



LUNDS
UNIVERSITET

Medvetna mentala tillstånd och superveniens

En undersökning av David Chalmers zombieargument

Examensarbete för kandidatexamen i teoretisk filosofi

Julia Sjödahl

Filosofiska institutionen
Handledare: Martin Jönsson

2014-01-22

INNEHÅLLSFÖRTECKNING

| | |
|--|----|
| 1. Inledning | 2 |
| 2. Chalmers zombieargument | 2 |
| 2.1. De enkla problemen och det svåra | 3 |
| 2.2. Problemet med reduktiva förklaringar | 3 |
| 2.3. Naturalistisk dualism | 4 |
| 2.4. Zombieargumentet | 4 |
| 3. Exempel på svag kritik av zombieargumentet | 6 |
| 3.1. Alla ni zombies | 6 |
| 3.2. Varför zombies inte kan existera | 8 |
| 3.3. Zombiedöden | 12 |
| 4. Kan möjlighet härledas från tänkbarhet? | 15 |
| 4.1. Standardinvändningen | 15 |
| 4.2. Chalmers svar på standardinvändningen | 16 |
| 4.2.1. Tvådimensionell semantik | 16 |
| 4.2.2. Chalmers lösningsförslag | 18 |
| 5. Kritik av TM-tesen | 19 |
| 5.1. Tänkbart, möjlighet och kropp-själ-problemet | 19 |
| 5.2. Chalmers tänkbarhetsargument för dualism | 23 |
| 5.3. Tänkbart, möjlighet och fysikalism | 24 |
| 5.4. Anti-zombieargumentet | 27 |
| 6. Diskussion och slutsats | 30 |
| 7. Källförteckning | 33 |

1. INLEDNING

Debatten kring kropp-själ-problemet, det vill säga sättet på vilket relationen mellan det fysiska och det mentala förhåller sig, är en långt tillbakagående diskussion inom filosofin.¹ Många försök att förklara denna relation har gjorts, och det råder traditionellt sett stark oenighet mellan två näst intill diametralt motsatta sätt att se på problemet. Det första är monism, synen på medvetna mentala tillstånd som något som inte är distinkt från det fysiska, och det andra är dualism; synen på medvetna mentala tillstånd och fysiska tillstånd som separata entiteter.²³

Denna uppsats kommer bestå av en undersökning av ett klassiskt argument för dualism i modern tappning, nämligen David Chalmers zombieargument. Detta argument kan i korta drag sägas bestå av påståendet att man kan tänka sig att det skulle kunna existera varelser som saknar medvetna mentala tillstånd, att dessa varelser givet sin tänkbarhet är möjliga, och att mentala egenskaper därmed inte determineras av fysiska.

Renée Descartes presenterade ett av de första argumenten för dualism av detta slag i verket *Betraktelser över den första filosofien* (1641). Argumentets poäng är att givet tänkbarheten av själen som något som skulle kunna existera utan kroppen kan man dra slutsatsen att kropp och själ verkligen är separata entiteter, och därmed att dualismen är riktig. Descartes argument kan skrivas ut på följande sätt:

P1: Jag kan tänka mig att jag skulle kunna existera utan min kropp

P2: Allt jag kan tänka är logiskt möjligt

Så

S: Min kropp och jag är separata entiteter

Den här uppsatsens syfte är att undersöka Chalmers zombieargument för att se om det är ett giltigt och sunt argument eller inte. Detta kommer jag göra genom att först presentera problematiken som får Chalmers att anta en dualistisk position, för att sedan redogöra för zombieargumentet samt vad det är Chalmers försöker belysa med det. Därefter kommer jag undersöka en rad invändningar mot zombieargumentet; först tre svagare argument, som görs på tre olika grunder, och sedan fyra starkare, som har en gemensam nämnare i att alla angripa det faktum att Chalmers menar att möjlighet kan härledas från tänkbarhet. I uppsatsens sista del presenterar jag ett lösningsförslag genom vilket jag tror att man skulle kunna visa att zombies är logiskt möjliga utan att involvera denna tänkbarhetsaspekt.

2. CHALMERS ZOMBIEARGUMENT

I artikeln *Facing Up to the Problems of Consciousness* (1995) presenterar David Chalmers vad han menar är det svåra problemet med medvetna mentala tillstånd samt ett lösningsförslag i form av en position han kallar för naturalistisk dualism. I boken *The Conscious Mind* (1996)

¹ ”Det mentala” är översättningar av engelskans ”mind”.

² ”Medvetna mentala tillstånd” är en översättning av ”conscious states of mind”.

³ I denna uppsats kommer fysikalism, även kallad materialism, det vill säga synen på medvetna mentala tillstånd som fysiska tillstånd, vara den enda variant på monism som tas upp.

presenteras utöver detta ett antal argument för den egna positionen, varav ett är det så kallade zombieargumentet.

2.1. DE ENKLA PROBLEMEN OCH DET SVÅRA

Det finns enligt Chalmers många problem som handlar om medvetna mentala tillstånd. Somliga av dem skulle teoretiskt gå att besvara med hjälp av kognitionsvetenskapliga metoder, och dessa kallar Chalmers för de enkla problemen. Exempel på sådana problem är bland annat vad det innebär att fokusera samt skillnaden mellan att vara vaken och att sova. Trots att vi ännu inte har någon förklaring på dessa fenomen är det enligt Chalmers inte svårt att föreställa sig sätt på vilka de skulle kunna förklaras (1995, s. 200).

Det svåra problemet med medvetna mentala tillstånd är, enligt Chalmers, problemet med *upplevelser*; att det finns en slags subjektiv aspekt när vi tänker och varseblir. Ett exempel är att när elektromagnetiska vågor träffar ögat, bryts av hornhinna och lins och projiceras på hornhinnan så *upplever* vi något; hur rött något är, att något är ljust eller mörkt eller dylikt; vi får en känsla. När vissa fysiska processer sker i kroppen så upplever vi sådant som smärta, njutning, smaker, ljud och så vidare. Vad som är gemensamt för dessa tillstånd, menar Chalmers, är att *det är på ett visst sätt att ha dem* (1995, s. 201).⁴ De är tillstånd som vi *upplever*. Problemet formulerar Chalmers på följande sätt:

”Why is it that when our cognitive systems engage in visual or auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion? It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises.” (Chalmers, 1995, s. 201.)

2.2. PROBLEMET MED REDUKTIVA FÖRKLARINGAR

Gemensamt för de enkla problemen är enligt Chalmers att de skulle kunna förklaras funktionellt; man beskriver helt enkelt en mekanism. För att förklara något på detta sätt behöver man specificera de involverade mekanismerna. Chalmers ger ett exempel med gener och mekanismer som överför arvs massa från en generation till nästa – DNA; när vi har förklarat hur funktionen fungerar har vi förklarat fenomenet vi ville undersöka från början. På samma sätt fungerar det enligt Chalmers med de enkla problemen; när mekanismen är förklarad är problemet löst (1995, s. 206)

Chalmers menar dock att denna typ av förklaring inte fungerar för att förklara medvetna upplevelser. Det svåra problemet går *bortom* utförandet av funktioner. Även efter att man har förklarat alla inblandade kognitiva funktioner i exempelvis perception så kvarstår frågan om *varför dessa funktioner ackompanjeras av en upplevelse*. Det finns ett glapp i förklaringen mellan funktionen och upplevelsen, och vägen dem emellan kan inte vara en funktion, utan måste bestå av något annat, en *extra ingrediens*, som Chalmers kallar det (1995, s. 207).

⁴ ”Att det är på ett visst sätt att ha dem” är en översättning av ”that there is something it is like to be in them”, se Thomas Nagels *What Is It Like To Be A Bat?* (1974).

Problemet med alla sådana som hittills har föreslagits är dock detsamma som tidigare; varför skulle just denna extra ingrediens ge upphov till medvetna mentala tillstånd?

2.3. NATURALISTISK DUALISM

Chalmers förslag är att se upplevelser som fundamentala och icke-analyserbara entiteter, och konstruera en icke-reduktiv teori om medvetna mentala tillstånd utifrån detta.⁵ Syftet med denna teori är att visa hur fysiska egenskaper ger upphov till upplevelser genom att postulera så kallade psykofysiska principer. De psykofysiska principerna ska ses som komplement till de fysiska lagarna, som en form av bro mellan dem och det mentala, och Chalmers menar att dessa principer är den extra ingrediens som behövs för att förklara hur något fysiskt kan ge upphov till en upplevelse (1995, s. 213).

Genom att på detta sätt postulera egenskaper och lagar utöver de fysiska antar Chalmers en form av dualism för att kunna förklara hur upplevelser kan komma ur fysiska processer, som en form av epifenomenalism; de medvetna mentala tillstånden uppkommer i hjärnan men är något utöver de fysiska hjärntillstånden. Chalmers kallar sin position för *naturalistisk dualism*, och menar att den är dualism av ett oskyldigt slag som är kompatibel med den vetenskapliga världsbilden; inget i den motsäger något i fysikens teorier om hur världen förhåller sig, utan det handlar bara om att utöka dem med ytterligare principer (1995, s. 214).

2.4. ZOMBIARGUMENTET

För att bevisa att reduktiva förklaringar inte fungerar behöver Chalmers visa att medvetna mentala tillstånd inte logiskt supervenierar på fysiska tillstånd, det vill säga visa att medvetna mentala tillstånd inte kan härledas ur mikrofysiska fakta i världen.⁶

Chalmers argumenterar för att fenomenella fakta inte kan härledas *a priori* från fysiska fakta. Nödvändigheten som definierar den relevanta superveniensrelationen är enligt Chalmers *a priori*-versionen av logisk nödvändighet, där primära intensioner (se s. 16) är centrala (1996, s. 94). Som en del i att visa att medvetna mentala tillstånd inte logiskt supervenierar på fysiska tillstånd argumenterar Chalmers för zombiers logiska möjlighet. Chalmers zombieargument bygger på idén om att zombier är *tänkbara*; det vill säga att vi utan problem kan föreställa oss dem.⁷ Argumentet har följande form:

⁵ Icke-analyserbara entiteter är sådana entiteter som inte kan förklaras i termer av något enklare.

⁶ Superveniens är en relation där en typ av egenskaper determinerar en annan sorts egenskaper. B-egenskaper supervenierar på A-egenskaper om det inte finns två möjliga situationer som är identiska i termer av A-egenskaperna men som är olika i termer av B-egenskaper. *Lokal superveniens* handlar om när en individs A-egenskaper determinerar dess B-egenskaper, exempelvis har objekt med samma fysiska egenskaper samma form; formen supervenierar på de fysiska egenskaperna. *Global superveniens* handlar om att A-fakta i en hel värld determinerar B-fakta (1996, s. 32-34). Chalmers menar att lokal och global superveniens står och faller tillsammans, och att det därför inte spelar någon roll om han utför argumentet på global eller lokal nivå. Om medvetna mentala tillstånd över huvud taget supervenierar, menar Chalmers, så supervenierar det förmodligen lokalt, och därför kan han argumentera på individnivå, men argumentet kan även föras på global nivå (1996, s. 93).

⁷ "Tänkbar" är en översättning av "conceivable".

P1: Zombies är tänkbara

P2: Om något är tänkbart så är det logiskt möjligt

Så

S1: Zombies är logiskt möjliga

S2: Det mentala supervenierar inte på det fysiska

En zombie är något eller någon som är identiskt med en vanlig person, bortsett från att den saknar medvetna mentala tillstånd. Där mitt medvetande finns är det för zombien bara ”svart”; den har inga medvetna upplevelser. Dessa så kallade *fenomenella zombies* är inte som ”Hollywoodzombies”, som saknar förmåga till introspektion och beteendekontroll (men som skulle kunna ha medvetna upplevelser; det skulle mycket väl kunna finnas ett *sätt på vilket det är* för en Hollywoodzombie att äta en människa till exempel), utan vad Chalmers är intresserad av är sådana zombies som är fysiskt, funktionellt och psykologiskt identiska med vanliga människor, men som saknar upplevelser (1996, s. 95). Vi kan på global nivå till och med tänka oss en hel zombievärld; en värld som är helt och hållet identisk med vår egen, allt där är intill minsta molekyl exakt likadant som i vår värld, bortsett från att alla är zombies; det finns i denna möjliga värld inte några medvetna upplevelser alls.

Min zombietvilling är alltså fysiskt, funktionellt och psykologiskt identisk med mig, men den har inga medvetna upplevelser. När jag tittar ut genom fönstret och ser några träd, äter något och känner smaksensationer eller har ont i armen gör min zombietvilling funktionellt sett precis samma saker; den bearbetar samma information och reagerar likadant som jag på samma typer av input. Zombien kan rapportera innehållet i sina mentala tillstånd, fokusera sin uppmärksamhet och så vidare, men inget av dessa funktionella tillstånd åtföljs av några medvetna upplevelser för zombien. Det finns inget *sätt på vilket det är* att vara en zombie.

Chalmers menar att det är högst osannolikt att zombies skulle kunna existera i vår värld, utan en fysisk kopia av mig i vår värld skulle förmodligen ha medvetna mentala tillstånd (1996, s. 96). Detta är dock inte relevant för zombieargumentets slutsats, utan frågan är om idén om zombieidén är begreppsmässigt koherent; att det är en så kallad meningsfull möjlighet. Detta är det enda som krävs för logisk möjlighet och för att därmed fastställa Chalmers slutsats att medvetna mentala tillstånd inte logiskt supervenierar på det fysiska.

Det är i regel inte svårt för oss att tänka oss att andra människor skulle kunna sakna medvetna mentala tillstånd, och Chalmers menar att zombiers logiska möjlighet är uppenbar; en varelse som är fysiskt identisk med mig men utan medvetna upplevelser är möjligen empirisk omöjlig, men enligt Chalmers definitivt koherent; det finns ingen motsägelse i beskrivningen. Att man kan begripa zombiebegreppet och tänka sig dem är enligt Chalmers tillräckligt för att fastställa deras logiska möjlighet (1996, s. 96). Om ingen intern inkohärens kan detekteras finns det enligt Chalmers starka skäl att tro att zombievärlden är logiskt möjlig (1996, s. 99).

Enligt Chalmers är det faktum att medvetna mentala tillstånd inte supervenierar på fysiska tillstånd skälet till att ingen reduktiv teori om medvetna mentala tillstånd någonsin kan vara framgångsrik (1996, s. 106). Givet vilket försök som helst att säga att fysiska processer

underligger medvetna mentala tillstånd uppkommer, som tidigare nämnt, alltid en vidare fråga: varför ger denna process upphov till medvetna mentala tillstånd? Varje fysisk process är lika applicerbara på mig, som har medvetna mentala tillstånd, som på min zombietvilling, som saknar dem. Ur detta följer, enligt Chalmers, att inget i dessa fysiska processer kan förklara varför det för mig uppstår upplevelser (1996, s. 107).

Så, Chalmers syfte med zombieargumentet är att genom att hävda att zombies är logiskt möjliga varelser bevisa att medvetna mentala tillstånd inte logiskt supervenierar på fysiska tillstånd, och därmed att reduktiva förklaringar av medvetna mentala tillstånd inte fungerar. Om det finns en enda möjlig värld där zombier existerar är detta tillräckligt för att fastställa slutsatsen att medvetna mentala tillstånd inte logiskt supervenierar på fysiska tillstånd, och därmed att det inte är möjligt att förklara medvetna mentala tillstånd i reduktiva termer.

3. EXEMPEL PÅ SVAG KRITIK AV ZOMBIEARGUMENTET

Jag har nu presenterat David Chalmers zombieargument. Många personer har sedan det lanserades kritiserat detta argument på en rad olika sätt, och i detta avsnitt följer tre, i min mening, mindre starka invändningar som görs på tre olika grunder. Steffen Borge försöker tillbakavisa zombieargumentet genom att hävda att det leder till ett dilemma, Nigel Thomas hänvisar till betänkligheter om zombies yttrande av satser som innehåller fenomenella termer, och Robert Kirk hävdar att zombies trots allt måste ha perceptuella medvetna mentala tillstånd.

3.1. ALLA NI ZOMBIES

I en artikel från 1999 (*All You Zombies: David Chalmer's Metaphysical Solipsism*) försöker Steffen Borge visa att Chalmers argument bäst förstås som en skeptisk utmaning till materialismen. Om den accepteras leder detta, enligt Borge, till metafysisk solipsism, det vill säga tesen att det enda jag kan vara säker på existerar är jag själv. Borge menar att Chalmers därför står inför ett dilemma: om vi är rättfärdigade att acceptera andras själsliv i den här världen så är vi det också i alla andra möjliga världar med samma fysiska konstitution som vår egen, och om vi inte är rättfärdigade att acceptera andras själsliv i den här världen så följer metafysisk solipsism.

Genom att ha egna medvetna mentala tillstånd står det klart för mig att jag inte lever i en zombievärld. Jag vet att *jag* inte är en zombie, men vad är det som säger att *andra* inte är zombies? Det skulle ju faktiskt kunna vara så att jag lever i en värld som delvis består av zombies. Borge menar att det enligt Chalmers inte finns någon anledning att tro att vi lever i en delvis-zombievärld; genom våra egna mentala upplevelser kan vi konstruera så kallade psykofysiska lagar, och från dem härleda att andra också har mentala upplevelser. Genom dessa lagar är vi i vår värld alltså rättfärdigade att tro att andra personer än vi själva har själsliv. De psykofysiska lagarna kräver inget utöver det fysiska, och därför borde materialism vara den bästa förklaringen, menar Borge (1999, s. 80).

Regeln för universell generalisering lyder som följer: om man för *en godtycklig individ* ur en domän lyckas bevisa något, så gäller det bevisade för *alla* individer inom samma domän.⁸ Borge argumenterar genom denna regel för att om vi för *en godtycklig värld* ur domänen av *alla möjliga världar med samma fysiska konstitution som vår egen* kan bevisa att vi är rättfärdigade att tro att andra personer än vi själva har mentala upplevelser så är vi rättfärdigade till detta i alla möjliga världar i samma domän. Borge menar att Chalmers, genom att hänvisa till att vi i vår värld genom de psykofysiska lagarna är rättfärdigade att tro att andra personer har mentala upplevelser, har bevisat detta för en godtycklig individ i domänen av alla möjliga världar med samma fysiska konstitution som vår egen (1999, s. 81). Genom att vi i *en enda värld* inom domänen är rättfärdigade att tro att vi inte lever i en delvis-zombievärld så finns det ingen värld i domänen där vi är rättfärdigade att tro att zombies är tänkbara.

Borge menar att detta leder till ett dilemma. Antingen kan vi i vår värld postulera enkla psykofysiska lagar och behöver inte oroa oss över möjligheter i andra världar, och om vi då endast beaktar vår egen värld, isolerad, så skulle man därmed kunna säga att medvetna mentala tillstånd inte är någonting utöver det fysiska. Detta skulle enligt Borge vara den bästa och enklaste förklaringen för de här lagarna (1999, s. 81). Härledning till bästa förklaring leder alltså tillbaka till materialism. Om vi är epistemiskt förpliktigade att tro på den bästa tillgängliga förklaringen så är vi enligt Borge förpliktigade att tro att andra personer har medvetna mentala tillstånd i alla relevanta möjliga världar, *och* att tro på materialism i alla relevanta möjliga världar (1999, s. 81).

Å andra sidan, menar Borge, gäller att även om zombievärldar inte är epistemiskt tillgängliga för oss så är de fortfarande logiskt möjliga. Givet zombievärldarnas tänkbarhet kommer det finnas åtminstone en möjlig värld där våra psykofysiska lagar misslyckas med att peka ut sambandet mellan det fysiska och det mentala, eftersom det i denna värld inte finns några mentala egenskaper. Dessutom är dessa lagar enligt Borge inte tillräckligt bevis för andras medvetna mentala tillstånd. Om det för en enda värld i domänen gäller att rättfärdigandet inte finns där så gäller det genom lagen för universell generalisering för alla världar i domänen. Detta är dilemmats andra alternativ; i ett sådant scenario, menar Borge, behåller Chalmers argument sin skarphet gentemot materialismen, *men* detta sker på bekostnad av skepticism mot alla andra personers själsliv, och därmed metafysisk solipsism; jag vet med säkerhet att jag själv har medvetna mentala tillstånd, men resten av alla människor skulle kunna vara zombies (1999, s. 82).

Jag håller med Borge i att *om* vi är rättfärdigade att tro att andra personer än vi själva har medvetna mentala tillstånd i en godtycklig värld så måste vi vara det i alla möjliga världar med samma fysiska konstitution, och vice versa; om vi inte är rättfärdigade att tro det i en enda godtycklig värld så är vi det inte i någon med samma fysiska konstitution. Dilemmat är därmed verkligt. Jag tror däremot inte att vi någonsin kan vara rättfärdigade om andras själsliv, varför zombies med utgångspunkt i Borges artikel inte verkar omöjliga.

Jag tror inte att metafysisk solipsism över huvud taget är särskilt problematiskt; det finns hittills ingen metod för att med säkerhet uppnå kunskap om andra personers själsliv. Vi kan säga att

⁸ Regeln är densamma som introduktion av allkvantifikator (\forall Intro) (Barker-Plummer et al., 2011, s. 576).

det är *sannolikt* att andra personer har ett själsliv eftersom vi kan dra slutsatser om deras beteende och även förutsäga deras beteende genom att dra paralleller till vårt eget själsliv, men vi kan aldrig säga något med säkerhet om huruvida andra personer faktiskt *har* ett själsliv eller hur detta i så fall förhåller sig. Vi kan aldrig sätta oss in i *hur det är* att vara någon annan, och frågan är om detta egentligen spelar någon roll.

Så länge andra människor fyller den funktion i subjektets tillvaro som de gör genom att bete sig på ett visst sätt och vi tillförlitligt kan tolka deras beteende på basis av slutsatser man drar utifrån kunskap om sitt eget själsliv verkar det inte spela någon som helst roll huruvida andra personer är som jag eller om ljuset är släckt inom dem. Om jag en morgon skulle vakna upp och utan min vetskap ha transporterats till zombievärlden, bytt plats med min zombietvilling och därmed vara den enda personen i världen som har medvetna mentala tillstånd, så är det mycket tveksamt om jag någonsin skulle upptäcka detta. Så länge zombierna beter sig exakt som människorna gjorde i den vanliga världen kommer jag inte upptäcka att de saknar medvetna mentala tillstånd, och huruvida de faktiskt har medvetna mentala tillstånd eller ej är därför irrelevant.

Jag tror att dilemmat är verkligt men ser inte problemet med dilemmats ena horn (metafysisk solipsism, om den förstås som jag beskrivit). Jag tror därför inte att det finns någon anledning att tillbakavisa zombies med utgångspunkt i denna artikel.

3.2. VARFÖR ZOMBIES INTE KAN EXISTERA

Robert Kirk vill i en artikel från 1999 visa att zombiekonceptet är logiskt omöjligt därför att det innehåller motsägelser, och därefter förklara varför det är så genom att visa att zombies, precis som vi människor, har perceptuella medvetna mentala tillstånd.⁹

Kirk menar att det är problematiskt att försöka undersöka andra varelser som medvetna observatörer genom ett *det är på ett visst sätt*-kriterium. Anledningen till detta är att förklara någon annans *sätt på vilket det är* skulle innebära att använda begrepp man själv inte kan förstå ur förstapersonsperspektiv, exempelvis ekolokalisering för att förklara hur det är att vara en fladdermus.¹⁰ Man kan inte förklara *hur det är* att vara någon annan, och detta är anledningen till att vi borde överge detta kriterium och snarare använda oss av ett *vad som krävs*-kriterium för vad det är att vara perceptuellt medveten. På så sätt behöver man inte göra någon slags neutrala översättningar eller logiska ekvivalenter för att tillskriva någon medvetna perceptuella upplevelser menar Kirk, utan bara tillräcklig förklaring av vad som är involverat när man har en sådan (1999, s. 4). Begreppen med vilka vi tänker på medvetna mentala tillstånd är mystiska, men så är det inte för vardagliga psykologiska begrepp; för dessa kan vi känna till deras applikation ganska väl, och Kirk vill använda sig av dessa begrepp för att förstå det begrepp som förvirrar oss (1999, s. 5).

Kirk menar att enligt zombieförespråkare är det fenomenella egenskaper – qualia, som utgör skillnaden mellan zombies och vanliga människor. Det verkar enligt Kirk därmed som att om min zombietvilling skulle få tillgång till qualia så skulle den få medvetna mentala tillstånd,

⁹ ”Har perceptuella medvetna mentala tillstånd” är en översättning av ”perceptually conscious”.

¹⁰ ”Medveten observatör” är en översättning av ”conscious observer”.

eftersom den enda skillnaden mellan mig och den är att den inte har qualia (1999, s. 5).¹¹ Problemet är, menar Kirk, att om zombies är möjliga så kan detta inte hända; qualian utgör ingen skillnad för zombien eftersom den bara är ett fysiskt system som inte kan påverkas av icke-fysiska entiteter. För att få tillgång till medvetna mentala tillstånd genom icke-fysisk qualia så måste man kunna påverkas av den, men eftersom zombien per definition inte kan ta till sig qualia påverkas den inte av det. Om allt var mörkt inuti zombien innan den påträffade qualian så är det så även efteråt, menar Kirk; den blir inte medveten av att få tillgång till qualia (1999, s. 6).

Zombieförespråkarna menar enligt Kirk att även om man skulle ta bort all qualia från en människa och göra den till en zombie så består människan trots allt av en speciell förening av en fysisk kropp ("zombiekompanjonen") och ickefysisk qualia. Jag är inte identisk med min kropp utan består av dessa två komponenter. Frågan är, menar Kirk, om den här föreningen, det vill säga människan, kan avgöra skillnaden mellan exempelvis doften av te och doften av kaffe, det vill säga inte bara diskriminera mellan dem (vilket även zombies kan) utan faktiskt *skilja* deras subjektiva karaktärer – deras qualia - åt? För att göra detta krävs enligt Kirk att man kan påverkas av qualia, vilket vi redan vet att zombiekroppen inte kan. Men qualia kan inte heller avgöra sådana skillnader, menar Kirk, som skriver: "*Qualia are, at best, that between which there are differences to be detected; they cannot also be that which does the detecting.*" (1999, s. 7).

Det kan alltså enligt Kirk inte vara så att qualia är vad som avgör skillnaden mellan två olika qualia, eftersom att avgöra skillnad på det här sättet är en kognitiv handling som involverar jämförande och särskiljande (1999, s. 7). Det antas enligt Kirk av zombieförespråkare att inget utöver zombiekroppen och icke-fysisk qualia är involverade i en människa med medvetna mentala tillstånd, och av denna anledning så kan man dra slutsatsen att inte heller en människa med medvetna mentala tillstånd kan upptäcka skillnader mellan qualia. Men det kan vi; de flesta av oss kan faktiskt skilja mellan doften av kaffe och doften av te. Ur detta följer, enligt Kirk, att vi inte är föreningar av den sort som zombieförespråkare menar att vi är, och att zombies i deras mening inte är möjliga (1999, s. 8).

Kirk vill nu visa på vad som krävs för att något ska sägas ha perceptuella medvetna mentala tillstånd, för att ur detta dra slutsatsen att zombies definitivt inte är möjliga. Kirk menar att vad han kallar *baspaketet*; ett komplex av kapaciteter, är centralt när det gäller zombietvillingen.¹² Dessa kapaciteter beskrivs i vanliga psykologiska termer, som måste innehas av varje varelse som kan sägas ha perceptuella medvetna mentala tillstånd. Perception innebär enligt Kirk att lära sig om omgivningen, och består av förändringar i individen som får den att kontrollera sitt beteende på sätt den annars inte skulle ha gjort (1999, s. 9). För att ha perceptuella medvetna mentala tillstånd måste ett system enligt Kirk vara förmöget att:

¹¹ "Qualia" är detsamma som "fenomenella egenskaper".

¹² "Baspaketet" är en översättning av "The Basal Package".

- (i) control its own behaviour
- (ii) acquire information about its environment that it can use
- (iii) retain such information
- (iv) interpret such information
- (v) initiate and modify its behaviour on the basis of incoming and retained information
- (vi) assess its situation
- (vii) make decisions on the basis of retained and incoming information.

Om en enda av dessa kapaciteter är applicerbar på ett givet system så är enligt Kirk alla det; de hänger ihop. Kirk menar att baspaketet karaktäriserar något mer basalt än den sort av intentionalitet vi har, och som Kirk förstår det så har zombietvillingen det här baspaketet. Baspaketet är att betrakta som neutralt och inte något som implicerar medvetna mentala tillstånd; det är inte question-begging att applicera det på zombies, menar Kirk (1999, s. 11).

Baspaketet själv är inte tillräckligt för att fastställa om något har perceptuella medvetna mentala tillstånd, utan det behövs något mer: en specialegenskap, vilken består i att inkommande perceptuell information tvingar sig på zombietvillingen utan att den själv behöver söka upp den.¹³ Att söka upp information själv är något som vi aktivt *gör*, medan medvetna mentala tillstånd bara kommer till oss (1999, s. 12). Att det finns perceptuell information kräver att omgivningen orsakar förändringar hos subjektet. Men för att sådana förändringar ska ske krävs det att subjektet har baspaketet. De måste påverka subjektets beteende, tolkande, bevarande och beslutsfattande. Sådana effekter skulle kanske kunna ske som ett resultat av att subjektet själv söker upp informationen, men då skulle de vara indirekta och försenade, inte direkta och omedelbara, menar Kirk (1999, s. 13).

Kirk tror att baspaketet och specialeenskapen är logiskt nödvändiga och tillräckliga för perceptuella medvetna mentala tillstånd, och att fenomenella zombies har båda. Om zombietvillingen presenteras för två koppar fyllda med brun vätska, luktar på båda och korrekt beskriver den ena som kaffe kan den på något sätt *avgöra skillnaden* mellan doften av kaffe och doften av te (1999, s. 13). Att avgöra skillnad är inte detsamma som att bara reagera olika på olika stimulus, utan när zombietvillingen säger ”mm, kaffe!”, eller låter bli att göra det, när den luktar på kaffe så beror det på vad den *väljer* att göra. Den kan avgöra skillnad mellan vätskorna, men det är inte detta som får den att handla. Hur zombien betar sig beror på hur den *tolkar*, *bedömer* och *fattar beslut* i den givna situationen, tillsammans med dess preferenser, menar Kirk. Den *beaktar* alltså de här effekterna medan den *bestämmer* sig för att rikta sig först mot den ena doften och sen den andra och för att *avgöra* vilken som är vilken (1999, s. 14).

Kirk ser inte hur det skulle kunna vara så att tidigare beskrivna processer skulle vara på *ett annat sätt* för min zombietvilling än för mig; det är på ett sätt för den att lukta på teet och på ett annat sätt att lukta på kaffet, så zombietvillingen är inte bara ett system beskrivet i tredjepersonstermer utan en förstaperson med perceptuella medvetna mentala tillstånd. Det vore

¹³ ”Specialegenskap” är en översättning av ”special feature”.

omöjligt att rikta sig mot lukten av te eller kaffe utan att faktiskt rikta sig mot dem *på riktigt*. Alltså; zombies är enligt Kirk omöjliga eftersom de har perceptuella medvetna mentala tillstånd (1999, s. 14).

Jag tror att Kirk har en viktig poäng i att man aldrig kan beskriva hur det är att vara någon annan; man kan tänka sig eller föreställa sig detta men, som jag tidigare påpekat, aldrig vara riktigt säker på huruvida någon annan har ett själsliv och hur det i så fall förhåller sig. I undersökningar av vad det är att vara fenomenellt medveten tror jag att det är fördelaktigt att istället hålla sig till ett *vad som krävs*-kriterium. I zombiefallet ställer jag mig däremot tveksam; zombies saknar per definition medvetna mentala tillstånd, och att zombien verkar kunna skilja mellan olika qualia, som i kaffe- och te-fallet, beror på att den är kognitivt oskiljbar från mig; inte på att den självständigt skulle fatta egna beslut.

Jag tror vidare att ett centralt problem för Kirk är att ett av hans kriterier för perceptuella medvetna mentala tillstånd är att man behöver vara förmögen att *tolka* information (iv). Zombien kan uppenbart förändra sitt beteende på basis av saker som händer den, men jag tror att tolkningssteget i denna förändringsprocess måste saknas hos zombien. Jag tror att självmedvetande är centralt för tolkningsförmågan, och att zombien per definition måste sakna denna förmåga eftersom den saknar medvetna mentala tillstånd. Att förändra sitt beteende på basis av en tolkning av information man tar in kräver att man kan förhålla sig till sig själv och sin egen roll i världen. Jag tror att förändring av beteende på basis av tolkning, till skillnad från att förändra sitt beteende ”mekaniskt”, involverar en medveten process där man testat omgivningen mot den egna, subjektiva, aspekten, och denna förmåga tror jag att zombien per definition måste sakna.

Zombien kan uppenbart förhålla sig till information på samma sätt som en människa i situationer där man reagerar instinktivt för att t.ex. snabbt fånga något som kastas mot ens ansikte eller omedelbart drar bort handen från en het spisplatta (där medvetenhet om vad som händer inte uppstår i stunden utan först efteråt), det vill säga den kan funktionellt reagera på input, men eftersom zombien saknar medvetna mentala tillstånd kan den inte tolka, reflektera över och dra slutsatser om vad som händer; även om den förändrar sitt beteende på basis av inkommande information så kan den inte vara medveten om varför den gör det.

Dessutom är jag övertygad om att även om man hade kunna säga att zombien är förmögen att tolka på något omedvetet sätt och därmed dra slutsatsen att baspaketet och specialfunktionen faktiskt fanns hos zombien så skulle dessa finnas där enbart därför att zombien är kognitivt oskiljbar från mig. Det skulle fortfarande kunna vara så att det finns något speciellt hos mig som gör att jag utöver detta har medvetna mentala tillstånd, vilket zombien per definition saknar. Vi delar kausal historia, men att min zombietvilling spelar samma roll i zombievärlden som jag gör i den vanliga världen innebär inte att den är förmögen till medvetna processer som att tolka eller välja. Den utför handlingarna men ljuset är, för att ta en användbar metafor i bruk, släckt.

3.3. ZOMBIEDÖDEN

Nigel J.T. Thomas fokuserar i ett konferensbidrag (*Zombie Killer*) från 1998 på zombiens språk, och argumenterar för att oavsett om man beskriver yttranden gjorda av zombier som sanna, falska eller meningslösa innebär detta allvarliga problem för zombieförespråkaren. Thomas menar att zombies är omöjliga därför att språket för zombies, liksom för oss människor, omöjligt kan vara något som bara finns inuti mitt eget huvud utan måste vara något som sker i samverkan med andra människor (eller i zombievärlden: med andra zombies).

Jag och min zombietvilling är kognitivt oskiljbara; de processer som pågår i mig pågår alltid också i min zombietvilling. Eftersom zombierna i Chalmers zombievärld beter sig exakt som sin människotvilling i den vanliga världen så skulle somliga zombies hävda att de har medvetna mentala tillstånd, exempelvis ”jag upplever rött”. När en människa hävdar att den upplever rött är detta sant, och Thomas undersöker hur det förhåller sig när zombietvillingen yttrar detsamma. Thomas menar att oavsett om zombieförespråkaren väljer att påstå att zombieyttrandena är sanna, falska eller meningslösa så leder det till allvarliga problem för teorin om att zombies är möjliga (1998, s. 2).

Är zombieyttrandena falska?

När en person talar falskt förhåller det sig enligt Thomas vanligtvis så att personen antingen ljuger eller misstar sig. Men, menar Thomas, eftersom zombietvillingen är kognitivt oskiljbar från mig och alltid gör exakt samma saker som jag så kan den inte ljuga när den säger att den upplever rött eftersom att ljuga innefattar intentionen att göra det, och om jag inte har denna intention så har inte min zombietvilling det heller (1998, s. 2).

Men om zombien genuint *tror* att den har medvetna mentala tillstånd? Då borde samma kognitiva mekanismer som hos mig vara involverade. Frågan är, enligt Thomas, hur våra trosföreställningar uppkommer; drar jag och min zombietvilling slutsatsen att vi har medvetna mentala tillstånd från vår gemensamma observation av något som i vår värld är reliabelt korrelerat med medvetna mentala tillstånd men som inte är det i zombievärlden?

Zombieförespråkare menar, enligt Thomas, att vi känner till våra egna medvetna mentala tillstånd direkt. Thomas menar dock att *någon* kognitiv process måste ta mig från faktumet att jag har medvetna mentala tillstånd till trosföreställningen att jag har det. Eftersom zombien och jag är kognitivt oskiljbara äger samma process rum hos zombien. Processen leder hos zombien, till skillnad från hos mig, till en misstagen trosföreställning (1998, s. 3). Oberoende av om man faktiskt *har* medvetna mentala tillstånd eller inte leder processen alltså till trosföreställningen att man har det. Thomas menar att om processen kan förhålla sig på detta sätt så är antagandet om att vi själva inte är zombies grundlöst, och om vi skulle kunna vara zombies så har zombieargumentet förlorat sitt bett (1998, s. 3).

Är zombieyttrandena sanna?

Thomas tar upp ett synsätt han menar finns hos Moody (1994), som säger att man snarare borde kalla det zombies refererar till som medvetna mentala tillstånd, det vill säga när de yttrar sådant som ”jag upplever rött”, för *medvetna mentala tillstånd*⁷. Det kan härmed verka som att vi kan

göra en koherent tolkning av medvetna mentala tillstånd^z och zombiens användande av det, men Thomas menar att det faktum att vi över huvud taget kan göra en distinktion mellan *upplevelsen* av rött å ena sidan och *upplevelsen*^z av rött å andra bär med sig inkohärens (1998, s. 4). Om vi människor kan uppfatta och uttrycka den här skillnaden så kan zombietvillingarna det också. Zombietvillingen till en person som hävdar att den har medvetna mentala tillstånd kommer enligt Thomas alltid hävda exakt samma sak; att den upplever *rött*, inte att den upplever *rött*^z (1998, s. 4). Thomas menar därför att det inte går att göra någon distinktion mellan begrepp för oss och begrepp för zombies, och även om det skulle gå så skulle zombierna ändå alltid hävda att det är medvetna mentala tillstånd de har, och detta är inte sant (1998, s. 4).

Är zombietvillingarna meningslösa?

Det verkar enligt Thomas vid första åsyn plausibelt att yttranden om medvetna mentala tillstånd gjorda av zombies varken är sanna eller falska, eftersom ord som 'sensation' eller 'medvetna mentala tillstånd' saknar referens i zombievärlden. Men Thomas exemplifierar med att icke-refererande ord som "jabberwock" inte nödvändigtvis är meningslösa. Vi kan veta att det är ett slags monster i *Alice i Underlandet* och därför förstå satser som innehåller det här ordet. Jag kan till exempel svara "nej, det är inte sant" om någon påstår att hen har en jabberwock i fickan. På liknande sätt måste 'medvetna mentala tillstånd' enligt Thomas ha mening för en zombie. Och om en zombie över huvudet kan anta och veta saker, menar Thomas, så kommer dess hävdande att den har medvetna mentala tillstånd på sätt och vis vara meningsfulla nog för att föra oss tillbaka till redan diskuterad problematik om begrepp för oss och begrepp för zombierna (1998, s.4).

Men kanske kan en zombie inte veta och anta saker. Kanske är inget som helst som en zombie säger eller tänker meningsfullt. Thomas menar att om vi, i enighet med Searle (1992), accepterar att våra tankars meningsfullhet (vilken ordens meningsfullhet sedan bygger på) *är beroende av* våra medvetna mentala tillstånd så verkar det som att vi måste erkänna att zombiers talakter, och framförallt zombiers hävdande att de har medvetna mentala tillstånd, är meningslösa. Att de på detta sätt är meningslösa är därmed den enda koherenta beskrivningen av kognitiv signifikans hos zombietvillingen, menar Thomas (1998, s. 4).

Om zombieförespråkaren accepterar att zombietvillingen om fenomenella tillstånd är meningslös så är den dock bunden till att "när zombien yttrar samma sak som jag i en exakt likadan miljö och de andra zombierna reagerar på yttrandena på samma sätt som andra människor reagerar på mitt yttrande, så är orden i mitt fall meningsfulla men i zombiens fall meningslösa", menar Thomas. Detta går emot vad det är för ett språk att vara meningsfullt. Enligt Thomas bygger zombiekonceptet på att vi tänker på mentala upplevelser som något som uppkommer privat, inne i våra huvuden. Kanske är detta rätt sätt att tänka på medvetna mentala upplevelser, men det är enligt Thomas inte rätt sätt att tänka på språk, eftersom dessa förvisso delvis beror på vad som händer i huvudet på mig, men också på vad som sker i andra människors huvuden och vad som sker människor emellan.

Språk är en social aktivitet, inte något internt och mentalt, och detta verkar enligt Thomas gälla såväl i zombievärlden som i vår värld, eftersom zombien och dess zombievänner använder sig av språket på samma sätt som vi. Thomas frågar sig hur vi, när det förhåller sig på detta sätt,

kan säga att zombiernas yttranden inte är meningsfulla för dem. Det verkar enligt Thomas också felaktigt att förneka att zombieantagandena har sanningsvärde. När jag upplyser en annan person om något håller personen i regel vad jag säger för sant (om personen inte har skäl att tro att jag ljugar eller har fel), och våra zombietvillingar förhåller sig i en ekvivalent position, menar Thomas, och det är just detta som gör att zombies enligt Thomas är omöjliga:

”Normally (and justifiably) your twin would respond as if reliable information had been conveyed: i.e as if my twin had told the truth. Can the zombiphile reasonably deny that it is right to do so? If not, the case for zombies is lost.” (Thomas, 1998, s. 5).

Thomas menar att om zombies är omöjliga så har vi rätt att dra slutsatsen att komplex social och lingvistisk interaktion är omöjlig utan medvetna mentala tillstånd. Därmed behöver vi inte ge upp Searles princip om medvetna mentala tillstånd och den intrinsikala intentionaliteten som underliggjer lingvistisk mening. Vi kan till och med gå längre än Searle, menar Thomas, och säga att kvalitativa medvetna mentala tillstånd och intentionalitet snarast bör ses som olika aspekter av samma sak (1998, s. 5).

Jag tror att Thomas kritik är väldigt missriktad. Som jag förstår Chalmers är inte sanningsvärdet på vad zombierna säger något direkt relevant, utan det är faktumet *att* de säger och gör exakt samma saker som vi men att vi ändå kan *tänka oss* att de inte har några medvetna mentala tillstånd som är viktigt. Jag ser inte hur det kan vara så att zombies är omöjliga för att språklig mening har sitt upphov i medvetna mentala tillstånd; zombies är bara tänkbara varelser, Chalmers huvudpoäng är *att vi kan tänka oss* dem, och det verkar för mig i det här sammanhanget irrelevant hur språklig mening uppstår. Om språklig mening har sitt ursprung i medvetna mentala tillstånd saknar zombieyttranden helt enkelt referens även om zombien yttrar samma saker som jag, men om vi ändå kan *tänka oss* att de gör det så uppstår ingen inkoherens.

När Thomas går igenom zombieuttryck missar han ett alternativ. Zombien behöver inte nödvändigtvis ljuga eller missta sig; dess yttranden om medvetna mentala tillstånd kan vara falska ändå därför att de är kontextberoende. I vår värld finns medvetna mentala tillstånd och yttrandet ”jag upplever rött” är sant, men i zombievärlden finns inga medvetna mentala tillstånd och därför är samma sats yttrad där falsk. När en zombie säger att den upplever rött så är det fel, den har inga medvetna mentala tillstånd och kan därmed inte ha någon rödupplevelse. Jabberwock-exemplet verkar inte fånga vad det är som händer när en zombie yttrar sådant som ”jag upplever rött” därför att vi i regel inte tror att jabberwockar finns på riktigt när vi talar om dem. Genom att, i någon mening, veta vad en jabberwock är förhåller vi oss helt enkelt till utsagor om dem på ett annat sätt än zombies förhåller sig till utsagor om medvetna mentala tillstånd.

Det finns, förutom sådana utsagor som handlar direkt om en upplevelse och därmed är falska, en annan typ av satser om medvetna mentala tillstånd som zombien skulle kunna yttra, nämligen satser av typen ”min röda upplevelse gör mig glad”. Dessa satser är meningsfulla på samma sätt som om man talar om ”den nuvarande kungen av Frankrike är skallig”, men de är båda felaktiga därför att zombiens röda upplevelse, liksom den nuvarande kungen av Frankrike, inte

existerar.¹⁴ När en zombie på detta sätt talar om sin upplevelse har den helt enkelt fel; satsen ”min röda upplevelse gör mig glad” yttrad i zombievärlden är meningsfull men saknar sanningsvärde eftersom den inte refererar till något.

Ur detta sluter jag mig till att zombieyttrandena trots Thomas argumentation måste vara falska eller sakna sanningsvärde. Att missta sig behöver inte ligga i de interna kognitiva processerna, utan kan i hög grad ha att göra med hur den externa världen förhåller sig. Zombien och jag kan vara kognitivt oskiljaktiga trots att zombien talar falskt och jag sant, eftersom det faktum att jag har medvetna mentala tillstånd och inte zombien ligger i hur våra respektive *världar* förhåller sig snarare än i våra egna intentioner. Zombien behöver därmed inte vara kognitivt särskild från mig för att ha fel. Att mina utsagor om medvetna mentala tillstånd har sanningsvärde medan zombiens saknar det beror på det faktum att mina utsagor om detta refererar till mina medvetna mentala tillstånd medan zombiens utsagor om samma saker saknar referens.

Dessutom frågar jag mig om zombien verkligen är att betrakta som en kompetent språkanvändare. Den kan inte självständigt konstruera satser, utan anledningen till att den lyckas är att den är kognitivt oskiljbar från mig, som är en kompetent talare. Att tala om att zombies som *språkanvändare* verkar för mig felaktigt i grunden; de gör saker, de pratar, klättrar i träd, skriver, slår sig, tittar på tv och så vidare men allt de gör verkar sakna mening eftersom det bara är ”mörkt” inuti dem. Eftersom de producerar sina yttrande utan att kunna reflektera över vad de gör därför att de saknar medvetna mentala tillstånd *använder* zombierna inte språket på samma sätt som vi. Att språk är en social aktivitet för människor behöver inte betyda att detsamma gäller för zombies.

4. KAN MÖJLIGHET HÄRLEDAS FRÅN TÄNKBARHET?

Jag har nu gjort ett försök att visa att Chalmers zombieargument verkar stå emot några av de försök att vederlägga det som har gjorts. Det finns dock en särskild sorts invändningar som jag tror är mer lyckade, och gemensamt för dem är att de siktar in sig på Chalmers tes om att möjlighet kan härledas från tänkbarhet. Innan jag går in på dessa invändningar ska jag gå igenom några saker som är centrala för denna aspekt av Chalmers argument.

4.1. STANDARDINVÄNDNINGEN

Att möjlighet inte kan härledas ur tänkbarhet är en mycket vanlig invändning mot tänkbarhetsargument. Jag kommer hädanefter referera till tesen att möjlighet följer ur tänkbarhet som TM-tesen, efter inspiration från Keith Frankish (2007) som kallar den för *The CP-Thesis*.¹⁵¹⁶ Chalmers beskriver denna invändning i *The Conscious Mind* (1996), och hävdar att det förmodligen är den allra vanligaste invändningen från materialister som vill visa att det

¹⁴ Att satsen ”Kungen av Frankrike är skallig” varken är sann eller falsk men ändå meningsfull är ett exempel lånat av Strawson (1950).

¹⁵ C står för Conceivability och P för Possibility, och Frankish beskriver den som ”*the thesis that if a situation is ideally conceivable then it is metaphysically possible*” (2007, s. 652). Jag låter dock TM-tesen stå för härledningar ur vilken sorts tänkbarhet som helst till möjlighet.

¹⁶ TM-tesen, det vill säga att möjlighet kan härledas ur tänkbarhet, är ej att förväxla med dess motsats; att tänkbarhet kan härledas ur möjlighet (MT-tesen), som av uppenbara skäl är ogiltig. Alla möjligheter är inte tänkbara.

inte finns någon härledning mellan fysiska och fenomenella fakta. I en artikel av Anthony Brueckner (2001) omnämns denna invändning till och med som *standardinvändningen* mot tänkbarhetsargument.

Enligt Chalmers ackompanjeras standardinvändningen ofta av att zombies, i ljuset av Saul Kripkes *Naming and Necessity* (1980), endast är logiskt, och inte metafysiskt möjliga. Begreppslig koherens kan sägas vara tillräckligt för logisk möjlighet, men inte för metafysisk. På samma sätt som det är logiskt, men inte metafysiskt, möjligt att vatten inte är H₂O är det naturligt att anta att zombies är logiskt, men inte metafysiskt möjliga (1996, s. 131).

4.2. CHALMERS SVAR PÅ STANDARDINVÄNDNINGEN

Chalmers svar på standardinvändningen vilar på hans tvådimensionella semantik. I denna sektion kommer jag följaktligen presentera den tvådimensionella semantiken som Chalmers beskriver i *The Conscious Mind* (1996), samt det svar Chalmers ger på standardinvändningen i samma bok.

4.2.1. TVÅDIMENSIONELL SEMANTIK

Det finns vissa begreppsliga sanningar som vi inte kan få kunskap om *a priori*. Chalmers menar i enighet med Kripke (1972) att ”vatten är H₂O” är exempel på en sådan sanning. Enligt Kripke avgör det faktum att vatten är H₂O i den aktuella världen att vatten är H₂O i alla möjliga världar, och att vatten är H₂O är därmed en nödvändig sanning. För att lösa detta introducerar Chalmers sitt tvådimensionella ramverk för mening och nödvändighet (1996, s. 56). Chalmers menar att det finns två intensioner associerade till varje given term; den *primära* intensionen, som bestäms av hur dess referens fixeras i den *faktiska* världen, beroende på hur den förhåller sig, samt den *sekundära* intensionen, som bestäms av hur dess referens fixeras i *kontrafaktiska* världar, givet att referensen i den faktiska världen redan är fixerad (1996, s. 57).

De två intensionerna kan enklast demonstreras genom att betänka termen ”vatten”. Den *primära* intensionen pekar ut det som skulle vara termens referent i en given värld om denna värld hade varit den faktiska. Om hav och sjöar i denna värld består av XYZ så skulle ”vatten” referera till XYZ, och om de består av H₂O är detta vad termen ’vatten’ skulle referera till. ”Vatten” pekar ut ”den klara drickbara vätskan som finns i hav och sjöar” eller ”sådant som är av vattenkaraktär” oberoende om det består av XYZ eller H₂O; i en värld där denna beskrivning är applicerbar på XYZ pekar ”vatten” ut XYZ, och i en värld där den är applicerbar på H₂O pekar ”vatten” ut H₂O.¹⁷ Men givet att ”vatten” visar sig referera till H₂O i den aktuella världen så pekar den *sekundära* intensionen av ”vatten” ut ”vatten” i varje kontrafaktisk värld, menar Chalmers; så ”vatten” pekar ut H₂O i varje möjlig värld (1996, s. 57). Referensen för ”vatten” i den faktiska världen fixeras genom att det pekar ut sådant som är av vattenkaraktär i den faktiska världen, och om vatten är H₂O i den faktiska världen så är vatten H₂O i alla möjliga världar. I en värld där den klara drickbara vätskan är XYZ är detta enligt Chalmers inte vatten, utan bara något av vattenkaraktär; den sekundära intensionen pekar ut vatten i alla världar och därmed H₂O i alla världar. Vad som är vatten avgörs inte *a priori*, men har enligt Chalmers

¹⁷ ”Sådant som är av vattenkaraktär” är en översättning av ”watery stuff”.

ändå en nära relation till den primära intensionen; sekundär intension determineras genom att man utvärderar primärintensionen i den faktiska världen och *rigidifierar* utvärderingen så att samma sak pekats ut i alla möjliga världar (1996, s. 59).

Synen på dessa två meningskomponenter liknar enligt Chalmers Kaplans (1979) distinktion mellan *ytrandekontext* och *utvärderingsomständigheter*.¹⁸ När vi tänker oss en kontrafaktisk värld behåller vi den aktuella världen som ytrandekontext och använder *W* som omständighet för utvärdering. Chalmers menar att om vi yttrar ”det finns vatten i havet” i den aktuella världen och *utvärderar* det i XYZ-världen refererar ”vatten” till H₂O och yttrandet är falskt. Om vi däremot betraktar *W* som aktuell så tänker vi på den som en potentiell ytrandekontext och beaktar hur saker skulle förhålla sig om uttryckets kontext skulle visa sig vara *W*. Om kontexten för mitt yttrande ”det finns vatten i havet” hade varit XYZ-världen så skulle påståendet vara sant när det utvärderas i den världen. Chalmers menar att primärintensionerna därmed är nära besläktade med vad Kaplan kallar för en terms *karaktär*, medan sekundärintensionen motsvarar vad han kallar för en terms *inhåll* (1996, s. 60).

Chalmers menar i enighet med Quine (1969) att det finns viss asymmetri i att ytrandekontexten, men inte utvärderingsomständigheterna, är en så kallad *centrerad* möjlig värld; ett ordnat par som består av en värld och ett *centrum* som representerar en synvinkel i världen för den agent som använder termen i fråga. Centrat består alltså av åtminstone en ”markerad” individ och tidpunkt, och är enligt Chalmers nödvändigt för att kunna fånga det faktum att en term som ”vatten” pekar ut olika saker för mig och för min tvilling på Tvillingjorden, trots att vi lever i samma universum. Det enda som skiljer sig åt är vår position i universum, och det är denna skillnad som är relevant för fixeringen av referens, menar Chalmers (1996, s. 60).

Chalmers menar att detta fenomen uppkommer på ett särskilt uppenbart sätt när vi behandlar indexikala termer såsom ”jag”, vars referens beror på vem som använder termen och inte endast tillståndet i världen. Den primära intensionen av ordet ”jag” pekar ut individen i centrum av en centrerad värld (den sekundära intensionen av ”jag” yttrar av David Chalmers pekar ut David Chalmers i varje möjlig värld och yttrar av mig pekar det ut Julia Sjödal i varje möjlig värld). Hos ord som ”vatten” är det indexikala elementet inte lika uppenbart. Detta element kan analyseras som ”*dhen* (klara, drickbara vätskan i vår omgivning)”, och det är enligt Chalmers detta indexikala element som kräver att primära intensioner beror på centrerade världar (1996, s. 61).¹⁹ Men, menar Chalmers; när referensen är fixerad i den faktiska världen behövs inget centrum för att utvärdera referens i kontrafaktiska världar, och därför kan utvärderingsomständigheterna representeras av en enkel möjlig värld utan ett centrum (1996, s. 61).

Både primära och sekundära intensioner är enligt Chalmers möjliga kandidater för ett begrepps ”mening”, och det finns ingen poäng med att välja ett av dem som *meningen*, utan primära och sekundära intensioner är snarare *a priori*- och *a posteriori*-aspekter av mening (1996, s. 62). Därmed menar Chalmers att båda intensionerna för oss till en särskild sorts begreppsliga

¹⁸ ”Ytrandekontext” och ”utvärderingsomständigheter” är översättningar av ”context of utterance” och ”circumstances of evaluation”.

¹⁹ ”Dhen” är en försvenskning av vad Kaplan kallar för ”dthat”.

sanningar. Primärintensionen berör *a priori* sanningar, såsom ”vatten är sådant som är av vattenkaraktär”, som är sant oavsett hur den aktuella världen är. Sekundära intensioner berör sådana sanningar som gäller i alla kontrafaktiska världar, såsom ”vatten är H₂O”. Chalmers menar att båda intensionerna är sanna i värde av sin mening; det handlar bara om olika sorts mening.

De primära intensionerna i ”vatten” och ”H₂O” är olika, så vi kan inte veta *a priori* att vatten är H₂O; den primära positionen är enligt Chalmers inte nödvändig (1996, s. 64). Att de sekundära intensionerna sammanfaller är däremot nödvändigt; att ”vatten är H₂O” är sann i alla möjliga världar när den utvärderas i förhållande till sekundära intensioner.

Chalmers menar att den primära intensionen är det mest centrala för de egna syftena. Om någon säger ”förklara vad vatten är” så behöver vi inte veta något om H₂O för att kunna göra detta, utan vad som efterfrågas är en förklaring av den klara drickbara vätskan som finns i hav och sjöar. Den primära intensionen är enligt Chalmers oberoende av empiriska faktorer; den specificerar hur referens beror på hur den externa världen förhåller sig, och beror inte själv på hur världen förhåller sig (1996, s. 57).

4.2.2. CHALMERS LÖSNINGSFÖRSLAG

Chalmers menar att *a posteriori*-nödvändighet inte klarar av att bära den börda som standardinvändningen kräver, och att den är en felaktig invändning (1996, s. 131). Logisk möjlighet är den möjliga sanningen hos ett påstående när det utvärderas i termer av involverade primära intensioner, menar Chalmers. De primära intensionerna hos ”vatten” och ”H₂O” är olika, så det är logiskt möjligt att vatten inte är H₂O. Metafysisk möjlighet däremot, handlar enligt Chalmers om ett påståendes möjliga sanning när det utvärderas i termer av dess sekundära intensioner. Eftersom de sekundära intensionerna hos vatten och H₂O är samma så är det enligt Chalmers metafysiskt nödvändigt att vatten är H₂O (1996, s. 132).

Standardinvändningen handlar därför enligt Chalmers om att han själv, genom att använda sig av ett tänkbarhetsargument, endast har visat att zombievärlden är möjlig genom att använda sig av primära intensioner, och inte de mer lämpliga sekundära; även om de primära intensionerna av fenomenella begrepp eventuellt skiljer sig från fysiska begrepp så kanske dess sekundära intensioner är samma. I så fall pekar fysiska och mentala begrepp ut samma egenskaper *a posteriori*.

Chalmers menar dock att invändningen inte påverkar zombieargumentet därför att detta argument fokuserar på primära intensioner och ignorerar sekundära. Oavsett om de primära och sekundära intensionerna sammanfaller så determinerar den primära intensionen enligt Chalmers en egenskap för objekt i möjliga världar (1996, s. 132). Egenskapen att vara något vattenaktigt är en egenskap, även om det inte är samma egenskap som att vara H₂O. Om vi kan visa att det finns möjliga världar som är identiska med vår men där egenskapen den primära intensionen introducerar inte existerar så följer enligt Chalmers dualism.

Chalmers menar att detta är precis vad som sker med medvetna mentala tillstånd i zombieargumentet; Chalmers har visat på att det finns världar som är exakt likadana som vår men där medvetna mentala tillstånd inte finns, enligt dess primära intension. Olikheten i dessa

världar är enligt Chalmers tillräcklig för att visa att det finns egenskaper i vår värld utöver de fysiska. Denna skillnad kan visas oberoende av *a posteriori*-faktorer, och därmed är enligt Chalmers betänkligheter om *a posteriori*-nödvändighet irrelevanta (1996, s. 133).

5. KRITIK AV TM-TESEN

Jag har nu presenterat den så kallade standardinvändningen mot modala tänkbarhetsargument samt Chalmers försök till lösning av denna på basis av hans tvådimensionella semantik. Följande avsnitt innehåller referat av fyra artiklar som på olika sätt samtliga berör TM-tesen. Katalin Balog hänvisar till betänkligheter av zombies yttranden av fenomenella termer och Anthony Brueckner menar att Chalmers lösningsförslag av TM-invändningen inte rättfärdigar zombies möjlighet, medan Sara Worley hävdar att Chalmers argument inte alls är tillgänglig för standardinvändningen men att det finns andra problem med den. Keith Frankish menar att zombieargumentet kan vederläggas genom att man, givet TM-principen, lika enkelt som att konstruera ett zombieargument kan konstruera ett anti-zombieargument.

5.1. TÄNKBARHET, MÖJLIGHET OCH KROPP-SJÄL-PROBLEMET

Katalin Balog vill i en artikel från 1999 visa att tänkbarhetsargument inte är lyckade därför att vad de säger inte har att göra med medvetna mentala tillstånd i sig (det vill säga *hur det är*-egenskapen), utan snarare med fenomenella begrepp.

Balog riktar in sig tänkbarhetsargumenten från David Chalmers och Frank Jackson, och försöker visa att principen de förlitar sig på för att binda samman tänkbarhet och möjlighet är felaktig. Balog följer Jacksons framställning av argumentet, men hennes rekonstruktion är menad att tillskrivas dem båda (1999, s. 502). Enligt Balog så är det för Jackson och Chalmers centralt att fysikalism är ekvivalent med påståendet att varje sanning om vår värld nödvändiggörs av ett fysiskt påstående som ger den fulla fysiska beskrivningen av världen, och är sann i alla och endast de fysiska kopiorna av vår värld. För detta syfte använder Balog följande definition:

För vilket sant påstående T som helst: $\Box (K \supset T)$ ²⁰

I vilken K står för alla fysiska sanningar om världen och T är en sanning.

Nödvändigheten av detta är inte bara fakta utan behöver förklaras, och förklaringen måste vara begreppslig, menar Balog. Om fysikalismen är sann så är ' $K \supset T$ ' inte bara metafysiskt nödvändig utan också en *a priori*, begreppslig sanning, det vill säga om fysikalismen är sann så är alla sanningar *a priori* härledningsbara från den fullkomliga fysiska beskrivningen av världen.²¹ Detta kallar Balog för *a priori-härledningstesen*, och menar att denna är huvudpremiss i Jacksons och Chalmers argument som länken mellan tänkbarhet och möjlighet (Balog, 1999, s. 505). Många nödvändigheter av formen ' $K \supset T$ ' stämmer väl överens med denna tes, och enligt Jackson och Chalmers, menar Balog, finns det inte någon anledning att

²⁰ $\Box (K \supset T)$ läses som *det är nödvändigt att T följer av K* (\supset betecknar här materiell implikation).

²¹ En begreppslig sanning är sann per definition.

anta att det finns undantag från den. Balog menar dock att det visst finns undantag från *a priori*-härledningstesens (1999, s. 508).

Om *a priori*-härledningstesens är sann innebär detta problem för fysikalismen genom att det i så fall finns fenomenella egenskaper i världen. Balog visar detta genom ett exempel med fenomenet smärta. Utsagan 'K \supset vatten täcker 60 % av jordens yta' bestäms av den begreppsliga sanningen "vatten är en klar, luktfri vätska". Sanningen i fenomenella utsagor, såsom 'K \supset x känner smärta', beror däremot på någon begreppslig sanning som kopplar samman 'smärta' med en icke-fenomenell beskrivning som är *a priori*-tillräcklig för 'känner smärta'. En sådan finns inte. Termer som 'smärta' menar Balog är direkta; vi använder dem inte för att beteckna beteende eller något annat som är distinkt från själva erfarenheten; 'smärta' refererar direkt till smärtan själv. Men ur *a priori*-härledningstesens följer att om 'x känner smärta' inte kan härledas *a priori* från K så är 'K \supset x känner smärta' falsk och därmed gäller att om 'x känner smärta' är sann så är fysikalismen falsk. Balog ställer upp Jacksons och Chalmers argument på följande sätt:

- (1) If physicalism is true, then for any true T, statements of the form K \supset T are conceptual truths.
- (2) There are some true statements Q to the effect that phenomenal conscious experience occurs (eliminativism about phenomenal experience is false).
- (3) If Q is a phenomenal statement, then 'K \supset Q' is not a conceptual truth.

So

- (4) Physicalism is false

Balog menar att det kan visas att detta argument för dualism är osunt, och introducerar nu sitt motargument: zombieverlägningen. Med detta har Balog för avsikt att visa att Jacksons och Chalmers tänkbarhetsargument är självunderminerande; man kan med hjälp av några, enligt Balog, godtyckliga tilläggspremissar härleda en kontradiktion ur det. Balog vill visa att själva antagandet att det finns en möjlig värld som är precis som vår men utan fenomenella egenskaper underminerar argumentet som leder till att en sådan möjlig värld skulle kunna existera.

I zombievärlden finns zombie-Jackson, som är precis som Jackson men som saknar fenomenella tillstånd. Professor zombie-Jackson ger en serie seminarier i zombie-Oxford (precis som Jackson gör i Oxford) och argumenterar för *a priori*-härledningstesens. Enligt Balog har zombie-Jackson intentionala tillstånd – när han talar är orden inte meningslösa, utan zombie-Jacksons intentionala tillstånd är identiska med den vanliga Jacksons, bortsett från dem som involverar fenomenella koncept. De utsagor som verkar handla om fenomenella koncept refererar för zombie-Jackson istället till sakförhållanden i zombievärlden. Zombie-Jacksons argument är lika meningsfulla som Jacksons men inte identiskt med dem – vissa ord har annan mening. Dessa ord markerar Balog med ett +, till exempel 'Smärta+'.

Zombie-Jacksons argument blir som följer:

(1*) If physicalism is true, then for any true T, statements of the form $K \supset T$ are conceptual truths.

(2*) There are some true statements Q^+ to the effect that a phenomenal state occurs (eliminativism about phenomenal states is false).

(3*) If Q^+ is a phenomenal statement, then ' $K \supset Q^+$ ' is not a conceptual truth.

So

(4*) Physicalism is false.

Balog menar att om en premiss är sann i Jacksons fall så är den också sann i zombie-Jacksons fall. Vi vet att den dualistiska slutsatsen av zombie-Jacksons argument är falsk i zombievärlden. Som en konsekvens av detta vet vi att zombie-Jacksons argument inte kan vara sunt. Givet att argumentet är meningsfullt är det giltigt, så en av premisserna måste vara falska menar Balog. Härav följer enligt Balog att en av Jacksons premisser också måste vara falska. Detta är Balogs försök till en *reductio* av Jacksons argument; det är alltså enligt Balog möjligt att få en kontradiktion från Jacksons premisser tillsammans med några tilläggsantaganden (1999, s. 513). Balog menar att givet dessa antaganden så följer premiss (1*)-(3*) från originalpremisserna. Men i så fall följer dualistiska slutsatsen från zombie-Jacksons argument också, vilken motsäger konsekvensen av Jacksons och Chalmers argument; att zombievärlden existerar och är fysikalistisk. Balogs tilläggsantagandena är dessa:

Assumption 1: Jackson and zombie-Jackson share most of their intentional states except those involving phenomenal concepts.

Assumption 2: Those concepts of zombie-Jackson that correspond to Jackson's phenomenal concepts will refer in the zombie to some (physical) state of the zombie.

Assumption 3: A priority for thoughts supervenes on the conceptual roles of their constituent (and related) concepts.

Zombie-Jackson och Jacksons argument är enligt Balog nu identiska. Genom antagande 1 menar de samma saker med sina ord förutom när fenomenella termer är involverade. Detta innebär att premiss (1*) på zombiespråket uttrycker *a priori*-härledningstesens, som om den är sann är nödvändigt sann; om den är sann i den vanliga världen så är den sann i zombievärlden också. Genom antagande 2 får vi premiss (2*), och givet antagande 3 så har premiss (3*) i zombie-Jacksons argument enligt Balog lika stor sannolikhet att vara sann som den vanlige Jacksons, trots att ' $K \supset Q^+$ ' har annan mening än ' $K \supset Q$ ' (1999, s. 514). Att de är lika sannolika beror på att fenomenella koncept+ och fenomenella koncept har parallella begreppsliga roller. Båda saknar länk till fysiska, funktionella eller beteendemässiga begrepp för att grunda *a prioritet* av ' $K \supset Q^+$ '.

Genom antagande 3 supervenierar enligt Balog *a prioritet*, eller begreppslig nödvändighet, på de begreppsliga rollerna hos relevanta koncept. Detta innebär, och här kommer själva poängen i Balogs zombievederläggning, att om ' $K \supset Q$ ' inte är härledningsbar från begreppsliga sanningar givet att Q saknar tillräckliga begreppsliga länkar till fysiska, funktionella och beteendemässiga begrepp så är inte heller ' $K \supset Q^+$ ' härledningsbar från

begreppsliga sanningar eftersom Q+ också kommer sakna de rätta funktionella banden (1999, s. 521).

Jag är skeptisk till Balogs tilläggsantagande nummer 2, där hon hävdar att zombiernas utsagor refererar till fysiska tillstånd i hjärnan. Balog erkänner i artikeln att tilläggsantagande 2 skulle kunna vara falskt om 'smärta+' refererar till icke-fysisk fenomenell smärta eller om 'smärta+' inte refererar till något alls, och jag tror att utsagor om fenomenella tillstånd i zombievärlden snarare saknar referens än att de skulle referera till något fysiskt hos zombien (1999, s. 517). Utsagor om fenomenella termer i zombievärlden kan enligt Balog inte sakna referens, eftersom de då skulle sakna sanningsvärde, men jag tror, som jag beskrev i den tidigare diskussionen av zombieyttranden, att det är just det de gör.

Vidare menar Balog att zombiernas fenomenella+-yttranden och icke-fenomenella+-yttranden har begripliga kopplingar. När zombien till exempel yttrar "sist jag hade tandvärk+ så gick jag till tandläkaren" så verkar detta riktigt, eftersom varje gång zombie-Jackson säger att han har ont+ är han i ett hjärntillstånd eller funktionellt tillstånd som är pålitligt korrelerat med hans term 'smärta+'. Enligt Balog är den naturliga kandidaten för referens av zombie-Jacksons term 'smärta+' samma fysiska tillstånd i hjärnan eller funktionella tillstånd som hos den vanliga Jackson. Detta innebär att närhelst Jacksons påstående 'jag har ont' är sant så är zombie-Jacksons påstående 'jag har ont+' också sant, om han befinner sig i samma hjärn- eller funktionella tillstånd som Jackson (1999, s. 519). Jag är tveksam till detta. Även om det stämmer att Jackson och zombie-Jackson förvisso befinner sig i samma fysiska eller funktionella tillstånd när de yttrar samma saker så är det inte detta gemensamma tillstånd som *är* det medvetna mentala tillståndet som Jacksons term refererar till. Jacksons term refererar till något utöver detta, som zombien inte kan få tillgång till. Utsagor om medvetna mentala tillstånd kan omöjligt referera till något fysiskt eller funktionellt eftersom de, som Balog själv säger när hon exemplifierar med termen smärta, verkar referera direkt till själva tillståndet.

Hade man på något sätt kunnat göra Balogs poäng utan tilläggsantagande nummer 2 så tror jag att man skulle ha ett mycket starkt argument, men med detta antagande faller hela argumentet därför att zombieyttranden per definition måste sakna referens och sanningsvärde. Vad Jacksons utsaga om ett medvetet mentalt tillstånd refererar till finns helt enkelt inte där för hans zombietvilling när denne yttrar samma sak. Zombieyttranden om fenomenella tillstånd måste vara antingen falska eller sakna sanningsvärde, just därför att vad utsagor om medvetna mentala tillstånd refererar till inte existerar i zombievärlden.

Jag tror däremot att en viktig sak Balog belyser vad gäller zombieyttrandena är att de fyller samma funktionella roll för vanliga människor som för zombies. Zombiens yttranden av fenomenella termer är *meningsfulla* därför att de används på samma sätt som för oss, som i Balogs exempel med tandvärken. Yttrandet "jag har tandvärk+" är meningsfullt, inte för att det refererar till ett fysiskt eller funktionellt tillstånd, utan för att det används på rätt sätt. Zombiens yttranden av fenomenella termer har samma kausala historia som fenomenella termer yttrade av dess människotvilling, och kan på grund av detta sättas i ett meningsfullt sammanhang.

5.2. CHALMERS TÄNKBARHETSARGUMENT FÖR DUALISM

I en artikel från 2001 argumenterar Anthony Brueckner för att Chalmers inte lyckas bortförklara standardinvändningen. Brueckner går igenom Chalmers svar på den, och menar att även efter Chalmers förklaringsförsök återstår endast zombiernas tänkbarhet; inte deras möjlighet, och tänkbarhet är inte tillräcklig för att vederlägga den materialistiska superveniens-tesen.

För att förstå Chalmers svar på standardinvändningen måste vi enligt Brueckner beakta hur tvådimensionssemantiken hanterar satsen '(P & ~C)', det vill säga: 'det fysiska som det är & inga medvetna mentala tillstånd', alltså zombies. Enligt Brueckner menar Chalmers att referensfixeraren för 'medvetna mentala tillstånd' (c-begreppet) är 'har en fenomenell aspekt'.²² Så givet en möjlig värld har den primära intensionen hos 'medvetna mentala tillstånd' en extension bestående av de tillstånd i den världen som har fenomenell aspekt. Dessa tillstånd är förmodligen själva medvetna mentala tillstånd, och den sekundära intensionen är därmed samma som den primära; den sekundära intensionen leder till en extension som består av medvetna mentala tillstånd.

För begreppet 'att vara fysisk på det sätt som mänskliga kroppar är' (p-begreppet) så tror Brueckner att Chalmers skulle mena att referensfixerare i den faktiska världen refererar 'genom en extrinsikal karakterisering' av den relevanta entiteten. Ett exempel är 'elektron', som refererar till 'entiteten som spelar elektronrollen'. Därmed, menar Brueckner, kan vi anta att p-begreppets primära intension leder till extensioner som bara består av mänsklig kropp-fysik. Den primära intensionen för p-begreppet skiljer sig från den primära intensionen för c-begreppet, och ur detta faktum, menar Chalmers enligt Brueckner, följer att vi kan föreställa oss zombies. Detta är enligt Brueckner nyckeln till Chalmers svar på standardinvändningen (2001, s. 190).

Brueckner kallar den egenskap som pekas ut av p-begreppets primära intension för p1-egenskapen och den som pekas ut av c-begreppets primära intension för c1-egenskapen, och sammanfattar Chalmers argument som följer:

- (1) The primary intension of the c-concept differs from the primary intension of the p-concept. [from the conceivability of the truth of (P & ~C)]
- (2) The c1-property differs from the p1-property. [from (1)]
- (3) The c1-property applies to things in our world. [obvious]
- (4) The only physical property of things in our world that is a candidate for being the c1-property is the p1 property. [obvious]
- (5) There is a property of our world over and above the physical properties, viz. the c1-property of having a phenomenal feel. [from (2), (3), (4)]

Nu kan man tänka sig en situation där våra fysiska dubbelgångare är zombies, men, menar Brueckner, trots att vi kan *tänka oss* dess sanning så är (P & ~C) inte möjlig. Det finns enligt Brueckner inte en möjlig zombievärld, och detta på grund av standardinvändningen; att den är

²² "Har en fenomenell aspekt" är en översättning av "has a phenomenal feel".

tänkbar är inte tillräckligt. Hela Chalmers argument vilar enligt Brueckner på (1), där antagandet om att det finns en möjlig värld där p-begreppets primära intension leder till en icke-tom extension som innehåller våra fysiska dubbelgångare, medan c-begreppet leder till en tom extension är centralt. Så vi behöver anta att det finns en möjlig värld som innehåller våra fysiska kopior men som saknar tillstånd med fenomenella aspekter, det vill säga medvetna mentala tillstånd. Men Chalmers enda grund till detta antagande är att vi kan *tänka oss* att $(P \ \& \ \sim C)$ är sann genom att föreställa oss dubbelgångare utan medvetna mentala tillstånd. Detta är enligt Brueckner inte tillräckligt för att garantera en sådan möjlig värld (2001, s. 192).

Chalmers svar på standardinvändningen vilar alltså på en härledning ur den tänkbara sanningen av $(P \ \& \ \sim C)$ till möjligheten att det skulle kunna finnas fysiska kopior av oss som saknar medvetna mentala tillstånd för att fastställa divergensen mellan de primära intensionerna som antas i (1). Men det är precis den här härledningen som ifrågasätts av standardinvändningen när den appliceras på Chalmers argument. Chalmers tänkbarhetsargument mot tesen att medvetna mentala tillstånd supervenierar på det fysiska tillbakavisar inte standardinvändningen, menar Brueckner; tvådimensionssemantiken ger inget nytt svar (2001, s. 193).

Jag håller helt och fullt med Brueckner om att Chalmers svar inte tillbakavisar standardinvändningen. Vad Chalmers gör i sitt svar är endast att tillbakavisa *en viss sorts* invändning som ofta görs tillsammans med standardinvändningen, nämligen invändningar på *a posteriori*-grunder, som involverar metafysisk möjlighet. Chalmers visar med hjälp av tvådimensionssemantiken att zombies är logiskt möjliga genom att hänvisa till primärintensioner och att hävda att sekundärintensioner inte är relevanta, men frågan om huruvida slutsatsen att zombies är möjliga kan *härledas från* påståendet att de är tänkbara står i min mening, precis som Brueckner påpekar, fortsatt obesvarad även efter Chalmers lösningsförsök.

5.3. TÄNKBARHET, MÖJLIGHET OCH FYSIKALISM

Sara Worley argumenterar i en artikel från 2003 för att Chalmers zombieargument inte påverkas av den så kallade standardinvändningen, därför att Chalmers förstår 'tänkbarhet' på ett sätt på vilket möjlighet kan härledas från det. Men, menar Worley, Chalmers sätt att förstå tänkbarhet leder till att vi inte kan vara säkra på att zombies faktiskt *är* tänkbara.

Worley diskuterar begreppet tänkbarhet, och menar att detta ofta förstås på ett epistemiskt sätt, eftersom vad ett subjekt håller för tänkbart i stor utsträckning beror på dess kunskap och trosföreställningar. Enligt Worley är det ganska väntat att tänkbarhet realiserar på dessa sätt, eftersom vad vi är kapabla att tänka i mångt och mycket beror på vilka koncept som är tillgängliga för oss, samt på vad vi vet och inte vet. Med den här epistemiska inställningen står det enligt Worley klart att tänkbarhet inte härleder möjlighet, eftersom vad som är tänkbart beror på vilka koncept eller bakgrundsinformation som är tillgänglig för oss. Standardinvändningen är enligt Worley en plausibel invändning för den epistemiska förståelsen av tänkbarhet därför att vissa satser kan verka tänkbara för oss trots att de är omöjliga (2003, s. 17). Worley menar dock att Chalmers förståelse av tänkbarhet är icke-epistemisk, och exemplifierar med följande citat från *The Conscious Mind*:

”In determining whether it is logically possible that some statement is true, the constraints are largely conceptual. The notion of a male vixen is contradictory, so a male vixen is logically impossible; the notion of a flying telephone is conceptually coherent, if a little out of the ordinary, so a flying telephone is logically impossible. (Chalmers, 1996, s. 35)”²³

Det är enligt Worley uppenbart att Chalmers menar att världars möjlighet, precis som påståendens möjliga sanningar, bestäms av begreppsliga begränsningar; det finns inga möjliga världar i vilken det finns tikar av hankön på grund av inkompatibiliteten mellan begreppet att vara en hane och begreppet att vara en tik. Det enklaste sättet att utöka denna förståelse av möjlig sanning om påståenden till världars möjlighet verkar enligt Worley vara att utöka begreppslik koherens eller inkoherens till att innefatta hela världar genom att säga att världen själv, som helhet, är koherent eller inkoherent.

Påståendet att en värld är inkoherent tolkar Worley som att det finns någon inkoherens i beskrivningen av den världen. Men detta leder enligt Worley till problem; det skulle kunna vara så att två eller flera beskrivningar fanns tillgängliga av en ensam situation; en som är koherent och en som är inkoherent. Worley menar att om det finns en enda inkoherent beskrivning av en situation så är det tillräckligt för att visa att situationen är omöjlig (2003, s. 18). Vi vet att det inte finns några möjliga världar där tikar kan vara hanar, på grund av inkoherensen när det kommer till en tik som är en hane. Om det finns andra beskrivningar varav någon eller några är koherenta är enligt Worley inte relevant (2003, s. 18).

Worley menar att eftersom Chalmers hävdar att vi bör förstå logisk möjlighet genom begreppsliga termer måste det vara så att om det förekommer någon begreppslik inkoherens så är den beskrivna situationen omöjlig. Vi kan förstå varför tikar som är hanar är omöjliga, men tanken på att det skulle kunna finnas nödvändigheter som inte på så sätt är grundade i begreppslik inkompatibilitet är svår. Att anta att det finns sådan är att anta att det finns restriktioner i rymden av möjliga världar över och under de begreppsliga restriktionerna, menar Worley, men att sådana restriktioner skulle vara godtyckliga.

Enligt Worley bör vi anta att restriktionerna om vad som är möjligt helt och hållet är begreppsliga för att undvika godtycklighet, och att något är omöjligt om det finns åtminstone en beskrivning av den som visar på konceptuell inkoherens. Worley menar att vi nu kan se att det inte finns några konstigheter i härledningen från tänkbarhet till möjlighet; för Chalmers är tänkbarhet och möjlighet enligt Worley samma sak. Att vara tänkbar är att vara sann i en tänkbar värld, men de tänkbara världarna är de möjliga världarna; så att vara tänkbar är att vara sann i en möjlig värld och därmed möjlig. Att möjlighet kan härledas från tänkbarhet är därmed, enligt Worley, utan undantag (2003, s. 19).

Worley ser dock av två andra problem med Chalmers argument. Först; även om en situation är omöjlig endast om det finns någon inkoherent beskrivning av den så kanske vi inte känner till den beskrivningen. En situation kan *verka* tänkbar på grund av sättet på vilket den presenteras inte är tillräckligt för att avslöja dess inkoherens, även om det finns någon annan beskrivning där inkoherensen avslöjas. Worley menar att även om tänkbarhet inte relativiseras av vår

²³ Exemplet med ordet ”vixen”, det vill säga rävhona, är inte lika träffande på svenska som på engelska. För att tydliggöra Worleys poäng kommer jag därför istället använda ordet ”tik”, det vill säga en hund av hankön.

epistemiska status så är våra *utsagor om vad som är tänkbart* nödvändigt relativiserad av detta; våra utsagor kanske inte är pålitliga indikatorer om vad som faktiskt är tänkbart. Vad vi kan säga, menar Worley, är bara att det verkar tänkbart för oss, eller att vi inte kan avgöra huruvida det är tänkbart eller inte. Vi kan inte förlita oss på vad som verkar tänkbart som ett bevis på vad som faktiskt är tänkbart (2003, s. 20).

Vidare, menar Worley, finns det andra orsaker att vara skeptisk till adekvansen i våra tänkbarhetsutsagor, givet Chalmers förståelse av tänkbarhet. Om att vara tänkbar är att vara sann i en tänkbar värld så måste man för varje värld undersöka huruvida en viss proposition i en värld skulle motsäga något annat som är sant i den världen, och detta kan vi inte avgöra endast genom att undersöka våra koncept (2003, s. 20). Worley visar på problemet genom ett exempel med Goldbachs förmodan.²⁴ För Chalmers hänger enligt Worley omöjlighet och otänkbarhet ihop med begreppslig inkoherens, så om Goldbachs förmodan är sann så måste det finnas någon form av inkoherens i påståendet att den är falsk. Men enligt Worley beror att det inte finns någon tänkbar värld där Goldbachs förmodan är falsk inte på att påståendet att den skulle kunna vara falsk är självmotsägande eller inkoherent, utan på att Goldbachs förmodan kan härledas ur andra matematiska sanningar som är sanna i varje värld (2003, s. 20).

Worley menar att det inte räcker att tänka sig fysiska kopior av människor utan fenomenella upplevelser för att se om detta involverar någon motsägelse. För varje möjlig värld måste vi kontrollera huruvida sanningen av detta fenomen skulle motsäga sig andra sanningar i den världen. Om det visar sig att fysiska strukturer härleder fenomenella upplevelser så är zombievärldar inkoherenta. Det är enligt Worley oklart om vi borde förlita oss på våra intuitioner när det gäller zombiers tänkbarhet mer än vi litar på våra intuitioner när det gäller någon fråga om vad som kan härledas ur det fysiska, eller om huruvida matematiska satser är sanna (2003, s. 21). Mänskliga intuitioner är inte av någon betydelse. Worleys slutsats är att omöjlighet, givet Chalmers förståelse av tänkbarhet, kan härledas ur tänkbarhet, men på bekostnad av att vi inte kan veta om zombies är tänkbara eller ej (2003, s. 23).

Jag är inte så säker på att Worley räddar Chalmers undan standardinvändningen. Om Worley har rätt i att Chalmers har en annan förståelse av tänkbarhet (vilket jag inte alls tror) så är frågan om denna förståelse är korrekt. Varför ska vi anta en specialdefinition av tänkbarhet om vår epistemiska status, givet att Worley har rätt, ändå har så pass stark implikation på våra utsagor om vad som är som tänkbart samt begränsar oss i att se huruvida något är tänkbart eller ej?

Jag tror definitivt inte att tänkbarhet och omöjlighet för Chalmers är samma sak. Det vore för det första onödigt att ens ha med premiss 1 i argumentet om han menade att tänkbarhet och omöjlighet när allt ställs på sin spets ändå är samma sak. Chalmers premiss är att *om* något är tänkbart *så* är det logiskt möjligt, och detta är inte ett påstående om att tänkbarhet och omöjlighet skulle vara identiska. Dessutom så följer inte bara TM-tesen, att omöjlighet kan härledas ur tänkbarhet, ur att tänkbarhet och omöjlighet skulle vara samma sak, utan också den högst osannolika MT-tesen, att tänkbarhet kan härledas ur omöjlighet. Detta tror jag inte att Chalmers

²⁴ *Goldbachs förmodan* är ett matematikproblem som består i att varje jämnt tal från och med 4 kan skrivas som summan av två primtal. Denna hypotes har visat sig stämma i alla hittills testade fall, men man har inte kunnat bevisa att den är sann.

är intresserad av. Chalmers poäng är att givet ett begrepps tänkbarhet *så* är detta begrepp logiskt möjligt, inte att tänkbarhet är ekvivalent med möjlighet.

Vad gäller Worleys slutdiskussion om tänkbarhet som något som inte relativiseras av mänskligt tänkande utan av något slags tänkbarhet *i sig* ställer jag mig frågande till om det verkligen finns en distinktion mellan tänkbart för oss och tänkbart i sig. Definitionen av tänkbarhet torde vara just att det är något som *kan tänkas av* ett subjekt. Vad som däremot kanske inte relativiseras av mänskligt tänkande är *möjligheter*, och jag tror snarare att detta är en av anledningarna till att standardinvändningen har uppkommit; att härledningar av vad som är möjligt ur vad som är tänkbart kanske inte är så bra.

Eftersom Chalmers omöjligt kan ha förståelsen av tänkbarhet och möjlighet som samma sak berör standardinvändningen i lika hög grad hans argument efter som före Worleys artikel. Att vi, som Worley menar, inte skulle kunna vara säkra på om zombies är tänkbara eller ej verkar inte heller stämma eftersom tänkbarhet är något som uppenbarar sig direkt för subjektet och omöjligt kan bestämmas av någon slags tänkbarhet i sig.

5.4. ANTI-ZOMBIEARGUMENTET

Keith Frankish drar i en artikel från 2007 slutsatsen att zombieargumentet förvisso är ett elegant och förföriskt filosofiskt argument, men att det faktum att vi skulle kunna avgöra medvetna mentala tillståndets natur på grund av vår föreställningsförmåga verkar för bra för att vara sant. Frankish menar att det faktum att vi kan konstruera ett anti-zombieargument visar att det inte är sant. Zombies och anti-zombies ”tar ut” varandra, och detta visar att tänkbarhet inte spelar någon roll i debatten om medvetna mentala tillståndets natur.

Frankish skriver ut zombieargumentet på följande sätt:

1. Zombies are conceivable
2. If zombies are conceivable, then zombies are metaphysically possible
3. If zombies are metaphysically possible, then consciousness is non-physical
4. So consciousness is non-physical.

Tänkbarheten som avses här är, menar Chalmers enligt Frankish, *ideal* och *primär* (2007, s. 652). Ett sakförhållande är idealt tänkbart om dess tänkbarhet inte kan besegras av bättre resonemang, och primärt tänkbart om det för allt vi vet *a priori* skulle kunna hålla. Primär tänkbarhet är sådan tänkbarhet som har att göra med primära intensioner; de begränsas eller styrs endast av logik och *a priori*-meningen hos involverade termer.

Frankish menar att zombieargumentet kan göra anspråk på att vara sunt. Premiss 1 är plausibel. Premiss 2 är en instans TM-tesen, som är kontroversiell, men det finns enligt Frankish starka argument för den. Den viktigaste invändningen mot denna tes kommer ur *a posteriori*-nödvändigheter, men vi missleds bara om vi felaktigt beskriver detta som den metafysiska möjligheten att vatten inte är H₂O, menar Frankish (2007, s. 652). Premiss 3 är okontroversiell,

menar Frankish; om det är möjligt för en varelse att ha alla fysiska egenskaper utan att ha fenomenella upplevelser så är fenomenella egenskaper något utöver fysiska.

Vad Frankish vill göra för att motbevisa detta argument är att visa att vi precis lika enkelt kan konstruera ett likadant till förmån för fysikalismen. Frankish introducerar begreppet *anti-zombie*, en varelse som är exakt som en vanlig människa, lever i ett universum som är precis som vårt, och som har medvetna upplevelser som vi. I anti-zombievärlden är medvetna mentala tillstånd ett fysiskt fenomen, som supervenierar metafysiskt på de mikrofysiska egenskaperna. Anti-zombieargumentet mot dualism går såhär:

5. Anti-zombies are conceivable
6. If anti-zombies are conceivable, then anti-zombies are possible
7. If anti-zombies are possible, then consciousness is physical
8. So consciousness is physical.

Detta argument gör också starkt anspråk på att vara sunt. Beskrivningen av en anti-zombie är lika koherent som den av en zombie, menar Frankish; något som är fysiskt identiskt med mig och som *har* medvetna mentala tillstånd – ljuset är tänt inom den. De andra premisserna är också plausibla på samma sätt som premisserna i zombieargumentet. I anti-zombievärlden är medvetna mentala tillstånd fysiska, så de mikrofysiska egenskaperna i den världen är metafysiskt tillräckliga för medvetna mentala tillstånd, och vilken annan värld som helst med samma mikrofysiska egenskaper kommer ha samma distribution av fenomenella egenskaper. Vår värld har detta, så de mikrofysiska egenskaperna i vår värld är tillräckliga för medvetna mentala tillstånd. Så följande är, eftersom metafysisk superveniens härleder fysikalism, ekvivalent med (7), menar Frankish:

9. If anti-zombies are possible, then the microphysical features of our world are metaphysically sufficient for the existence of consciousness (2007, s. 654).

Detta är anti-zombieargumentet. Frankish menar att det kanske verkar konstigt att hävda att endast tänkbarheten av fysikalism härleder dess sanning, men denna strategi är exakt samma som används i zombieargumentet. I båda fallen för TM-tesen oss från en psykologisk utsaga till en metafysisk; vi föreställer oss något och härleder något annat ur det. Om denna strategi är sund i det ena fallet så är den enligt Frankish det också i det andra (2007, s. 655).

Argumentets syfte är att sätta TM-tesen under press. Både zombie- och anti-zombieargumentet kan enligt Frankish inte vara sunda. Frankish menar att om både zombies och anti-zombies är tänkbara så härleder tänkbarhet inte möjlighet därför att båda enligt Frankish inte kan motsvara en genuin möjlighet eftersom nödvändigt existerande entiteter måste existera i alla världar om de existerar i någon (2007, s. 656).

Zombieförespråkarnas enda sätt att lyckas motbevisa anti-zombieargumentet är enligt Frankish att hävda att anti-zombies är otänkbara. Genom att hävda att zombies är möjliga så måste zombieförespråkare enligt Frankish hävda att anti-zombies är omöjliga, eftersom han menar att dessa möjligheter är inkompatibla. Om det finns en anti-zombie-värld så är dess mikrofysiska

egenskaper metafysiskt tillräckliga för medvetna mentala tillstånd, och därmed kan det inte finnas någon värld med samma mikrofysiska egenskaper men som saknar medvetna mentala tillstånd. Man kan också formulera problematiken som att det bara finns en enda möjlig värld som är en kopia av den faktiska världen och frågan är om invånarna på denna värld har medvetna mentala tillstånd eller ej. De kan enligt Frankish inte båda ha det och inte, och han menar därför är idén om zombies inkompatibel med idén om anti-zombies (2007, s. 658).

Frankish menar att om zombieargumentet hade varit sunt så hade zombieförespråkarna bara behövt visa att begreppet om anti-zombies är inkoherent. Detta ger en förändring i bevisbördan, som Chalmers förklarar på följande sätt:

”In general, a certain burden of proof lies on those who claim that a given description is logically impossible... If no reasonable analysis of the terms in question points towards a contradiction, or even makes the existence of a contradiction plausible, then there is a natural assumption in favour of logical possibility.” (Chalmers, 1996, s. 96).

Frankish argumenterar nu för att zombieförespråkare inte kan hantera denna börda och att zombieargumentet därför faller (2007, s. 659). Det är enligt Frankish inte alls omöjligt, eller ens särskilt svårt, att formulera ett begrepp om hur medvetna mentala tillstånd skulle kunna vara något fysiskt; det är bara att föreställa sig en sammansättning av en fysisk dubbelgångare och medvetna mentala tillstånd, och detta är enligt Frankish inte svårare än att skilja dem åt i zombiefallet (2007, s. 660).

Frankishs idé med att konstruera anti-zombieargumentet är att visa upp det som ett taktiskt försök att neutralisera zombieargumentet. Dess primära funktion är att visa att TM-tesen är ett tvåeggat svärd som inte borde användas. Den naturliga slutsatsen att dra, menar Frankish, är att både zombieargumentet och anti-zombieargumentet är opålitliga av samma anledning (2007, s. 665). Om man välkomnar anti-zombie argumentet så måste man naturligtvis visa att zombies är inkoherenta, men grunderna för att göra detta är, enligt Frankish, bättre än för att göra motsatsen för anti-zombies. Frankish menar därför att dualister inte borde vara zombieförespråkare. Tanken att vi kan avgöra medvetna mentala tillstånds natur genom att föreställa oss något verkar för bra för att vara sann, och det faktum att vi kan konstruera ett anti-zombieargument visar att det faktiskt är för bra för att vara sant. Därmed, menar Frankish, har tänkbarhet ingen roll att spela i debatter om medvetna mentala tillstånds natur (2007, s. 666).

Jag ställer mig först och främst frågande till Frankishs påstående att Chalmers skulle mena att zombies är metafysiskt möjliga; det är ju precis det han säger att de inte är genom att besvara standardinvändningen i termer av tvådimensionell semantik. Genom att visa att sekundära intensioner, och därmed metafysisk möjlighet, inte är involverade i zombieargumentet, utan att det som är relevant är primära intensioner och därmed logisk möjlighet, visar Chalmers klart och tydligt att zombies är logiskt, inte metafysiskt, möjliga.

Jag tror vidare att Frankish har fel i att anti-zombies är tänkbara. Det är naturligtvis lika möjligt att medvetna mentala tillstånd skulle kunna bestå av något fysiskt som något annat. Problemet är att hur medvetna mentala tillstånd är konstituerade är otänkbart för oss i samma bemärkelse som de kemiska beståndsdelarna i den banan jag just nu äter eller hur en flygplansmotor

fungerar är otänkbara för mig. Anledningen till att detta är otänkbart för mig har att göra med min avsaknad av kunskap. Jag har helt enkelt inte den blekaste aning om hur det faktiskt förhåller sig, och jag tror att detsamma gäller när vi ska försöka tänka oss hur medvetna mentala tillstånd är konstituerade. Vi kan ha kunskap om *att* medvetna mentala tillstånd finns och tänka oss saker om vilken funktion de fyller, på samma sätt som jag vet att bananen gör mig mätt eller att motorn har del i att få flygplanet att lyfta, men på samma sätt som jag inte kan tänka mig hur bananen eller flygplansmotorn är sammansatt kan vi inte tänka oss *av vad* medvetna mentala tillstånd är konstituerade.

Problemet med anti-zombies tänkbarhet är, som jag ser det, det faktum att erkännandet av anti-zombies som tänkbara innebär att ta ställning till hur medvetna mentala tillstånd är konstituerade. Detta är något helt annat än att tänka bort medvetandet. Att tänka sig varelser utan medvetna mentala tillstånd är inte särskilt komplicerat (jag tror till och med tanken på att andra människor skulle kunna sakna själsliv är något de flesta av oss någon gång har snuddat vid), medan att tänka sig hur medvetna mentala tillstånd faktiskt förhåller sig kräver något helt annat av subjektet. Att tänka sig in i hur något *faktiskt är* innebär att ta hänsyn till helt andra saker än när man bara tänker sig att man tar bort något, och därmed verkar zombie- och anti-zombieargumentet inte vara likvärdiga på det sätt Frankish menar att de är.

Om möjlighet kan härledas ur tänkbarhet, och om vi för stunden går med på att såväl zombies som anti-zombies är tänkbara, så stämmer det att båda begreppen är möjliga. Frankish menar att detta är problematiskt; att zombie- och anti-zombiebegreppet inte samtidigt kan motsvara genuina möjligheter. Jag tror att detta är felaktigt, och att flera inkompatibla begrepp visst kan vara möjliga samtidigt. Ett exempel är att det är möjligt att Sveriges kung just nu står upp, men det är också möjligt att Sveriges kung just nu sitter ner.²⁵ Zombies och anti-zombies kan inte vara möjliga *i samma värld*, men i olika möjliga världar kan de motsvara genuina möjligheter, och därför verkar Frankishs anti-zombieargument inte kunna vederlägga zombieargumentet. Om möjlighet kan härledas ur tänkbarhet och båda begreppen är tänkbara så motsvarar zombies och anti-zombies helt enkelt möjligheter i olika världar.

6. DISKUSSION OCH SLUTSATS

Denna uppsats syfte var att undersöka huruvida Chalmers zombieargument är ett sunt och giltigt argument eller ej. I sektion 1 presenterade jag Chalmers position, för att i sektion 2 undersöka tre invändningar mot zombieargumentet som jag betraktar som mindre starka. Därefter presenterade jag den så kallade standardinvändningen mot tänkbarhetsargument, det vill säga påståendet att möjlighet inte kan härledas ur tänkbarhet, samt Chalmers svar på denna invändning i termer av tvådimensionell semantik. I sektion 5 undersökte jag fyra argument som alla behandlade den så kallade TM-tesen, som premiss 2 i Chalmers argument är en instans av. Trots att jag visade på problem hos tre av fyra presenterade argument mot TM-tesen (alla utom Brueckner) så tror jag ändå att kritiken av den är befogad, på ett sätt som jag ska förklara nedan.

²⁵ Ett exempel som min handledare Martin Jönsson presenterade för mig.

Jag börjar med en påminnelse om hur Chalmers zombieargument ser ut:

P1: Zombies är tänkbara

P2: Om något är tänkbart så är det logiskt möjligt

Så

S1: Zombies är logiskt möjliga

S2: Det mentala supervenierar inte på det fysiska

Med premiss 1 ser jag inga problem. Vi kan tänka oss många olika mer eller mindre konstiga saker, som till exempel att vara någon annan, mig själv i ett framtida skede, mig själv utan armar, tidsresor och enhörningar. Tanken på mig själv utan medvetna mentala tillstånd (det vill säga min zombietvilling) verkar inte vara en konstigare tanke än något av detta. Premiss 1 i Chalmers argument verkar alltså plausibel.

Jag tror dessutom att Chalmers har rätt i att zombies är logiskt möjliga; det finns ingen motsägelse i deras begrepp, och detta tror jag i enighet med Chalmers är tillräckligt för att fastställa deras logiska möjlighet. Vad jag däremot betraktar som problematiskt är hur Chalmers når slutsatsen att zombies är logiskt möjliga. Som jag ser det är premiss 2 nämligen inte rättfärdigad. Med sitt försök till svar på standardinvändningen visar Chalmers inte att man kan härleda logisk möjlighet från tänkbarhet, utan vad han gör är helt enkelt att värja sig från en viss sorts invändning, nämligen den sort som görs på *a posteriori*-grunder och därmed involverar metafysisk möjlighet.

Däremot så påstår Chalmers själv att det är fullt tillräckligt att det inte finns någon motsägelse i zombiebegreppet för att fastställa slutsatsen att zombies är logiskt möjliga, och detta tror jag kan vara den springande punkten. För mig verkar premiss 1, att zombies är tänkbara, av denna anledning överflödig; att zombies är logiskt möjliga verkar kunna fastställas på ett sätt som inte bara involverar tänkbarhet, och därmed inte involverar den problematiska TM-tesen. Mitt förslag är att göra om zombieargumentet på följande sätt:

P1: Om ingen inkonsistens kan detekteras i begreppet x så är existensen av x logiskt möjlig²⁶

P2: Ingen inkonsistens kan detekteras i zombiebegreppet

Så

S1: Zombies är logiskt möjliga

S2: Det mentala supervenierar inte på det fysiska

Jag tror nämligen att tänkbarhet är en sak och att undersöka huruvida ett begrepp innefattar motsägelser eller inte är en annan, och att det är ur det senare vi kan nå evidens för möjlighet. Chalmers position när det gäller härledning till möjlighet från tänkbarhet verkar vara att blotta uppenbarelsen av en tanke fastställer dess logiska möjlighet, det vill säga att blotta tanken på

²⁶ Med att ingen inkonsistens ska kunna detekteras i begreppet x avser jag att ingen inkonsistens ska kunna detekteras i den mängd satser vi använder för att beskriva x.

något är tillräckligt för att nå slutsatsen att det är logiskt möjligt. Att däremot nå slutsatsen att något är logiskt möjligt på basis av att medvetet kontrollera huruvida ett begrepp är konsistent eller inte innebär en kritisk granskning av ett begrepp. Jag tror att det senare är ett betydligt mer tillförlitligt sätt att fastställa logisk möjlighet än det faktum att något är tänkbart någonsin kan ge oss.

Tänkbarhet verkar vara något som uppenbarar sig internt och mentalt för subjektet, medan kritisk granskning av ett begrepp för att detektera eventuell inkonsistens kan ske intersubjektivt, exempelvis genom filosofisk diskussion. Förutom detta så skulle en dator mycket väl kunna programmeras till att detektera eventuell inkonsistens hos ett begrepp, och därmed fastställa logisk möjlighet eller omöjlighet. Datorn skulle däremot aldrig kunna få en plötslig uppenbarelse om att något är tänkbart, och därmed aldrig utifrån tänkbarheten av ett begrepp kunna dra slutsatsen att detta begrepp är logiskt möjligt eller inte. Tänkbarhet verkar därtill i hög utsträckning kunna vara en gradfråga; något kan vara mer eller mindre tänkbart, medan att detektera eventuell inkonsistens är ett absolut kriterium för logisk möjlighet. Antingen så finns motsägelsen där eller så gör den det inte.

Det skulle vidare kunna vara så att en ond demon lurar mig till trosföreställningen att vissa begrepp är tänkbara trots att de egentligen inte alls är det, eller att jag på grund av tillfällig sinnesförvirring når samma slutsats. Huruvida något av detta är vad som har skett är naturligtvis omöjligt för mig att kontrollera därför att det bara är begreppets tänkbarhet, så som det uppenbarar sig för mig, som kan användas som evidens för att detta begrepp är just tänkbart. Eftersom det endast är inför ett subjekt något kan vara tänkbart så kan tänkbarhet aldrig testas eller kontrolleras intersubjektivt. Det är av denna anledning möjligt att tala om såväl ”tänkbart för mig” och ”tänkbart för dig” som, om möjlighet kan härledas från tänkbarhet, ”möjligt för dig” och ”möjligt för mig”.

Påståenden om huruvida begrepp är konsistenta eller inte är däremot tillgängliga för intersubjektiv granskning. Detta kan endast ses som något positivt eftersom man, till skillnad från med tänkbarhet, kan visa på inkonsistens genom att helt enkelt peka ut den motsägelse man har detekterat. Det är omöjligt att på samma sätt dela med sig av tänkbarhet. Eftersom vi genom kritisk granskning av ett begrepp kan uppnå en intersubjektiv, och kanske i någon mening objektiv, bild av huruvida begreppet är logiskt möjligt eller ej står det klart att detta är ett sätt att göra härledning till logisk möjlighet säkrare.

Jag har, åtminstone för stunden, sällat mig till den mängd filosofer som menar att möjlighet inte kan härledas från tänkbarhet. Jag tror att zombies är såväl tänkbara som möjliga, men att det inte finns någon länk däremellan. Jag tror dessutom att tänkbarhetspremissen förvisso är sann, men onödigt därför att jag menar att Chalmers slutsats; att zombies är logiskt möjliga och att det mentala inte supervenierar på det fysiska, skulle kunna fastställas utan såväl denna premiss som den falska TM-premissen. Jag tror alltså att Chalmers zombieargument är osunt därför att en av dess premisser, att möjlighet kan härledas från tänkbarhet, inte är sann. Jag har utifrån detta nått slutsatsen att Chalmers zombieargument inte är något bra argument för dualism, men därmed inte sagt att zombies är omöjliga.

7. KÄLLFÖRTECKNING

- Balog, K. (1999) *Conceivability, Possibility and The Mind-Body Problem*, The Philosophical Review, Vol. 108, No. 4
- Barker-Plummer, D. Barwise, J. & Etchemendy, J. (2011) *Language, Proof and Logic*, 2:nd Edition, CSLI Publications
- Borge, S. (1999) *All You Zombies: David Chalmers' Metaphysical Solipsism*, Uwe Meixner Peter Simons (ed.); *Metaphysics in the Post-Metaphysical Age*. Austrian Ludwig Wittgenstein Society
- Brueckner, A. (2001) *Chalmer's Conceivability Argument for Dualism*, Analysis, Vol. 61, No. 3
- Chalmers, D. (1995) *Facing up to the Problems of Consciousness*, Journal of Consciousness Studies 2 (3), pp. 200–219
- Chalmers, D. (1996) *The Conscious Mind*, Oxford University Press
- Descartes, R. (1641) *Betraktelser över den första filosofien*, ur samlingen *Valda skrifter*, översatt av Konrad Marc-Wagau (1990), BTJ-Tryck, Lund
- Frankish, K. (2007) *The Anti-Zombie Argument*, The Philosophical Quarterly, Vol. 57, No. 229
- Kaplan, D. (1979) *Dthat*, In P. Cole, ed. *Syntax and Semantics*, New York: Academic Press
- Kirk, R. (1999) *The Inaugural Adress: Why There Couldn't Be Zombies*, Proceedings of the Aristotelian Society, Supplementary Volumes, Vol. 73
- Kripke, S. (1980), *Naming and Necessity*, Harvard University Press
- Nagel, T. (1974) *What Is It Like To Be a Bat?* Philosophical Review, 4:435-50
- Quine, W. V. (1969) *Propositional Objects, Ontological Relativity and Other Essays*, New York: Columbia University Press
- Searle, J. R. (1992). *The Rediscovery of the Mind*, MIT Press.
- Strawson, P. F. (1950) *On Referring*, Mind, New Series, Vol. 59, No. 23, pp. 320-34
- Thomas, N. J. T. (1998) *Zombie Killer*, in S. R. Hameroff, A. W. Kaszniak, and A. C. Scott (eds.), *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*. Cambridge, MA: MIT Press
- Worley, S. (2003) *Conceivability, Possibility and Physicalism*, Analysis, Vol. 63, No. 1