

Humes utelämnade nyans av blått

Pär Sundström, Umeå universitet

1 Introduktion

Enligt Humes begreppsempirism kommer allt tankematerial från erfarenheten: vi kan tänka det vi har erfårit, och kombinationer av det vi har erfårit, men ingenting mer.

Eller som han själv uttrycker det:

"Alla våra enkla idéer vid sina första framträdanden härrör från enkla intryck, vilka de motsvarar och exakt representerar" (Hume 1739, bok 1, del 1, sektion 1).

Hume uppmärksammar själv en typ av fall som verkar utgöra ett motexempel till denna tes. Anta att en person har upplevt alla färger utom en nyans av blått. Hume kommenterar fallet sålunda:

"Jag frågar nu om han med hjälp av sin fantasi kan fylla denna lucka och själv skapa idén om denna särskilda nyans, trots att den aldrig förmedlats honom genom sinnena? Jag tror att få skulle förneka att han kan det, och detta kan tjäna som ett bevis för att de enkla idéerna inte alltid härrör från de motsvarande intrycken (Hume 1739, bok 1, del 1, sektion 1).

Jag ska undersöka utsikterna för en begreppsempiristisk teori i ljuset av detta berömda fall. I sektion 2 urskiljer jag några varianter av begreppsempirism, och hävdar att några av dessa har i alla fall en initial plausibilitet. Sektion 3 förklarar hur dessa initialt plausibla teser hotas av fallet med den utelämnade nyansen av blått. Sektion 4 diskuterar vilken attityd man bör inta till Humes fall i förhållande till en eller annan begreppsempiristisk tes. Jag urskiljer tre möjligheter. Man kan hävda (i) att Humes

fall utgör ett motexempel till den aktuella tesen, och att tesen därför bör avvisas, (ii) att Humes fall utgör ett motexempel till tesen, men att tesen ändå bör accepteras, eller (iii) att fallet inte utgör ett motexempel till tesen ifråga. Jag ska förespråka en variant av (iii). Min slutsats är att det finns begreppsempiristiska teser som har en initial plausibilitet, och för vilka fallet med den utelämnade nyansen av blått inte är ett motexempel.

Jag bör kanske redovisa att jag framträder här som djävulens -- eller i alla fall en motståndares -- advokat. Av skäl som jag inte kan gå in på här är jag benägen att tro att de teser jag ska diskutera, även de som har en initial plausibilitet, i slutändan bör avvisas. Men jag uppfattar ändå att det resultat jag tror mig ha kommit fram till -- att en begreppsempirist kan hantera fallet med den utelämnade nyansen av blått -- är intressant, av minst två skäl. Dels är det av värde att vara klar över *varför* en tes bör avvisas, om den bör avvisas. Om jag har rätt så är Humes fall inte ett bra skäl att överge begreppsempirismen. Dels lägger Humes fall, som vi ska se, vissa begränsningar på en hur en begreppsempiristisk tes bör preciseras, för att vara plausibel. Fallet hjälper oss därmed att utkristallisera de preciseringar av begreppsempirismen som är mest intressanta att fokusera på i en mer omfattande utvärdering.¹

2 Varianter av begreppsempirism

2.1 Fenomenell begreppsempirism

Man kunde möjligen misstänka att fallet med den utelämnade nyansen av blått bara hotar en historiskt omfattad åsikt som ingen numer försvarar. Men så är det inte. Beakta t.ex. följande tes:

¹ Jag har presenterat delar av ovanstående material på *Filosofidagarna* i Umeå i juni 2007, och på workshopen *Perception and Content* i Stockholm i september 2007, och vill tacka deltagare vid båda föredragen för värdefulla synpunkter och diskussioner. Artikeln har skrivits med finansiellt stöd från Riksbankens jubileumsfond.

Fenomenell begreppsempirism: Det finns en uppsättning "fenomenella begrepp" $Q1 \dots Qn$ för en uppsättning enkla sensoriska kvaliteter $Q1 \dots Qn$ sådana att nödvändigtvis: om ett subjekt S har något av $Q1 \dots Qn$ så har S upplevt den kvalitet som begreppet står för.

Jämfört med Humes begreppsempiristiska påståenden, som verkar vara mer generella (även om det kanske inte är glasklart hur de ska förstås), är denna tes tydligt begränsad två avseenden. För det första uttalar den som bara om begrepp rörande en viss klass av egenskaper, nämligen sensoriska kvaliteter, såsom rödhet och smärta. Den säger alltså ingenting om huruvida det finns något erfarenhetsberoende begrepp för t.ex. demokrati, inflation, eller rättvisa. För det andra säger tesen bara att det finns *ett* slags begrepp för sensoriska kvaliteter vars innehav kräver att man har haft en viss upplevelse. Tesen utesluter inte att det kan finnas andra begrepp för dessa kvaliteter som man kan ha utan att ha upplevt kvaliteterna. Anta att t.ex. begreppet "moster Barbro's favoritkvalitet" är ett begrepp för rödhet. Den fenomenella begreppsempirismen medger att en person kan ha detta begrepp utan att någonsin ha upplevt rödhet.

Begränsade begreppsempiristiska teser av detta slag har haft flera försvarare under senare år. David Papineau skriver t.ex.:

Den inflationistiska teorin [som är Papineaus egen] om fenomenella begrepp påminner på vissa sätt om den traditionella empiristiska teorin för idéer. ... Inflationister behöver inte hålla med Hume om att *alla* begrepp kommer ur erfarenheten. Inflationism är en tes specifikt om fenomenella begrepp för medvetandeegenskaper, och har inga implikationer för *icke*-fenomenella begrepp för medvetandeegenskaper eller för någonting annat. Men när det gäller de fenomenella begreppen själva hävdar inflationister att du behöver en initial upplevelse ... för att förvärva ett fenomenellt begrepp (Papineau 2002, sid. 68; se även Tye 1995, sid. 167, och 2000, sid. 24).

Även en sådan begränsad tes berörs av Humes fall med den utelämnade nyansen av blått -- vilket kommer att illustreras nedan.

Jag ska huvudsakligen fokusera på Fenomenell begreppsempirism i den här uppsatsen, eftersom det är en någorlunda väldefinierad och i samtiden populär tes. Mycket av det jag ska säga kan dock tillämpas även på andra begreppsempirismer.

Jag ska härnäst urskilja några varianter av Fenomenell begreppsempirism.

2.2 Okvalificerade respektive antropologiska varianter

Vi kan till att börja med urskilja vad jag ska kalla *okvalificerade* respektive *antropologiska* versioner av Fenomenell begreppsempirism. Okvalificerad fenomenell begreppsempirism säger att det finns en uppsättning begrepp för en uppsättning sensoriska kvaliteter som är sådana att *ingen möjlig* tänkare kan ha ett begrepp av den aktuella typen utan att ha upplevt den kvalitet som begreppet står för. Antropologisk fenomenell begreppsempirism är en svagare tes, som säger att det finns en uppsättning begrepp för en uppsättning kvaliteter som är sådana att det är en *psykologisk lag för människor* att om man har ett begrepp av den aktuella typen så har man upplevt den relevanta kvaliteten.

Några ord om hur jag kommer att förstå den antropologiska tesen. Jag ska förstå den som att den fäller ett påstående inte bara om faktiska, utan även kontrafaktiska omständigheter. (Detta är innebörden i att den är en "lag".) Samtidigt ska jag förstå den så att den medger *exceptionella undantag*. För att illustrera tesens karaktär, tänk på påståendet att: om en människa S är en kompetent talare av japanska, så har S varit i kontakt med det japanska språket. Detta är rimligen sant för alla faktiska människor, och regelbundenheten verkar tillräckligt robust för att vi ska anta att den gäller även kontrafaktiska omständigheter: det är rimligt att anta att om en viss talare av japanska aldrig någonsin *hade* varit i kontakt med det japanska språket, så *skulle* hon inte ha varit en kompetent japansk talare. Men det är samtidigt rimligt att anta att tesen medger *exceptionella undantag*; det är rimligen i princip möjligt för en människa att bli en kompetent talare av japanska genom någon slumpmässig händelse eller planerad procedur (t.ex. hjärnkirurgi) som omstrukturerar hennes hjärna på rätt sätt men som inte involverar att hon kommer i kontakt med det japanska språket.

2.3 "Uppleva"

Man får starkare respektive svagare versioner av Fenomenell begreppsempirism i en annan dimension genom att variera förståelsen av vad det innebär att "uppleva", eller "erfara", en viss sensorisk kvalitet. Tänk t.ex. på en veridikal perception av blått, en hallucination av blått, och en avsiktlig, viljestyrd visualisering av blått. Räkna man alla dessa som upplevelser av blått får man en viss begreppsempiristisk tes, räknar man bara (t.ex.) de två första så får man en annan, starkare version.

2.4 Konstitutiva respektive icke-konstitutiva varianter och versioner av dessa

Fenomenell begreppsempirism kan -- liksom andra begreppsempirismer -- vidare vara antingen *konstitutiv* eller *icke-konstitutiv*. Konstitutiv fenomenell begreppsempirism säger att det är *konstitutivt* för att ha ett fenomenellt begrepp att man har haft en viss upplevelse; d.v.s. att ha haft en viss upplevelse är en del av vad det *är* att ha begreppet. Icke-konstitutiv fenomenell begreppsempirism säger att en viss typ av erfarenhet *krävs*, men är inte konstitutivt för, att ha ett visst fenomenellt begrepp.

Av utrymmesskäl ska jag här sätta Konstitutiv fenomenell begreppsempirism åt sidan, och fokusera på den icke-konstitutiva varianten.

En icke-konstitutiv fenomenell begreppsempirist kan ha olika idéer om vad som är konstitutivt för att ha ett fenomenellt begrepp och olika beledsagande idéer om vad exakt upplevelser förser en med och varför de är nödvändiga för att hamna i det tillstånd som är konstitutivt för det relevanta begreppsinnehavet. Betrakta t.ex. följande tre teser:

Icke-konstitutiv *imagistisk* fenomenell begreppsempirism: Att ha ett fenomenellt begrepp Q för en enkel sensorisk kvalitet Q består (delvis) i att kunna avsiktligt skapa en inre "fantasibild" av Q (t.ex. visualisera Q), och för att kunna avsiktligt skapa en inre fantasibild av Q , så måste man ha upplevt Q .

Icke-konstitutiv *intellektualistisk* fenomenell begreppsempirism: Att ha ett fenomenellt begrepp Q för en enkel sensorisk kvalitet Q består (delvis) i att veta eller förstå hur Q sensoriskt "framträder" (ser ut, smakar, etc.), och för att veta eller förstå hur Q sensoriskt framträder måste man ha upplevt Q .

Icke-konstitutiv *igenkänningsbaserad* fenomenell begreppsempirism: Att ha ett fenomenellt begrepp Q för en enkel sensorisk kvalitet Q består (delvis) i att kunna "känna igen" Q , och för att kunna "känna igen" Q måste man ha upplevt Q .

Några anmärkningar om dessa teser:

(i) Den imagistiska tesen är ett fall där man behöver överväga exakt hur mycket som ska räknas som en "upplevelse av en sensorisk kvalitet". Man kan med viss rimlighet argumentera för att uttrycket, i denna tes, bör förstås i en något insnävad mening (jfr. 2.3 ovan). Om uttrycket förstås i vidast möjliga betydelse så medger tesen bl.a. att en person förvärvar förmågan att avsiktligt skapa en inre fantasibild av en kvalitet Q genom att avsiktligt skapa en inre fantasibild av Q ; d.v.s. förvärvar en förmåga genom att utöva denna förmåga. Det kan förefalla omöjligt, vilket skulle tala för att tesen kunde förstärkas utan att förlora plausibilitet. Man kunde t.ex. föreslå att för att förvärva förmågan att avsiktligt skapa en inre fantasibild av Q måste man ha "upplevt" Q i någon mer begränsad mening -- t.ex. genom att presenteras med Q i en veridikal perception eller hallucination.

(ii) Jag kan inte här diskutera i detalj vad det kan innebära att "veta eller förstå hur en kvalitet sensoriskt framträder". Jag ska istället bara anta följande *test* för när denna typ av kunskap eller förståelse föreligger:

Förståelse-testet: Ett subjekt S vet eller förstår hur en kvalitet Q sensoriskt framträder om och endast om: om S presenterades med och uppmärksammade Q skulle hon inte *känna sig upplyst* om hur Q sensoriskt framträder. D.v.s. om S inte är benägen att ha reaktion av typen 'aha, blått ser ut sådär', eller 'någonting ser ut sådär' när hon uppmärksammar blått, så förstår S redan hur blått sensoriskt framträder; om hon är benägen att ha någon sådan "aha-upplevelse" så förstår hon inte hur blått sensoriskt framträder.

(iii) "Igenkänningsbaserade" teorier om fenomenella begrepp framförs ofta i litteraturen (se t.ex. Loar 1990/1997, och Levin 2006), men det är inte alltid så klart -- i alla fall inte för mig -- exakt hur de ska förstås. Jag kan dock, åter igen, inte rota i denna fråga utan ska istället bara diskutera *en* möjlig kandidat till vad det kan innebära att "känna igen" (t.ex.) blått, nämligen denna: Du kan "känna igen" blått om och endast om du kan på ett riktigt sätt tillämpa termen 'blått' (och relaterade termer, som 'blå', och översättningar, som 'blue') på basis av dina synupplevelser (t.ex. säga 'jag upplever blått' när du upplever blått, men inte när du upplever grönt).

(iv) De tre teserna -- den imagistiska, den intellektualistiska, och den igenkänningsbaserade -- representerar *olika* påståenden, men de är uppenbart *förenliga*. Man kan t.ex. ställa sig bakom en konjunktion av vilket som helst par av dem, och även konjunktionen av alla tre. (Det finns tendenser till sådana konjunktiva påståenden i litteraturen; se t.ex. Tye 2003, sektion 15-6.)

(v) Såvitt jag kan se är de tre teserna i viss, men i väldigt begränsad utsträckning, beroende av varandra.

Det är möjligen rimligt att man kan *avse* bara det man kan *förstå*, och att man därför kan avsiktligt skapa en inre fantasibild av en kvalitet Q bara om man förstår hur Q sensoriskt framträder. Om det är riktigt så implicerar den konditionala delen av den intellektualistiska teorin (att om man förstår Q så har man upplevt Q) den konditionala delen av den imagistiska teorin (att om man kan avsiktligt skapa en inre fantasibild av Q så har man upplevt Q). Det är i så fall ett beroendeförhållande mellan teserna.

Utöver denna eventuella koppling är dock de aktuella förmågorna att skapa inre fantasibilder, att förstå hur kvaliteter sensoriskt framträder, och att känna igen sådana kvaliteter, oberoende av varandra, såvitt jag kan se, och följaktligen är de tre teserna också oberoende av varandra.

Till att börja med verkar det klart att man kan förstå hur en kvalitet sensoriskt framträder utan att kunna avsiktligt skapa en inre fantasibild av den. Jag kan t.ex. vara väl bekant med kvaliteten blå och inte alls vara benägen att tänka 'aha, blått ser ut sådär' när jag upplever den, och ändå sakna förmågan att spontant och på egen hand kunna visualisera blått. (Man stöter nu och då på dem som hävdar sig vara beskaffade på detta vis.)

Det verkar vidare klart att man kan både förstå hur blått ser ut och kunna avsiktligt visualisera blått utan att kunna känna igen blått -- i alla fall enligt det test för igenkänning som jag arbetar med här. För att illustrera: Frank Jacksons Mary (Jackson 1982, 1986) har som bekant växt upp i ett svart-vitt rum där hon aldrig har upplevt några kromatiska färger. Anta nu att vi sticker in ett blått spelkort under Marys dörr, men inte talar om för Mary vilken färg kortet har. Det verkar fullt möjligt att Mary, genom att granska kortet, förvärvar en förståelse för hur blått ser ut och en förmåga att visualisera blått, men inte förvärvar förmågan att tillämpa ordet 'blått' på basis av sina synupplevelser (om vi frågar henne om kortet är blått eller grönt så kan hon bli svaret skyldig).

Det verkar till sist klart att man kan ha förmågan att känna igen blått (enligt mitt aktuella test) utan att vare sig förstå hur blått ser ut eller kunna visualisera blått. För att illustrera, tänk på "Klärvoajanta Mary". Klärvoajanta Mary har, liksom Jacksons Mary, aldrig upplevt någon kromatisk färg. Men hennes hjärna är så beskaffad att, när hon upplever blått så blir hon omedelbart och regelmässigt övertygad om att det hon upplever kallas (på svenska) 'blått'. (Det är inte svårt att tänka sig detta. Våra upplevelser av blått har rimligen en eller annan neural realisering, kalla den N1. Vår benägenhet att med övertygelse tillämpa 'blått' på något vi ser har rimligen en annan neural realisering, kalla den N2. I normala fall uppstår en kausal koppling mellan N1 och N2 genom inlärning. I Marys fall finns den där utan någon inlärning. När N1

första gången uppstår ger det omedelbart upphov till N2.) Vi kan vidare tänka oss att det är väl belagt och allom bekant att Mary är beskaffad på detta vis, och att bl.a. hon själv känner till det. Det förefaller klart att Klärvoajanta Mary kan på ett riktigt sätt tillämpa termen 'blått' på basis av sina synupplevelser: hennes instinkter rörande termen är helt korrekta -- hon är benägen att tillämpa 'blått' på blått och inte på någonting annat än blått -- och med tanke på hennes bakgrundsinformation har hon all anledning att lita på dessa instinkter (hon vet att den typiska orsaken till hennes benägenhet att tillämpa 'blått' är att hon upplever blått). Och därmed kan Klärvoajanta Mary, i vår aktuella mening, *känna igen* blått. Men det verkar klart att hon kan ha denna förmåga utan att kunna visualisera blått, eller förstå hur blått sensoriskt framträder; när hon först presenteras med blått kan hon förväntas säga till sig själv något i stil med: *aha, blått ser ut sådär*.

Sammanfattningsvis föreslår jag alltså att det finns följande *oberoenden* mellan de tre aktuella förmågorna: (a) Man kan förstå hur blått ser ut utan att kunna avsiktligt visualisera blått. (b) Man kan förstå hur blått ser ut utan att kunna känna igen blått. (c) Man kan ha förmågan att avsiktligt visualisera blått utan att kunna känna igen blått. (d) Man kan ha förmågan att känna igen blått utan att förstå hur blått ser ut. (e) Man kan ha förmågan att känna igen blått utan att kunna visualisera blått. Till detta kommer följande möjliga beroende mellan två av förmågorna: Man kan inte ha förmågan att avsiktligt visualisera blått utan att förstå hur blått ser ut.

(vi) Vid närmare reflektion verkar det rimligt att avvisa alla *okvalificerade icke-konstitutiva* versioner av fenomenell begreppsempirism. Klärvoajanta Mary illustrerar att det i princip är möjligt att känna igen blått utan att någonsin ha upplevt blått. Det verkar likaledes vara i princip möjligt att förstå och ha förmågan att visualisera blått utan att någonsin ha upplevt blått. Tänk t.ex. på följande Mary-scenario (från Dennett 2005, kap. 5, som är inspirerad av Davidson 1984; liknande fall diskuteras i Unger 1966). I Dennetts exempel slår blixten ned i Marys svart-vita rum. Av en händelse orsakar detta att Marys hjärna omstruktureras så att den hamnar i precis det tillstånd den *skulle* ha varit i om Mary hade varit fri och haft rikliga erfarenheter av kromatiska färger. Vi kan anta att, under en viss period efter blixtnedslaget har Mary fortfarande inte upplevt några kromatiska färger -- hon är fortfarande instängd i rummet, och har händelsevis inte visualiserat några kromatiska

färger -- men trots det har hon förmågan att känna igen och avsiktligt visualisera blått, och hon inte är mer benägen än du eller jag att säga till sig själv när hon ser blått: aha, blått ser ut sådär.

(vii) Å andra sidan kan vi inte på dessa grunder utesluta några av de versioner av *Antropologisk icke-konstitutiv fenomenell begreppsempirism* som vi har urskiljt. Trots det som har sagts ovan kan det mycket väl vara en *psykologisk lag* för oss människor att, om någon av oss inte har upplevt blått så kan hon inte förstå eller veta hur blått ser ut, inte visualisera blått, och inte på ett riktigt sätt tillämpa termen 'blå' på basis av synupplevelser. Det är bara i extraordinära fall, som fallet med blixten som omstrukturerar Marys hjärna på ett mirakulöst sätt, som man kan tillägna sig dessa förmågor utan att ha haft den relevanta upplevelsen. I realiteten är detta dock inte mer möjligt än att lära sig tala japanska utan att ha varit i kontakt med det japanska språket (jfr. sektion 2.2 ovan).

D.v.s., utifrån vad som hittills har sagts kan Antropologisk icke-konstitutiv fenomenell begreppsempirism mycket väl vara riktig. I själva verket kan alla tre versioner av denna tes vara riktiga -- den imagistiska, den intellektualistiska, och den igenkänningsbaserade.

Humes fall med den utelämnade nyansen av blått, som vi nu ska återbesöka, är emellertid en utmaning för dessa teser.

3 Humes fall

Tänk igen på en människa som har sett alla färger utom ljusblått. Det skulle knappast vara förvånande eller exceptionellt om hon, på basis av sin bekantskap med de andra färgerna och den bakgrundsinformation vi kan förse henne med om hur ljusblått liknar och skiljer sig från dessa -- t.ex. att ljusblått förhåller sig till blått som ljusgrönt till grönt -- kunde förstå hur ljusblått ser ut, tillämpa 'ljusblått' på ljusblått när hon ser det, och t.o.m. visualisera ljusblått. I själva verket tycker jag det verkar troligt att det är på detta sätt vi faktiskt, typiskt sett utvecklas: när vi har upplevt en tillräckligt varierad

uppsättning färger, så kan vi förstå hur resten av färgerna ser ut, tillämpa rätt färgord på dem, och visualisera dem.

Det kan nu förefalla som att Humes fall falsifierar våra tre antropologiska begreppsempiristiska teser. Enligt de tre teserna är det en psykologisk lag för oss människor att (i) vi kan visualisera en enkel färgkvalitet bara om vi har presenterats med den, (ii) kan förstå hur en enkel färgkvalitet ser ut bara om vi har upplevt den, respektive (iii) kan tillämpa rätt färgord på enkla färgkvaliteter bara om vi har upplevt dem. Men Humes fall visar att vi troligen är högst förmögna att visualisera enkla färgkvaliteter som vi inte har presenterats med, förstå hur enkla färgkvaliteter ser ut trots att vi inte har upplevt dem, och tillämpa färgord på enkla kvaliteter som vi inte har upplevt. Eller?

4 Attityder till Humes fall

Det finns tre möjliga attityder som man kan inta till Humes fall i förhållande till en eller annan version av begreppsempirism.

4.1 Fallet är ett ödesdigert motexempel

En möjlighet är att hävda att fallet är ett motexempel till den relevanta begreppsempiristiska tesen, och att tesen därför bör avvisas. Mitt intryck från diskussioner om ämnet är att denna attityd är ganska spridd. Såvitt jag har kunnat se är den inte så ofta satt på pränt, men även det förekommer. T.ex. säger Torin Alter i en nylig uppsats:

Kräver kunskap om hur det är att ha en upplevelse med den fenomenella karaktären X att man har haft en upplevelse med X? Nej. Ett berömt motexempel är Humes utelämnade nyans av blått, i vilket det är möjligt att extrapolera från fenomenellt liknande upplevelser (Alter, under utgivning).

Från kontexten förefaller det klart att Alter här menar att (i) fallet med den utelämnade nyansen av blått är ett motexempel till en viss typ av begreppsempirism - - specifikt någon version av Intellektualistisk fenomenell begreppsempirism -- och (ii) denna begreppsempirism kan därför avvisas.

4.2 Fallet är ett motexempel, men inte ödesdigert

En annan möjlighet är att medge att Humes fall är ett motexempel till en viss typ av begreppsempirism, men hävda att teorin ändå inte behöver avvisas. En variant av det här svaret är att fallet är, inte ett *möjligt*, utan bara ett *tänkbart* motexempel till den aktuella teorin. Någoting liknande denna attityd har antytts av John Nelson i förhållande till Humes begreppsempirism:

Även om Hume talar om ett *kontradiktoriskt fenomen*, så talar han bara om ett *tänkbart* kontradiktoriskt fenomen (Nelson 1989, 356).

En annan variant av det här svaret är att fallet är ett möjligt motexempel, men att det är *exceptionellt* och därför inte falsifierar en viss begreppsempirism. En sådan attityd kan intas t.ex. av någon som försvarar en nomologisk variant av begreppsempirism -- t.ex. en antropologisk tes -- och som accepterar att lagar medger vissa undantag (jfr. sektion 2.2 ovan). Det är möjligt att detta var Humes egen attityd. Humes -- ofta citerade och omdiskuterade -- slutsats av fallet med den utelämnade nyansen av blått var att:

detta fall är så speciellt och säreget att det knappt är värt vår uppmärksamhet, och motiverar inte att vi för dess skull ändrar vår allmänna maxim [d.v.s. att enkla idéer alltid är härledda från enkla intryck] (Hume 1739, bok 1, del 1, sektion 1).

Av skäl som jag redan har redovisat uppfattar jag alla varianter av den här attityden som orimliga. Som jag betonade i sektion 3 förefaller det troligt att fallet med den utelämnade nyansen av blått är varken omöjligt eller exceptionellt, utan en del av vår *faktiska, typiska* mentala utveckling.

4.3 Fallet är inte ett motexempel

En tredje möjlighet är att hävda att Humes fall trots allt inte är ett motexempel till en given typ av begreppsempirism.

Det finns ett rätt stort antal förslag med denna inriktning i litteraturen. Många filosofer har argumenterat, i förhållande till en eller annan variant av begreppsempirism, att teorin kan modifieras på ett sådant sätt att den (i) medger Humes fall och samtidigt (ii) behåller den plausibilitet och det intressevärde den ursprungligen hade.

T.ex. har Don Garrett presenterat ett sådant förslag å Humes vägnar. Garrett menar att fallet inte är ett bekymmer för Humes begreppsempirism eftersom den utelämnade nyansen "ytterst härrör från en uppsättning väldigt *snarlika* impressioner" (1997, 52). Därigenom behöver inte Hume, enligt Garrett, oro sig över att den utelämnade nyansen ska kasta några tvivel över hans användning av begreppsempirism i behandlingen av teman som var mer centrala för honom, såsom frågorna om våra idéer om nödvändiga kausala förbindelser och substanser. I dessa fall kan det nämligen inte hävdas att de påstådda och kontroversiella idéerna skulle vara snarlika några av de impressioner som vi okontroversiellt har haft. Som jag förstår honom föreslår Garrett någonting i följande stil: Istället för att säga att varje enkel idé härstammar från en enkel impression som den exakt representerar skulle Hume lika gärna ha kunnat säga att varje enkel idé härstammar från en enkel impression som den exakt representerar *eller från en uppsättning väldigt snarlika impressioner*. Denna tes skulle vara förenlig med den utelämnade nyansen och ändå vara tillräckligt kraftfull för Humes vidare syften. Liknande förslag har för Humes räkning presenterats av Robert Cummins (1978), John Morreall (1982), och Robert Fogelin (1984). David Papineau (2002, 68) presenterar ett liknande förslag till försvar för hans egen fenomenella begreppsempirism.

Jag uppfattar att det ligger någonting i de här förslagen, men de förefaller mig vara underpreciserade. Det som behöver preciseras är exakt *hur* mycket likhet som behövs

mellan kvaliteter -- eller vilken typ av likhet som behövs -- för att ett medvetande ska kunna tillägna sig ett begrepp av en icke-upplevd kvalitet på basis av upplevelser av andra kvaliteter. Det finns trots allt en likhet mellan gult och rött. Med vissa måttstockar är likheten stor: gult och rött liknar varandra mer än någon av dem liknar en smärtförmimelse t.ex. Men det är knappast, typiskt sett, möjligt att härleda ett begrepp om gult på basis av en upplevelse av rött. Så exakt vilken likhet krävs mellan kvaliteter för att möjliggöra denna typ av "begreppshärledning"? En begreppsempiristisk teori borde ha resurser att ge ett principiellt svar på denna fråga. (Utan ett sådant svar är det bl.a. oklart om en begreppsempirist kan acceptera fallet med den utelämnade nyansen av blått utan att detta påverkar hennes användning av begreppsempirismen i diskussionen av andra frågor, som t.ex. begrepp för nödvändiga kausala förbindelser. För om vi kan härleda ett begrepp om den utelämnade nyansen av blått när vi har upplevt snarlika färger, varför skulle vi inte kunna härleda ett begrepp om nödvändig kausal förbindelse om vi har upplevt "snarlika" kontakter i tid och rum?)

Jag ska skissera en strategi som begreppsempirister kan använda för att precisera vilken likhet som krävs för att den typ av begreppshärledning som verkar ske i fallet med den utelämnade nyansen av blått ska vara möjlig.

Det verkar uppenbart att en begreppsempiristisk tes måste begränsa sig till att hävda att *enkla* begrepp (eller begrepp för enkla kvaliteter) är erfarenhetsberoende, och medge att ett medvetande kan kombinera dessa och därigenom tillägna sig komplexa begrepp (eller begrepp för komplexa kvaliteter). Det verkar t.ex. inte finnas något begrepp för röd-och-blå-randigt som vi människor kan ha bara om vi har upplevt något röd-och-blå-randigt; varje begrepp som vi kan tillägna oss genom att uppleva röd-och-blå-randigt kan vi rimligen också tillägna oss med hjälp av separata upplevelser av rött, blått, och randigt, och någon form av mental "kombineringsoperation".

Men vilka begrepp, eller kvaliteter, är enkla? När Hume diskuterar fallet med den utelämnade nyansen av blått så betraktar han varje färgnyans som en enkel kvalitet, som motsvaras av enkel "idé" (se t.ex. Hume 1739, sid. 637). Och det är rätt naturligt att tänka så. Men det är inte nödvändigt. Ett alternativ, som bygger på samtida fenomenologiska färgmodeller -- exempelvis det naturliga färgsystemet (The Natural

Color System) -- är att anta att det finns exakt sex enkla färger, nämligen svart, vit, röd, blå, grön, och gul, och att varje färgnyans är sammansatt av 1-4 av dessa (se t.ex. Hård och Svedmyr 1995). En ren röd kan t.ex. betraktas som fenomenologiskt sammansatt av enbart rödhet, nyanser av grått som fenomenologiskt sammansatta av vithet och svarthet, medan nyanser av gråaktig-orange kan betraktas som fenomenologiskt sammansatta av rödhet, gulhet, vithet och svarthet.

Med utgångspunkt i en sådan modell kan t.ex. en antropologisk, icke-konstitutiv fenomenell begreppsempirist begränsa sig till att hävda att det, *för var och en av de fenomenologiskt enkla färgkvaliteterna*, finns ett begrepp som en människa av psykologisk nödvändighet kan ha bara om hon har upplevt denna kvalitet. En sådan tes skulle medge att vi kan härleda begrepp för fenomenologiskt sammansatta färgnyanser om vi har upplevt de enkla färgkvaliteter som dessa är sammansatta av.

Denna teori är fortfarande underpreciserad. Man kan t.ex. fråga sig vad det innebär att *uppleva en fenomenologiskt enkel färgkvalitet*. Om Jacksons Mary presenteras med en gråaktig-orange färgnyans, i vilken rödhet enligt vår modell skulle vara en fenomenologisk komponent, har hon då upplevt rödhet? Eller upplever hon rödhet bara om hon presenteras med en ren röd? Jag avstår från att diskutera dessa komplikationer här. Även om en del arbete återstår tror jag det går att illustrera att teorin är lovande genom att betrakta några scenarier som teorin under *alla* preciseringar korrekt tillåter respektive korrekt utesluter.

Teorin tillåter -- som den bör (jfr. sektion 3) -- fallet med den utelämnade nyansen av blått. Om en människa har upplevt alla färgnyanser utom nyanser av ljusblått så är det okontroversiellt att hon har upplevt både vitt och blått. Från dessa upplevelser kan vi anta att hon har förvärvat fenomenella begrepp för vitt och blått. Genom att kombinera dessa i sitt medvetande kan hon förse sig med ett begrepp om ljusblått. På liknande sätt medger teorin att en människa kan förse sig med ett begrepp om t.ex. gråaktig-orange på basis av upplevelser av vitt, svart, rött och gult.

Teorin utesluter å andra sidan att en människa kan förvärva ett begrepp för t.ex. gult på basis av upplevelser av vitt och rött -- förutom i exceptionella fall. Och det verkar också plausibelt, under alla de sätt att förstå fenomenella begrepp som vi har

betraktat: det är rimligen bortom normala mänskliga förmågor att härleda en förståelse av hur gult ser ut på basis av en förståelse av hur vitt och rött ser ut, eller att utveckla en förmåga att visualisera gult på basis av förmågor att visualisera rött och vitt, eller att härleda en förmåga att känna igen gult på basis av förmågor att känna igen rött och vitt.

4.4 Är alla färgnyanser "kvalitativt enkla"?

Det förslag jag har skisserat för begreppsempirismens räkning har -- i mina ögon -- någonting ganska uppenbart över sig. Men det är inte allmänt accepterat. Såvitt jag vet har det inte tidigare försvarats i skrift. Till en del kan detta bero på att man ofta helt enkelt inte har uppmärksammat möjligheten av förslaget. Men det finns även fall där förslaget -- eller någonting mycket snarlikt -- har uppmärksammats men avvisats. Jag ska titta på en invändning till förslaget i denna sektion. Därefter ska jag, i nästa sektion, reda ut en sammanblandning som också kan bidra till (ogrundat) tvivel på förslaget.

I en opublicerad uppsats överväger Charles Johnson, å Humes vägnar, vad han kallar "primärfärgssvaret" till fallet med den utelämnade nyansen av blått. Detta svar postulerar

någon viss uppsättning fundamentala färger, från vilka John kan skapa en nyans i sin fantasi på samma sätt som han skulle göra med målarfärg: genom att blanda färger på sin palett till den rätta balansen av primärfärgsidéer (2005, 5).

Johnson avvisar detta svar med motiveringen att alla färgnyanser är kvalitativt enkla:

när jag ser turkos ser jag en unik, *solid* färg. ... Primärfärgssvaret förvanskar den *kvalitativt enkla* upplevelse som vi har av färger (2005, 6).

Johnson verkar betrakta som ett oavvisligt fenomenologiskt datum att varje färgnyans är "kvalitativt enkel". Jag tror att det är rätt vanligt att tänka så. Men det förefaller

mig väldigt tvivelaktigt att anta detta som ett oavvisligt datum. Såvitt jag själv kan bedöma så framstår turkos i min perception som en kvalitativ blandning av blåhet och grönhet. Och jag är knappast ensam om det. Det är precis denna typ av omdöme som ger stöd åt fenomenologiska färgmodeller: färgseende människor, även teoretiskt sofistikerade sådana, verkar helt förmögna att karakterisera alla färgnyanser -- inklusive olika nyanser av turkos, orange, och grått -- uteslutande i termer av vithet, svarthet, rödhet, gulhet, grönhet, och blåhet (se Hardin 1993, 42, och Hård och Svedmyr, kapitel 2).

4.5 Fenomenologiska, och andra, färgmodeller

Jag har stött på följande typ av invändning mot det förslag som jag har framfört för begreppsempirismens räkning: (i) Enligt förslaget kan en människa typiskt sett tillägna sig ett begrepp för en färgkvalitet Q_B om Q_B är en blandning av två enkla kvaliteter Q_{E1} och Q_{E2} som hon har upplevt. (ii) Grönt är en blandning av blått och gult. (iii) Alltså förutsäger förslaget att en människa typiskt sett kan tillägna sig ett begrepp om grönt på basis av upplevelser av gult och blått. (iv) Men (iii) är falsk, och alltså bör förslaget avvisas.

Invändningen baserar sig dock på en förväxling av olika modeller för färgblandning. Den modell som jag har hänvisat till är *fenomenologisk*. Modellen använder sig av en uppsättning *fenomenologiskt* grundläggande färger, and klagör hur enskilda nyanser av färg är *fenomenologiskt* sammansatta av en, två, tre eller fyra av dessa. Modellen syftar till att systematisera hur färger visuellt framträder, och den måste skiljas från två andra -- någorlunda bekanta -- färgmodeller, som syftar till att systematisera andra saker.

Så kallade *additiva* färgmodeller syftar till att systematisera de sätt på vilka man kan framställa färgnyanser genom att blanda monokromatiskt ljus i olika proportioner. Modellerna är användbara när vi t.ex. vill framställa färger på TV- eller datorskärmar, och de använder ofta röd, grön och blå som primärfärger. En sådan modell säger t.ex. att vi kan framställa gult på en skärm genom att projicera lämpliga proportioner av det monokromatiska ljus som ger upphov till grönt när det ensamt projiceras på en skärm

("grönt ljus") tillsammans med det monokromatiska ljus som ger upphov till rött när det ensamt projiceras på en skärm ("rött ljus"), och att vi kan framställa vitt på en skärm genom att projicera lämpliga proportioner av "grönt ljus", "rött ljus" och "blått ljus".

Subtraktiva färgmodeller syftar till att systematisera de sätt på vilka färgnyanser kan framställas från en begränsad uppsättning pigment. Sådana modeller är användbara när vi t.ex. vill producera färgskrivare. De använder ofta cyan, magenta, gult, och ibland svart som primärfärger. En sådan modell säger t.ex. att vi kan framställa rött tryck på ett papper genom att blanda lämpliga proportioner av magenta och gult pigment, och grönt tryck genom att blanda gult och cyan.²

Det är uppenbart att dessa modeller har olika syften, gör olika påståenden, och att de kompletterar snarare än konkurrerar med varandra. T.ex. är det uppenbart att additiva och subtraktiva färgmodeller inte syftar till att karakterisera hur färger visuellt framträder. Vitt framträder inte som en blandning av blått, rött, och grönt, även om vi kan framställa vitt genom att blanda blått, rött och grönt ljus. Inte heller framträder grönt som en blandning av cyan och gult även om vi kan framställa grönt genom att blanda cyan och gult pigment.

Med den här bakgrunden kan vi se att invändningen i början på denna sektion grundar sig på en förväxling av färgmodeller. Det aktuella förslaget är att en människa typiskt sett kan tillägna sig ett begrepp för en färgkvalitet Q_B om Q_B är en blandning av två *fenomenologiskt* enkla kvaliteter Q_{E1} och Q_{E2} som hon har upplevt. Grönt är inte en fenomenologisk blandning av blått och gult. Alltså gör inte förslaget den felaktiga förutsägelsen att en människa typiskt sett kan tillägna sig ett begrepp för grönt på basis av upplevelser av gult och blått.

² Det finns en hel del bra information på nätet om additiva och subtraktiva färgmodeller, och relationerna mellan dem. En interaktiv sajt som illustrerar ett vitt spektrum av möjliga additiva färgblandningar är <http://javaboutique.internet.com/ColorFinder/>. Jag känner inte till någon lika utförlig illustration av subtraktiv färgblandning. Men några illustrationer erbjuds på <http://home.att.net/~RTRUSCIO/MIXITUP.htm>. En interaktiv sajt som illustrerar fenomenologiska färgblandningar är <http://www.i2studios.com/vis/ncs/>.

4.6 Avslutande reflektion: Vår beskaffenhet och andra möjligheter

Man kan fråga sig, avslutningsvis, *varför* jag för begreppsempirismens räkning har använt mig av en fenomenologisk färgmodell snarare än någon annan modell. Svaret är att denna modell -- mer händelsevis än nödvändigtvis -- i det här sammanhanget leder till de mest plausibla förutsägelseerna.

Det är rimligen i princip *möjligt* att någon tänkare skulle vara mentalt beskaffad på så vis att hon kunde härleda färgbegrepp ifrån sin erfarenhet i enlighet med additiva eller subtraktiva färgmodeller. T.ex. verkar det möjligt att en tänkare skulle ha ett "additivt färgmedvetande" såtillvida att, om hon genom erfarenheten hade bekantat sig med rött och grönt så skulle hon automatiskt förvärva ett begrepp om gult -- där detta begrepp kunde förstås imagistiskt, intellektualistiskt eller som igenkänningsbaserat. På samma sätt verkar det möjligt att en tänkare skulle ha ett "subtraktivt färgmedvetande" och typiskt sett kunna förvärva exempelvis ett begrepp om grönt på basis av upplevelser av cyan och gult. (Om man tycker att detta verkar omöjligt så är det värt att påminna sig de olika möjliga scenarierna kring Mary, i sektion 2.4, punkterna (v) och (vi)).

Våra medvetanden förefaller emellertid inte vara på detta sätt additiva eller subtraktiva. Snarare är vi -- mer tillfälligtvis än nödvändigtvis -- konstruerade så att vi förvärvar förståelse av hur färger ser ut, samt förmågor att visualisera dem och känna igen dem, mer eller mindre i enlighet med fenomenologiska färgmodeller.

Litteratur

Alter, Torin (under utgivning) "Phenomenal knowledge without experience", online på <http://www.as.ua.edu/philos/talter/PKWE.pdf>.

Cummins, Robert (1978) "The missing shade of blue", *The Philosophical Review*, 87, 548-565.

Dennett, Daniel (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*, Cambridge, Mass.: MIT Press.

Fogelin, Robert (1984) "Hume and the missing shade of blue", *Philosophy and Phenomenological Research*, 45, 263-271.

- Garrett, Don (1997) *Cognition and Commitment in Hume's Philosophy*, Oxford: Oxford UP.
- Hardin, C.L. (1993) *Colors for Philosophers: Unweaving the Rainbow*, Indianapolis: Hackett.
- Hume, David (1739/1978) *A Treatise of Human Nature*, red. P. Nidditch, Oxford: Oxford UP. Jag har delvis följt den svenska översättningen *Avhandling om det mänskliga förståndet: Bok 1: Om Förståndet*, övers. R. Callergård, Stockholm: Thales 2002.
- Hård, Anders, och Åke Svedmyr (1995) *Färgsystemet NCS: Tanke, tillkomst, tillämpning*, Stockholm: Bygghälsningsrådet.
- Johnson, Charles (2005) "How I got the blue: Cognitive complexity, epistemic simplicity, and the missing shade of blue in Hume's skeptical empiricism", opublicerat manuskript, online på <http://radgeek.com/gt/2004/10/29/HowIGotTheBlue.pdf>.
- Levin, Janet (2006) "What is a phenomenal concept?", *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, red. T. Alter och S. Walter, Oxford: Oxford UP.
- Loar, Brian (1990/1997) "Phenomenal states", ursprungligen publicerad i *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*, J. Tomberlin (ed.), Atascadero, CA: Ridgeview, 1990. Återutgiven i reviderad form i *The Nature of Consciousness: Philosophical Debates*, N. Block et al. (eds.), Cambridge, Mass.: MIT Press, 1997.
- Morreall, John (1982) "Hume's missing shade of blue", *Philosophy and Phenomenological Research*, 42, 407-415.
- Nelson, John (1989) "Hume's missing shade of blue re-viewed", *Hume Studies*, 15, 353-364.
- Papineau, David (2002) *Thinking about Consciousness*, Oxford: Oxford UP.
- Tye, Michael (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, Cambridge, Mass.: MIT Press.
- (2000) *Consciousness, Color and Content*, Cambridge, Mass.: MIT Press.
- (2003) "A theory of phenomenal concepts", *Minds and Persons*, red. A. O'Hear, Cambridge: Cambridge UP.
- Unger, Peter (1966) "On experience and the development of the understanding", *American Philosophical Quarterly*, 3, 48-56.