andet argument mod globalitetsargumentet tager udgangspunkt i et bestemt uddrag af Fodors argumentation i *LOT 2*, som hun citerer. Hun mener at denne passage udgør et andet slags

globalitetsargument end det Fodor formulerer i *TMDWTW*. Schneider citerer Fodor på følgende måde:

“Fodor writes,

The Globality Problem is that I can't evaluate the overall simplicity of a belief system by summing the intrinsic simplicities of each of the beliefs that belongs to it. In fact, there is no such thing as the intrinsic simplicity of a belief. ... Nothing local about representations - in particular, nothing about the formal relations between the representations and its constituent parts - determines how much I would complicate or (simplify) my current cognitive commitments if I were to endorse it.

Notice that unlike the problem of relevance, this sort of worry about locality holds even for the very small systems of belief. ... Suppose that my sole belief is that P, but that I am now considering also endorsing either the belief that Q or the belief that R. What I therefore want to evaluate, if I'm to maximize simplicity overall, is whether P&Q is simpler than P&R. *But I can't do that by considering P, Q and R severally; the complexity of P&Q isn't a function of the simplicity of P and the simplicity of Q.* So the operations whereby I compute the simplicity of P&Q can't be local. (Fodor 2008, 122; italics mine)” (Schneider, 2011: 38).

Schneider mener at Fodors argumentation i denne passage er forfejlet, da der, ifølge Schneider, er en underliggende præmis i Fodors argumentation, der ikke er sand. Den underliggende præmis er, at Fodor antager at for at CTM er rigtig, må simpliciteten af P&Q relativt til P&R beregnes udelukkende gennem en læsning af P, Q og R (og operatoren '&') hver for sig (Schneider, 2011: 39). Ifølge Schneider undervurderer denne præmis kraftigt de ressourcer CTM har til rådighed. For at underbygge hendes påstand giver hun et eksempel, der skal vise hvordan beregningen af P&Q ikke behøver at hvile udelukkende på beregninger af P og Q separat. Hun beder læseren om at forestille sig en person, der har flere antagelser end P og gerne vil vurdere simpliciteten af en plan der består af P&Q. Der er ifølge Schneider ikke noget, der taler imod at vurderingen af simpliciteten af denne plan skulle inkludere andre repræsentationer fra ens 'database' (antagelsesystem). Schneider hævder at det ligefrem er plausibelt at antage at andre repræsentationer inkluderes i beregningen af simpliciteten. For eksempel kan de forskellige midler, der kan bruges til at udføre planen indgå i ens

beregninger. Simpliciteten af P&Q kan således beregnes ved at beregne på P, Q og andre antagelser som M (muligt middel til at udføre planen) og D (et andet muligt middel til at udføre planen) og så videre separat. Præmissen om at CTM kun kan være rigtig hvis P&Q beregnes ved udelukkende at læse P og Q separat, er altså forkert. Schneider mener således at have vist at globalitetsargumentet, som det fremgår i *LOT 2*, er forfejlet.

Mit svar

Jeg mener denne argumentation baserer sig på en misforståelse af det Fodor hævder i den citerede passage. Det er korrekt at Fodor mener at CTM er en teori, der implicerer at P&Q's egenskaber kun kan beregnes ved at beregne på P og Q separat, men kun i et hypotetisk eksempel *hvor personen ikke har andre antagelser end P*. Han mener ikke at simpliciteten af en given antagelse (eller plan) *i det hele taget* ikke kan hvile på andre repræsentationer end dem der (umiddelbart) beskriver antagelsen eller planen. Pointen i passagen er ikke at CTM kun kan være korrekt, hvis vi beregner simplicitet ud fra en *meget* lille gruppe af antagelser (hvilket tilsyneladende er sådan Schneider forstår den). Pointen er at simplicitet ikke kan beregnes ved at beregne på de enkelte repræsentationer, der udgør ens antagelsessystem og en inkluderet antagelse, *separat*. Simplicitet er en holistisk egenskab. Simpliciteten af P&Q eller P&Q&M&D er for eksempel ikke den samme, som summen af simpliciteten af P og Q eller P, Q, M og D. Schneider misforstår altså argumentet og hendes eksempel angriber en præmis, der ikke er i Fodors globalitetsargument. Hendes argumentation viser således ikke at Fodors globalitetsargument er forfejlet.

Schneider mener i øvrigt selv at Fodor formentlig ville indvende mod hendes argumentation, at globalitetsproblemet i vurderingen af simpliciteten af en plan, er at finde ud af hvilke af ens antagelser, der kan have betydning for simpliciteten af ens plan. Han ville altså være enig med Schneider i, at beregningen af simpliciteten af en plan kan inkludere andre repræsentationer end dem der (umiddelbart) beskriver planen. Problemet ifølge Fodor er imidlertid at det i princippet kan være alle ens antagelser, der kan have betydning for simpliciteten af en plan og vi ikke har nogen ide om hvordan vi vælger de antagelser, der er relevante for beregningen af simplicitet. Til dette svarer Schneider at det så ikke længere er et globalitetsproblem, men et relevansproblem, som hun har argumenteret for ikke er et problem for CTM (Schneider, 2011: 40)

Jeg er enig med Schneider i at Fodor formentlig ville indvende sådan⁴. Jeg mener imidlertid ikke, at han ville modsige hende at globalitetsargumentet også er en form for relevansargument. Globalitetsargumentet *er* en form for relevansargument. Schneider mener selvfølgelig at hun har afmonteret relevansargumentet og at hun således ikke behøver at argumentere for at relevansproblemet, der opstår i globalitetsargumentet, ikke er et problem. Jeg mener imidlertid ikke hun har afmonteret det, som jeg argumenterede for før. At hun her reducerer globalitetsproblemet til et relevansproblem hjælper altså ikke hendes argumentation.

3. indvending - En problematisk definition og usandsynligheden for globale egenskabers eksistens

Schneiders tredje argument mod globalitetsargumentet tager udgangspunkt i en grundigere fremstilling af Fodors argumentation i *TMDWTW*. Hun tager først og fremmest udgangspunkt i Fodors eksempel på hvordan simplicitet er en kontekstsensitiv egenskab. Hun fremlægger argumentets præmisser og konklusion således:

- “1. A belief can contribute to the complexity/simplicity of a theory. [fact]
2. The complexity/simplicity of a theory is relevant to cognition. [fact]
3. What a belief contributes to the complexity/simplicity of a theory is relevant to cognition. [1, 2]
4. Def: “The simplicity of a [belief is] whatever determines for any given theory that you add it to, how much it complicates(/simplifies) that theory” (p. 26).
5. The simplicity of a belief is relevant to cognition. [3, 4]
6. If the computational theory of mind is true, then the simplicity of a belief must be a context-invariant feature of it.
7. If the simplicity of a belief is a context-invariant feature of it, then it can only contribute to the complexity of a theory by contributing a “constant increment (/decrement) to the overall simplicity of” the theory (p. 26)

⁴ Det er imidlertid uklart hvordan Fodor mener simplicitetsproblemer opstår i forhold til planer og om han mener at simpliciteten af planer er den samme slags simplicitet som simpliciteten af antagesler. Den eneste gang han nævner og beskriver simplicitetsproblemer for planer er, da han beskriver eksemplet med ‘no wind tomorrow’ i *TMDWTW*, som jeg har citeret i dets fulde længde i Glob 2.

8. If adding a belief can contribute to the complexity of a theory only by contributing a constant increment or decrement, then adding a belief to one set of attitudes cannot make a minimal difference to the complexity of an agent's theory, while adding the same belief (a belief of the same type) to another set of attitudes makes a significant difference to the complexity of the agent's theory.

9. However, adding a belief to one set of attitudes can make a minimal difference to the complexity of an agent's theory, while adding the same belief to another set of attitudes makes a significant difference to the complexity of an agent's theory.

10. Therefore, CTM is false [6-9]" (Schneider, 2011: 70-71).

Som sagt opfatter hun globalitetsargumentet som et *in principle*-argument. Hun mener derfor ikke, at hun behøver vise andet end at Fodor tager fejl, når han argumenterer for at simplicitet ikke kan hvile på syntaks. Dette viser hun på to måder. Først og fremmest angriber hun præmis 7. Hun mener ikke at det er korrekt, at hvis simplicitet er en kontekstinvariant egenskab ved en antagelse, så vil antagelsen nødvendigvis bidrage med den samme forøgelse eller formindskelse af simplicitet til en teori (til et antagelsessystem), som antagelsen tilføjes (Schneider, 2011: 75). Når man tager Fodors definition af simplicitet i betragtning, kan det umiddelbart virke åbenlyst at Schneider ikke har ret i dette. Ifølge Fodors definition af en antagelses simplicitet, er simpliciteten af en antagelse tilsyneladende afhængig af den simplicitet en teori erhverver sig, når antagelsen tilføjes. Hvis kompleksiteten af en teori nogle gange forøges markant ved tilføjelsen af en antagelse og andre gange nærmest ikke ændrer sig, synes det åbenlyst at simpliciteten af en antagelse ikke kan være kontekstinvariant (Schneider, 2011: 76-77). Schneider påpeger imidlertid at Fodors definition siger, at den egenskab ved en antagelse der kaldes simplicitet, er 'whatever *determines*' at en given teori kompliceres eller simplificeres. Lægger man præmis 4 og 7 sammen hævder Fodor altså følgende:

"7.* If the property of a belief that determines any given theory that you add it to how much it complicates (/simplifies) that theory is a context-invariant feature of it, then it can only contribute to the complexity of a theory by contributing a constant increment or decrement" (Schneider, 2011: 76).

Ifølge Schneider er det klart, at dette ikke er korrekt. Dét ved en antagelse, eller den *bestanddel* af en antagelse, som bestemmer hvor meget en teori kompliceres, kan sagtens være kontekstinvariant

samtidig med at denne bestanddel forårsager forskellige grader af kompleksitet hos forskellige teorier. Som hun skriver:

“(...) clearly there can be intrinsic features of things that when added to some systems have a small effect on certain global properties, but that when added to others have a significant effect. Adding a small amount of mass, say, a gram, to the Moon will not have much effect on its size. But adding the same amount of mass to a star that is on the brink of collapsing in on itself due to its mass may have an enormous effect on its size.” (Schneider, 2011: 76).

Præmis 7 er altså ikke korrekt, og således kan simplicitet godt være en intrinsisk egenskab på trods af at sætninger som ‘no wind tomorrow’ medfører at en plan kompliceres og en anden ikke gør. En antagelses simplicitet kan altså godt være repræsenteret af antagelsens syntaks. Dermed har Fodor ikke vist at simplicitet ikke kan hvile på syntaksen af repræsentationer, og globalitetsargumentet med udgangspunkt i simplicitet er således ugyldigt (Schneider, 2011: 75-77).

Schneider bemærker derudover at Fodors definition af simplicitet, medfører at simplicitet er en intrinsisk egenskab. Hun giver selv denne formelle beskrivelse af definitionen:

“The simplicity of a belief B = the property P of B such that for any theory T, and any change in complexity R, if R is the change in the complexity of T on adding B to it, then that B has P determines that the change in complexity of T on adding B to it is R” (Schneider, 2011: 77-78).

Ud fra denne formalisering af definitionen mener Schneider at det er åbenlyst, at simplicitet ikke er en egenskab ved B, som skifter afhængigt af hvilken teori B tilføjes til. Simpliciteten er defineret som det, der *bestemmer* en teoris kompleksitet. Simpliciteten kan altså ikke være *afhængig* af teoriens kompleksitet. Fodors definition af simplicitet medfører derfor, sammen med præmis 9, at 7 og 7* er falske præmisser. Schneider ender således med at konkludere at globalitetsargumentet med udgangspunkt i simplicitet tilmed er “*self-defeating*” (Schneider, 2011: 78).

Schneider pointerer imidlertid at det er muligt at Fodor muligvis har formuleret sig forkert ved en fejl, da han formulerede definitionen af simplicitet. Hun omformulerer derfor hans argumentation uden definitionen (Schneider, 2011: 78-79). Ud fra den del i *TMDWTW*, der leder op

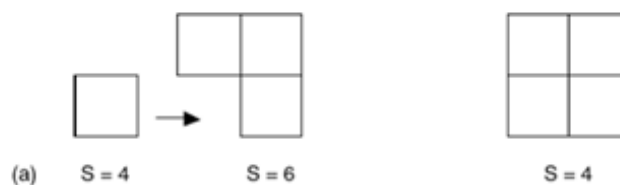
til simplicitetsargumentet, omformulerer hun argumentet så det består af følgende præmisser og følgende konklusion:

- “1. A belief can contribute to the complexity/simplicity of a theory. [fact]
2. The complexity/simplicity of a theory is relevant to cognition. [fact]
3. What a belief contributes to the complexity/simplicity of a theory is relevant to cognition. [1, 2]
4. If a belief contributes to the complexity/simplicity of a theory, then it does so by way of having a degree of simplicity of the same sort as the degree of simplicity the theory has.
5. If the property of an element and the property of a system to which it is added are determinants of the same determinable, then the contribution of that property of the element to that property of the system to it is added is additive.
6. What a belief contributes to the complexity/simplicity of a theory is additive. [4, 5]
7. If the computational theory of mind is true, then what a belief contributes to the complexity/simplicity of a theory must be a context-invariant property of it. [CTM, 3]
8. If the computational theory of mind is true, then what a belief contributes to the simplicity of a theory is always a constant increment/decrement. [6, 7]
9. If a belief can only contribute to the complexity of a theory by contributing a constant increment or decrement, then adding a belief to one set of attitudes cannot make a minimal difference to the complexity of an agent’s theory, while adding the same belief (a belief of the same type) to another set of attitudes makes a significant difference to the complexity of the agent’s theory.
10. However, adding a belief to one set of attitudes can make a minimal difference to the complexity of an agent’s theory, while adding the same belief to another set of attitudes makes a significant difference to the complexity of an agent’s theory.
11. Therefore, CTM is false. [8-10]” (Schneider, 2011: 81-82).

I omformuleringen er simpliciteten af en antagelse altså den samme *slags* som simpliciteten af en teori eller plan, antagelsen tilføjes. Dette var ikke givet i det oprindelige argument med Fodors definition (Schneider, 2011: 79). Derudover er simplicitet nu forstået som en additiv egenskab, hvor forøgelsen af antagelser er lig med forøgelsen af kompleksitet. Dette gør det ifølge Schneider mere

plausibelt at simpliciteten af en antagelse kun kan være intrinsisk, hvis den forårsager den *samme grad* af ændring i simpliciteten af en teori eller plan, som antagelsen tilføjes (Schneider, 2011: 79).

Schneider afviser også dette omformulerede argument. Hun bemærker først og fremmest at der kan stilles spørgsmålstejn ved præmis 4. Hun mener for eksempel ikke det er tydeligt at den simplicitet som 'no wind tomorrow' har, er den samme slags simplicitet som planerne at køre til Chicago og at sejle til Chicago har. Hun sætter også andre spørgsmålstejn for at illustrere uklarheden ved Fodors beskrivelser af simplicitet og eksempler på beregninger af simplicitet. De skal imidlertid ikke udgøre gendrivende indvendinger (Schneider, 2011: 82-83). Schneiders gendrivende indvending mod dette argument består i et angreb på præmis 5. Hun mener at det demonstrativt kan vises at denne præmis ikke er sand (Schneider, 2011: 84). Hun argumenterer for dette ved at give et eksempel på hvordan et element med en given egenskab (en form for simplicitet), kan tilføjes til et system af samme elementer, uden at det resulterer i et konstant additivt bidrag af den givne egenskab til dette system. Hun bruger firkanter som eksempler og deres (ydre) sider som den egenskab, der indikerer simplicitet. Det vil sige jo flere sider en firkant eller et system af firkanter har, jo mere kompleks er firkanten eller systemet af firkanter. I figur 1 kan man se at tilføjelsen af en firkant til system (a) ikke består i et additivt bidrag. Tværtimod. Antallet af sider i system (a) bliver formindsket fra 6 til 4 ved tilføjelsen af en firkant. Det er altså muligt at elementer (som antagelser) intrinsisk kan have den samme form for simplicitet som systemer af elementerne, samtidig med simplicitet ikke udelukkende er en additiv egenskab (Schneider, 2011: 84-85). Reformuleringen af Fodors globalitetsargument, med udgangspunkt i simplicitet, er altså også ugyldigt.



Figur 1

Fodor hævder imidlertid også at *konservatisme* og herunder *centralitet*, udgør eksempler på egenskaber som antagelser har gennem deres relation til hele det antagelsessystem, de er en del af. Schneider har altså ikke afvist hele globalitetsargumentet ved at vise at simplicitet *kan* hvile på syntaks (Schneider, 2011: 85). I stedet for at forholde sig direkte til Fodors argumentation med

konservatisme og centralitet, argumenterer hun for at det er usandsynligt, at der skulle findes egenskaber ved elementer, som er globale egenskaber sådan som (Schneider mener at) Fodor forstår dem. Det vil sige, det er usandsynligt at der findes egenskaber, som elementer har *gennem deres relation* til en større gruppe af disse elementer. Hvis der er et element som vi følger fra kontekst til kontekst (system af elementer til system af elementer), der har forskellige virkninger i de forskellige kontekster “(...) *then the overwhelmingly story is that a context-invariant property of it (...), together with the context-invariant properties of the other items in the system, explains why*” (Schneider, 2011: 87). Det er sådan naturen fungerer, så vidt vi ved indtil videre. Der er ikke noget, der tyder på at der findes elementer med egenskaber, som de *kun* har i bestemte kontekster. Schneider bemærker at videnskaben underbygger dette:

“(...) the history of science provides inductive evidence for the overhypothesis that relational properties like this do not figure in brute laws, namely, that we have not so far had to countenance them in any area as ultimate laws. For example, Kepler’s laws of planetary motion appeal to the property of being a planet but have been explained in terms of more fundamental laws which dispense with this higher-level relational property” (Schneider, 2011: 88).

I og med at det er meget usandsynligt at der findes elementer med relationelle egenskaber, og i og med at Fodors hovedeksempel på en relationel egenskab, simplicitet, ikke viste at simplicitet nødvendigvis er en relationel egenskab, mener Schneider således at have afvist hele globalitetsargumentet (Schneider, 2011: 87-89).

Mit svar

Jeg er enig i den del af Schneiders indvending der afviser Fodors argumentation, ud fra hans definition af simplicitet. Han har, som jeg ser det, unægteligt defineret simplicitet, sådan at det nødvendigvis er en kontekstinvariant egenskab. Jeg mener imidlertid ikke Schneiders argumentation som helhed får afvist globalitetsargumentet. For det første mener jeg at hendes omformulerede argument mangler en del af Fodors argumentation, hvilket medfører at hendes indvending mod det omformulerede argumentet mislykkes. Jeg er enig i at hendes argumentation med systemer af firkanter viser at antagelsers simplicitet godt kan være intrinsisk på trods af, at den samme antagelse forårsager en

minimal ændring i kompleksitet hos en teori og en signifikant ændring hos en anden teori. Jeg er således enig i at hun får vist at antagelsers og teoriers simplicitet *kan* have en syntaks. Hun får dermed også gendrevet selve omformuleringen af argumentet. Problemet er som sagt at hendes omformulering af argumentet ikke inkluderer hele Fodors argumentation. Fodor er enig i at teoriers simplicitet kan have en syntaks. Problemet er at syntaksen der i så fald repræsenterer teoriers simplicitet, er en syntaks teorierne har som helhed. Simplicitet er en holistisk egenskab, og det er (blandt andet) det, der er problemet. Hvis teoriers simplicitet er en holistisk egenskab, er et symbolmanipulerende system dermed nødt til at læse teorier som helhed, og det er et problem for CTM, både fordi

1. en hel teori således, for at CTM er korrekt, er nødt til at udgøres af én enkelt repræsentation som vores system beregner på på én gang, hvilket vores system givetvis ikke gør, da vi generelt ikke kan begribe en hel teori på en gang eller beregner globale egenskaber ud fra ideer der indeholder hele teorier

eller fordi

2. en teori ikke kan adskilles fra de antagelsessystemer de er indlejret i og et symbolmanipulerende system således er nødt til at beregne på hele vores antagelsessystem for at beregne en teoris simplicitet, hvilket et symbolmanipulerende system ikke kan gøre 'in real time'

Schneiders argumentation får *ikke* vist at simplicitet ikke er en holistisk egenskab. Tværtimod. Hvis simpliciteten af antagelser er af samme form som simpliciteten af firkanter, så illustrerer hendes eksempel netop hvordan simplicitet er en holistisk egenskab ved systemer af antagelser. Simpliciteten af et system af firkanter, som det Schneider giver et eksempel på, er en holistisk egenskab. Summen af simpliciteten af firkanterne hver for sig er en anden end simpliciteten af firkanteren sammen som en helhed. For sig har tre firkanter simpliciteten $(4 + 4 + 4) 12$, mens de sat sammen i et system har simpliciteten 6. Schneiders firkant-eksempel illustrerer dermed ikke andet end det Fodor hævder er et problem. At Schneider ikke selv ser dette, skyldes formentlig at hun opfatter Fodors argumentation som et *in principle-argument*, der har til formål at vise, at simplicitet ikke kan være en intrinsisk egenskab og dermed *umuligt* kan hvile på syntaks.

Udover at jeg mener hendes argumentation omkring omformuleringen af simplicitetsargumentet er forfejlet, mener jeg også hendes afvisning af den del af Fodors argumentation der hviler på konservatisme og centralitet, er forfejlet. For det første udgør hendes argumentation en svag afvisning. Argumentation består ikke i andet end en argumentation for, at det i det hele taget er højst usandsynligt, at der findes globale egenskaber. At hævde at en påstand er overvældende usandsynlig uden at forholde sig til argumenterne for påstanden direkte, er efter min mening en svag afvisning af en påstand. Hvis der er en der i ramme alvor hævder at en baby er bundet fast i toppen af en flagstang, og beder os om hjælp til at få babyen ned, så afviser vi ikke denne påstand fordi vores erfaringer fortæller os at det er overvældende usandsynligt at en baby er bundet fast til toppen af en flagstang. Vi går ud og ser om der rent faktisk er en baby bundet fast til flagstangen.

Derudover forholder hendes argumentation sig ikke til hele Fodors argumentation. Fodor hævder, at det er et problem for CTM, hvis systemer af antagelser har holistiske egenskaber, som vi beregner i vores daglige slutningsprocesser *uden* at begribe hele systemet på én gang. Hendes argumentation for at der sandsynligvis ikke findes globale egenskaber, afviser ikke at der findes holistiske egenskaber, og hendes forudgående eksempel med firkanterne viser, at simplicitet i princippet sagtens kan være en holistisk egenskab ved systemer af antagelser.

Schneiders argumentation gendriver altså ikke globalitetsargumentet. Schneider stiller imidlertid spørgsmålstejn ved Fodors afgrænsning af simplicitet og påpeger at det ikke er åbenlyst at 'no wind tomorrow' har samme simplicitet som planerne at køre til Chicago og at sejle til Chicago. Her mener jeg at hun har fat i noget. Jeg mener at det er åbenlyst at 'no wind tomorrow' har en anden slags simplicitet end planerne. Jeg mener derudover helt klart, at der er problemer ved at Fodor ikke afgrænser sin forståelse af simplicitet og hvornår man beregner simplicitet. Fodors påstand om at simplicitet er en global egenskab, og at vi således beregner simpliciteten af to teorier *ved* at beregne simpliciteten af teorierne og det antagelsessystem de er indlejret i sammen, mener jeg for eksempel er meget tvivlsom. Min intuitive forståelse af teoriers simplicitet er, at simpliciteten *er* et resultat af summen af antagelserne i teorierne. I mit eksempel med Pernille, der skal vurdere to teoriers simplicitet (Albertes eller Barbaras, "epicykler" vs. "heliocentrisme"), vurderer jeg at Barbaras teori er mere simpel end Albertes, netop fordi Barbaras teori indeholder færre antagelser. Albertes teori består groft sagt af følgende antagelser eller påstande:

- Planeterne cirkulerer konstant og i samme retning om jorden
- Solen cirkulerer konstant og i samme retning om jorden
- Månen cirkulerer konstant og i samme retning om jorden
- Fire planeter cirkulerer tilmed i epicykler

Barbaras teori består groft sagt af følgende antagelser:

- Planeterne cirkulerer konstant og i samme retning om solen
- Jorden cirkulerer konstant og i samme retning om solen
- Månen cirkulerer konstant og i samme retning om jorden

Alle andre antagelser har de to teorier tilfælles (og det er i øvrigt derfor at Pernille overhovedet ønsker at beregne simpliciteten). Det er denne, her ekspliciterede, forskel i antallet af antagelser, der gør at jeg mener Barbaras teori er mere simpel en Albertes teori. Jeg har svært ved at se hvad det ellers skulle være.

Det er altså ikke tydeligt at simplicitet er en global egenskab. Fodors eksempel på at antagelsen ‘one or two planetary regressions occurred’ tydeligvis medfører en minimal ændring i kompleksitet for et teorisystem og en signifikant ændring for et andet teorisystem, mener jeg ikke modsiger dette. For det der beregnes her, er, så vidt jeg kan se, ikke *simpliciteten* af teorisystemerne, men hvor mange antagelser i teorisystemerne, som kendsgerningen ‘one or two planetary regressions occurred’ modsiger. Hvis det var simpliciteten af teorisystemerne man vurderede, ville det kræve at man omformulerede teorisystemerne så de var tilpasset til ‘one or two planetary regressions occurred’. Det kan godt være, at vi forsøger at tilpasse teorisystemet i vores hovedet på forskellige måder, men det er ikke *det*, der gør at vi konkluderer at teorisystemet er mere eller mindre ‘kompliceret’ (i Fodors forstand). Det der gør at vi konkluderer at et teorisystem ‘kompliseres’ af inklusionen af en antagelse er, at vi kan ‘se’ at antagelsen *modsiges* en større eller mindre del af det teorisystemet hævder.

Denne argumentation er selvfølgelig ikke tilstrækkelig til at vise at simplicitet ikke er en holistisk egenskab, men jeg mener den er tilstrækkelig til at vise at Fodor ikke klart og tydeligt får vist at simplicitet *er* en holistisk egenskab. Schneiders pointe om at der kan stilles spørgsmålstejn ved hans forståelse af simplicitet, giver altså, som jeg ser det, god grund til at tvivle på den del af globalitetsargumentet, der hævder at simplicitet er en global(/holistisk) egenskab.

Opsamling

Jeg mener således at ingen af Schneiders indvendinger mod globalitetsargumentet er holdbare, men at hun har ret i at der kan stilles alvorlige spørgsmålstejn ved en del af globalitetsargumentet. Hun får gendrevet en del af Fodors argumentation, men denne del kan omformuleres og hun får ikke gendrevet denne omformulerede del af Fodors argumentation hvis man tager resten af hans argumentation i betragtning.

Hendes første indvending mod globalitetsargumentet består i at vise at et skakprogram 'løser' globalitetsproblemet. At et skakprogram løser' et globalitetsproblem viser imidlertid ikke i sig selv at vores kognitive system kan være et symbolmanipulerende system. Fodor afviser ikke at det er muligt at et symbolmanipulerende system kan 'løse' et globalitetsproblem. Han afviser at vores kognitive system gør, eller kan gøre det, sådan som et symbolmanipulerende system muligvis kan gøre det. Da Schneider ikke uddyber hvordan skakprogrammet løser det hun forstår ved et globalitetsproblem, viser hendes eksempel på et skakprogram ikke at Fodors globalitetsargument er ugyldigt.

Hendes anden indvending viser heller ikke at globalitetsargumentet er ugyldigt. Hun misforstår den del af Fodors argumentation hun forholder sig til. Hun tror at en af hans præmisser i argumentationen er, at CTM ikke kan være sand hvis vi beregner på mere end en meget lille gruppe af antagelser, når vi finder frem til hvilke planer der er de mest simple. Dette er ikke en præmis i Fodors argumentation. Hendes angreb på denne formodede præmis, er altså ikke et angreb på en præmis i Fodors argumentation. Schneider er selv opmærksom på at det muligvis er tilfældet, men påpeger at argumentet i så fald bliver til et relevansargument. Det er jeg enig i, men det hjælper ikke hendes argumentation, for hun har ikke afvist relevansargumentet.

Hendes sidste indvending er overordnet delt i to. Én indvending mod Fodors argumentation for at simplicitet ikke kan være en intrinsisk egenskab ved antagelser og en anden indvending der hævder, at der formentlig slet ikke findes globale egenskaber. Hendes indvending mod Fodors argumentation for at simplicitet ikke kan være en intrinsisk egenskab ved antagelser, mener jeg er korrekt. Fodor har defineret antagelsers simplicitet således at det nødvendigvis er en kontekstinvariant egenskab ved antagelser, og det ødelægger hans argumentation. Schneider omformulerer imidlertid argumentet (hvilket jeg i øvrigt mener formentlig er en omformulering, der svarer til Fodors intenderede argumentet). I forhold til omformuleringen får hun gendrevet argumentet som det er omformuleret. Omformuleringen mangler imidlertid den del af Fodors argumentation der påpeger at holistiske egenskaber ved systemer af antagelser er et problem for CTM. Dette gør at Schneiders

gendrivelse af omformuleringen ikke er en gendrivelse af globalitetsargumentet. Hendes gendrivelse implicerer nemlig at simplicitet er en holistisk egenskab ved systemer af antagelser. Hendes argumentation vedrørende det omformulerede simplicitetsargument er altså ikke holdbar.

Hendes argumentation for at der i det hele taget formentlig ikke findes globale egenskaber, er for det første en svag argumentation, da hun argumenterer for det uden at forholde sig direkte til Fodors argumentation for at centralitet (eller nogen andre egenskaber) er en global egenskab. Derudover implicerer hendes argumentation vedrørende simplicitet, at teorier og antagelsessystemer kan have holistiske egenskaber. Hvis dette er tilfældet, udgør det et problem for CTM ifølge Fodors argumentation, og hun får ikke afvist at holistiske egenskaber er et problem for CTM. Hendes argumentation mod resten af globalitetsargumentet er heller ikke holdbar.

Schneider påpeger imidlertid at der kan stilles spørgsmålstejn ved hvordan simplicitet skal forstås og om Fodors eksempler på beregninger af simplicitet giver mening. Det mener jeg hun har ret i, og jeg mener også at der er noget galt med den måde han forstår simplicitet på. Schneider får således vist at den del af globalitetsargumentet der hævder at simplicitet er en global egenskab, mildest talt ikke er klart og tydeligt og der derved kan stilles alvorlige spørgsmålstejn ved argumentet.

Konkluderende bemærkninger

Udgangspunktet for denne afhandling var at psykologien har problemer med at afgrænse sin genstand og med at forene sig om en forståelse af hvordan man kan undersøge denne genstand. Jeg hævdede at vi med grundige undersøgelser af de forskellige tilgange til psykologien, kan bevæge os mod en forenet forståelse af psyken og de forskellige måder at undersøge den på. Jeg ville med min afhandling tage et skridt i denne retning. Jeg valgt at undersøge et filosofisk grundlag for den kognitionspsykologiske tilgang for at tage dette skridt. Nærmere bestemt valgte jeg at undersøge CTM som udgør et filosofisk grundlag for påstanden om at vores kognitive system fungerer som en computer. Fodor, en af de mest fremtrædende filosoffer der har argumenteret *for* CTM, har imidlertid samtidig argumenteret for at CTM formentlig kun kan gælde for en meget begrænset del af vores kognitive processer. Jeg valgte derfor at undersøge CTM med udgangspunkt i hans to-delte argumentation for dette. Jeg undersøgte Fodors argumentation ved at undersøge fem indvendinger, som Schneider har fremført mod hans argumentation. Jeg gennemgik alle disse indvendinger og konkluderede at ingen af dem var holdbare. Schneider fik gendrevet en del af Fodors ene argument,

men den del omformulerede hun til et andet argument, og hendes gendrivelse af dette omformulerede argument konkluderede jeg var forfejlet. Jeg konkluderede imidlertid også at hun, ved siden af sin forfejlede gendrivelse, fik sat alvorlige spørgsmålstejn ved den omformulerede del af Fodors argumentation. Hun fik mere præcist sat alvorlige spørgsmålstejn ved Fodors påstand om, at vores beregninger af antagelsers og teories simplicitet udgør et problem for CTM.

Hvor efterlader dette os så i forhold til vores forståelse af psykologiens genstand og hvordan vi skal undersøge den? I og med Schneider kun får sat spørgsmålstejn ved en meget lille del af Fodors argumentation, peger det i retning af at vores kognitive system generelt ikke kan forstås som en computer. Vi har således taget et meget lille skridt (men *et skridt*) i en negativ retning: Vi har (fortsat) god grund til at tro at CTM ikke er (hele) sandheden om vores kognitive system. Dette indebærer desuden at vi formentlig ikke kan betragte vores epistemiske indstillinger som relationer mellem en organisme og en mental repræsentation (i CTMs forstand) og at forbindelserne mellem de epistemiske indstillinger ikke kan beskrives gennem symbolmanipulerende tilstande. Det lille skridt, som det er at have god grund til at opgive CTM, indebærer således en afgrænsning af hvad epistemiske indstillinger formentlig ikke er og hvordan de formentlig ikke forårsager hinanden. Det giver derudover en retning for nye undersøgelser: Hvad *er* epistemiske indstillinger? Hvad *er* antagaler? Hvad *er* ønsker? Hvad *er* frygt? Hvordan forårsager disse mentale tilstande hinanden? *Forårsager* de overhovedet hinanden?

Måske kan vi få svar på disse spørgsmål ved at undersøge de filosofiske grundlag for de andre tilgange til psykologien.

Litteraturliste

Bøger:

- Bailey, Andrew (2013): *Philosophy of Mind: The Key Thinkers*. 1. udgave. Bloomsbury Publishing Plc.
- Christensen, Gerd (2014): *Psykologiens Videnskabsteori: En Introduktion*. 2. udgave. Samfundslitteratur.
- Cooper, Stuart Barry & Leeuwen, Jan van (2013): *Alan Turing: His Work and Impact*. 1. udgave. Elsevier Science & Technology.
- Dreier, Ole (2007): *Psychotherapy in Everyday Life*. Cambridge University Press.
- Eysenck, Michael W. & Keane, Mark T. (2013): *Cognitive Psychology: A Student's Handbook*. 6. udgave. Psychology Press.
- Flyvbjerg, Bent & Sampson, Steven (2001): *Making Social Science Matter: Why Social Inquiry Fails and How it Can Succeed Again*. Cambridge University Press.
- Fodor, Jerry A. (1968): *Psychological Explanation*. New York: Random House.
- Fodor, Jerry A. (1978): *The Language of Thought*. Reprinted. The Harvester Press.
- Fodor, Jerry A. (1983): *The Modularity of Mind*. 1. udgave. A Bradford Book.
- Fodor, Jerry A. (1987): *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. 1. udgave. A Bradford Book.
- Fodor, Jerry A. (1998): *Concepts: Where Cognitive Science Went Wrong*. 1. udgave. Oxford University Press.
- Fodor, Jerry A. (2001): *The Mind Doesn't Work That Way: The scope and limits of computational psychology*. MIT Paperback 1. udgave. A Bradford Book.
- Fodor, Jerry A. (2008): *LOT 2: The Language of Thought Revisited*. 1. udgave. Oxford University Press.
- Holzkamp, Klaus, et al. (2013): *Psychology from the Standpoint of the Subject: Selected Writings of Klaus Holzkamp*. Redigeret af Schraube, E. & Osterkamp, U. Palgrave Macmillan.
- Pinker, Steven (1997): *How the Mind Works*. 1. udgave. Penguin Press.

- Schneider, Susan (2011): *The Language of Thought: A New Philosophical Direction*. 1. udgave. MIT Press.
- Sterelny, Kim (1990): *The Representational Theory of Mind: An introduction*. 1. udgave. Basil Blackwell.

Artikler:

- Fodor, Jerry A. & Pylyshyn, Zenon W.: “Connectionism and Cognitive Architecture: A Critical Analysis”. (mar. 1988) Publiceret af Elsevier i *Cognition*, Vol. 28, Issue 1-2.
- Fodor, Jerry A.: “Fodor’s Guide to Mental Representation: The Intelligent Auntie’s Vade-Mecum”. (jan. 1985) Publiceret af Oxford University Press i *Mind*, Vol. 94 No. 373, Mind Association.
- Fodor, Jerry A.: “You Can Fool Some of The People All of The Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations”. (jan. 1991) Publiceret af Oxford University Press i *Mind*, Vol. 100 No. 1, Mind Association.
- Fodor, Jerry A.: “Reply to Steven Pinker ‘So How Does the Mind Work?’”. (feb. 2005) Publiceret af Oxford i *Mind & Language*, Vol. 20, Issue 1.
- Pinker, Steven: “So How Does the Mind Work?” (feb. 2005) Publiceret af Oxford i *Mind & Language*, Vol. 20, Issue 1.
- Shanahan, Murray & Baars, Bernard: “Applying Global Workspace Theory to the Frame Problem”. (dec. 2005) Publiceret af Elsevier i *Cognition*, Vol. 98, Issue 2.

Internetsider:

- Web 1: [The Computational Theory of Mind \(Stanford Encyclopedia of Philosophy\)](#) Besøgt d. 02.04.2021.