Upplevelse-maskinen

—så formar och förutsäger hjärnan vår verklighet

Andy Clark

Översättning: Gösta Svenn och Helena Sjöstrand Svenn

fri tanke

FRI TANKE
Stockholm, 2026
Box 3020, 181 03 Lidingö
www.fritanke.se | info@fritanke.se

FRI TANKE *är ett förlag för vetenskap, filosofi, kultur och samhälle. Vi verkar för bildning och demokrati genom att lyfta fram viktiga natur- och samhällsvetenskapliga frågor i upplysningens anda.*

ISBN 978-91-8973-278-0
Copyright © 2023 by Andy Clark. All rights reserved.
© svensk utgåva Fri Tanke, 2026
ORIGINALTITEL *The Experience Machine: How Our Minds Predict and Shape Reality*
ÖVERSÄTTNING Gösta Svenn och Helena Sjöstrand Svenn
SAKGRANSKNING Hans Liljenström
FORMGIVNING OMSLAG Studio Liljemärker
SÄTTNING Magnus Frederiksen
TYPSNITT Sabon Next
TRYCK Livonia Print, Riga, 2026
Första tryckningen

Innehåll

Förord: Att forma upplevelser 11
1. Dags att packa upp förutsägelsemaskinen 19
2. Psykiatri och neurologi 53
3. Handling som självuppfyllande profetia 93
4. Att förutsäga kroppen 113
Mellanspel 143
5. Förvänta sig bättre 159
6. Bortom den nakna hjärnan 179
7. Att hacka förutsägelsemaskinen 217

Slutsatser 251
Appendix 259

Författarens tack 273
Noter 275
Bildrättigheter 337

Till Alexa och Gill, precis

För vad man hör och ser beror rätt mycket på var man står;
det beror också på vad man är för slags människa.

C.S. Lewis, *Min morbror trollkarlen*
Övers. Britt G. Hallqvist

FÖRORD:

Att forma upplevelser

FÖR NÅGRA ÅR SEDAN var jag inbjuden att tala vid en populärvetenskaplig festival som hölls i London. Jag är professor i kognitiv filosofi (en märklig titel inom ett tämligen eklektiskt intresseområde som spänner över filosofi, neurovetenskap, psykologi och artificiell intelligens) och skulle hålla ett föredrag om ett av mina favoritämnen: den mänskliga hjärnan som en »förutsägelsemaskin«. Festivalen drivs av en populärvetenskaplig tidskrift och heter New Scientist Live. Varje år bjuder *New Scientist* (tidskriften) in experter på många olika områden att göra offentliga framträdanden. Det året arrangerades festivalen i det enorma konferenscentret ExCel i London Docklands. Att kliva in i ExCel var som att stiga ombord på flera sammankopplade oceanångare, där det i var och en arrangerades ett eget storskaligt evenemang. Som universitetsprofessor är jag inte ovan vid att tala inför publik. Men när jag stod bakom scenen i en av de större föreläsningssalarna och tänkte på den fullsatta salongen bakom ridån, blev jag ändå skakig. Kanske borde jag ha gjort några sista minuten-ändringar i bildspelet. Kanske borde jag ha tagit på mig en mindre uppseendeväckande skjorta. Hade jag glömt tacka någon? Plötsligt avbröts mina ängsliga tankar av att mobilen surrade till i fickan.

Fast den låg inte i fickan. Som jag raskt erinrade mig hade jag inte bara tagit upp mobilen och lagt på en hylla i talarstolen, jag hade för säkerhets skull också satt den på flygplansläge under

hela eventet. Men jag hade känt att det surrade till – ett tydligt och starkt surrande dessutom. Det som jag hade upplevt var ett alltigenom modernt fenomen, ett påfallande vanligt spratt från hjärnan som kallas »fantomvibrationer«. Eftersom jag under lång tid ständigt använt mobilen har min hjärna med tiden börjat förvänta sig att bli störd av ett surrande i fickan, och det är jag inte ensam om. En studie från 2012 fann att 89 procent av universitetsstudenterna uppgav att de kände fantomvibrationer från mobilen, och det har visat sig vara särskilt utbrett bland AT-läkare där falska mobilvibrationer är starkt förknippade med stress.* Termen blev »årets ord 2012« hos den australiska ordboken *Macquarie Dictionary*.

Det var rätt passande att fantomvibrationerna störde mig precis när jag skulle inleda min presentation. För även om sådana fenomen är välkända inom psykologin och neurovetenskapen, ingår de numera i en mycket mer omfattande teori som jag själv har varit med om att skapa under det senaste årtiondet. Enligt denna övergripande teori (som också var ämnet för min föreläsning) är fantomvibrationer bara en distinkt demonstration av hur all mänsklig erfarenhet är uppbyggd. Enligt den nya teorin, *predictive processing*, eller *prediktiv bearbetning*, är verkligheten så som vi upplever den uppbyggd av våra egna förutsägelser. Det var min invanda förväntan om ett surr i fickan kombinerat med den stressiga situationen som skapade en tydlig, men helt inbillad, vibration.

Prediktiv bearbetning berör en av de mest utmanande frågorna inom vetenskap och filosofi: hur relationen mellan vårt mentala sinne och verkligheten är beskaffad. Teorin, som stadigt har vunnit mark, förändrar vår förståelse av denna relation

---

* Referenser, belägg och exempel finns samlade i slutnoterna längst bak i boken, där de är ordnade efter relevant sidnummer och ett kort utdrag ur texten.

på sätt som har långtgående konsekvenser. I motsats till den gängse uppfattningen att våra sinnen är ett slags passivt fönster mot världen, framträder en bild av en ständigt aktiv hjärna som hela tiden strävar efter att förutsäga vad omvärlden för närvarande kan ha att erbjuda. Dessa förutsägelser strukturerar och formar sedan hela den mänskliga upplevelsen, alltifrån hur vi tolkar en persons ansiktsuttryck till våra upplevelser av smärta och våra planer för ett biobesök.

Ingenting som vi gör eller upplever – givet att teorin stämmer – förblir opåverkat av våra egna förväntningar. Istället pågår ett ständigt givande och tagande där det som vi upplever inte bara speglar vad omvärlden förmedlar till oss just nu, utan också vad vi – medvetet eller omedvetet – förväntade oss att den skulle förmedla. En konsekvens av detta är att vi aldrig bara ser det som »verkligen finns där«, helt utan våra egna förväntningar eller isolerade från våra egna tidigare erfarenheter. All mänsklig erfarenhet är delvis inbillad – resultatet av djupt rotade förutsägelser. Lika lite som vi kan surfa utan vågor kan vi uppleva världen fri från förutsägelser och förväntningar.

När jag stod bakom scenen på New Scientist Live-festivalen gjorde stressen – väntan på att hålla min föreläsning – att min förutsägelsemaskin gick på högvarv. Med tanke på min livserfarenhet trodde jag inte att golvet plötsligt skulle förvandlas till gelé under mina fötter eller att en stenbumling skulle ramla ner på mitt huvud som i en tecknad film. Men min mobil vibrerar faktiskt irriterande ofta i fickan, vilket får hjärnan att skapa ett slags grundförutsägelse om upprepade vibrationer. Stress och koffein (jag hade gott om båda) brukar förstärka sådana effekter, och signaler från en orolig mage matas direkt in i förutsägelsemaskineriet i våra huvuden. När alla de där faktorerna samverkade blev grundförutsägelsen om vibrationer i fickan min verklighet för en kort stund. Men lika snabbt som

det inträffade kunde jag omorientera mig mot fakta och inse att det var en illusion.

Illusionen uppstod eftersom förutsägande hjärnor är gissningsmaskiner som proaktivt förutser signaler från kroppen och omvärlden. Gissandet är inte bättre än de antaganden som det bygger på, och även en kvalificerad gissning missar ofta målet. Det fanns ju ingen mobil i min ficka. När hjärnans bästa gissning missar målet innehåller den bristande överensstämmelsen med den faktiska sensoriska signalen viktig ny information. Informationen (prediktionsfel) kan användas för att försöka igen – för att åstadkomma en bättre gissning om sakernas verkliga tillstånd. Men upplevelsen speglar fortfarande hjärnans nuvarande bästa gissning. Det är bara det att varje ny gissningsrunda är lite mer kvalificerad.

Detta utmanar en tidigare traditionell bild av vad perception är. Medan sensorisk information ofta ansetts vara utgångspunkten för våra upplevelser, pekar den framväxande vetenskapen om den förutsägande hjärnan på en lite annan roll. Nu används den aktuella sensoriska signalen för att förfina och korrigera den redan pågående processen ifråga om kvalificerade gissningar (försöken att göra förutsägelser). Nu är det förutsägelserna som gör det tyngsta jobbet. Enligt denna nya bild är upplevelserna – av världen, oss själva och till och med våra egna kroppar – aldrig en enkel återspegling av yttre eller inre fakta. Alla mänskliga upplevelser uppstår i mötet mellan initierade förutsägelser och sensorisk stimulans.

Detta är ett djupgående skifte i vår förståelse av sinnet som fundamentalt förändrar hur vi bör tänka om perception och konstruktionen av människans verklighet. Under större delen av mänsklighetens historia har forskare och filosofer sett perception som en process som främst fungerat »utifrån och in«, när ljus, ljud, beröring och kemiska lukter aktiverar receptorer

i ögon, öron, näsa och hud och successivt förfinas till en rikare bild av omvärlden. En bra bit in på 2000-talet höll ledande modeller inom både neurovetenskap och artificiell intelligens fortfarande fast vid grunddragen i det synsättet.

Den nya vetenskapen om prediktiv bearbetning vänder upp och ner på den traditionella bilden. Perceptionen formas nu i hög grad från motsatt håll, eftersom förutsägelser som bildas djupt inne i hjärnan når ut och förändrar reaktionerna i områden nära huden, ögonen, näsan och öronen – de sinnesorgan som tar emot signaler från omvärlden. Inkommande sensoriska signaler hjälper till att korrigera fel i förutsägelserna, men det är förutsägelserna som styr. Detta innebär att det vi uppfattar idag är djupt rotat i det vi upplevde igår och alla dagar dessförinnan. Varje aspekt av våra dagliga upplevelser filtreras genom dolda nät av förutsägelser – hjärnans bästa förväntningar som har sin grund i vår egen historia.

För att förstå hur viktigt detta faktiskt kan vara kan vi tänka oss en värld där väderprognosen spelar en avgörande roll, inte bara när det gäller att förutsäga vädret utan också förorsaka det. I en sådan märklig värld åstadkommer en prognos med hög sannolikhet för regn en förändring i de flöden av materia och energi som avgör hur vädret förändras. Där har regnprognosen en orsakande kraft som gör att sannolikheten för *själva regnandet* ökar. Där, liksom här, är väderprognosen beroende av en (aldrig perfekt) modell över hur rådande väderförhållanden mest sannolikt kommer att förändras och utvecklas över tid. Men i den där världen speglar vädret som blir (här och nu) ett slags kombination av den tidigare väderprognosens effekter och de rådande förhållandena ute i världen.

Nu lever vi inte i den där märkliga världen. Vårt väder påverkas inte av våra bästa modellbaserade väderförutsägelser. Men vår mentala värld delar något med den där anmärkningsvärda

profilen. När hjärnan kraftfullt förutspår en viss syn, ett visst ljud eller en viss känsla spelar den förutsägelsen en roll i utformandet av det vi tycks se, höra eller känna.

Känslor, humör och till och med planering baseras också på förutsägelser. Depression, ångest och trötthet speglar förändringar i de dolda förutsägelserna som formar vår upplevelse. Om vi förändrar förutsägelserna (till exempel genom att »stöpa om« en situation med hjälp av andra ord) så förändras själva upplevelsen. Tänk på det där pirrande adrenalinpåslaget jag fick precis innan jag gick upp på scen för att hålla föredraget. Jag hade tränat på att uppmärksamma den där pirrande känslan och samtidigt stöpa om den verbalt, inte som förebud om ett misslyckande utan som ett tecken på min kropps kemiska beredskap att åstadkomma ett bra framträdande. Detta hjälper mig att förändra mina förutsägelser om mig själv och leder till ett mer avslappnat och ledigt framträdande. Vi kommer att utforska flera sådana ingripanden och betona både deras överraskande omfattning och deras obestridliga gränser.

Hur är ditt förhållande till den verklighet du uppfattar? På vilka sätt formar du den och på vilka sätt formar du i vidare mening dig själv, ofta utan att ens vara medveten om det? I den här boken utgår jag från paradigmskiftande forskning för att ta itu med dessa mycket viktiga frågor och se vad insikterna betyder för neurovetenskapen, psykologin, psykiatrin, medicinen och för hur vi lever våra liv. Vi kommer att granska kroppens och jagets upplevelser, från kronisk smärta till psykos, och se hur arbetet med den förutsägande hjärnan bidrar till att förklara ett brett spektrum av mänskligt beteende och neurologisk mångfald. Vi kommer att omvärdera våra egna upplevelser av världen, från social fobi och känslomässiga feedbackloopar till alla de olika slags förutfattade meningar som kan smyga sig in i våra omdömen. Vi kommer också att utforska några

sätt på vilka vår förutsägande hjärna kan stötta »ett utvidgat sinnen«, vilket suddar ut gränserna mellan oss själva och våra bäst anpassade verktyg och miljöer.

Boken avslutas med att omsätta några viktiga insikter i handling, titta på hur man kan »hacka« den förutsägande hjärnan genom att förändra våra vanor, omforma våra upplevelser med hjälp av olika slags språk och genom kontrollerad användning av psykedeliska droger. När dessa teman möts skymtar vi konturerna av en ny och mer enhetlig vetenskap om sinnet – en vetenskap som gör rättvisa åt den mänskliga upplevelsens hela bredd och mångfald och som får verkliga konsekvenser för hur vi ska tänka om oss själva och förbättra våra liv.

1

Dags att packa upp förutsägelsemaskinen

DET ÄR MORGON och jag ligger fortfarande och sover. Bredvid sängen i sovrummet balanserar en rejäl trave med arbete ostadigt. När jag sömndrucken vaknar till hör jag stilla fågelsång. Eller tycker jag mig i alla fall höra. Men snart upptäcker jag att jag hade fel. Jag lyssnar noga och inser att det är dödstyst. Inte ens mina katters tidiga jamande efter mat bryter tystnaden. Jag hallucinerade fågelsången.

Som tur är finns det en enkel förklaring. Min partner bestämde sig nyligen för att göra det lättare att vakna på morgonen genom att använda en mobilapp som spelar upp fågelsång istället för ett traditionellt alarm. Appens väckarklocka börjar med ett milt kvittrande som sakta men säkert växer till något som liknar en hel morgonkör av fågelkvitter. Idag hade alarmet inte satt igång – det var alldeles för tidigt. Och ljudet av riktig fågelsång brukar inte kunna tränga igenom tvåglasfönstret heller. Men jag har blivit så van vid att vakna till det försiktigt tilltagande kvittret från väckarklockan att min hjärna börjat spela mig ett spratt. Nu märker jag att jag ofta vaknar långt innan alarmet drar igång, eftersom jag tycker mig höra hur det förinspelade kvittret svagt börjar spela.

Det här är äkta hörselhallucinationer orsakade av min nya, starka förväntan om att vakna till de diskreta fågellätena. Det ligger troligen inget illavarslande bakom min benägenhet att

hallucinera fågelsången. Man har länge vetat att både hörsel- och synhallucinationer lätt kan framkallas med rätt sorts träning. Men dessa, liksom en myriad andra fascinerande fenomen, har på senare tid visat sig vara tecken på något mycket större – något som är väsentligt för alla mänskliga upplevelser.

Idén (bokens huvudtema) är att människans hjärna är en förutsägelsemaskin. Hjärnan är ett utvecklat organ som bygger upp och omformar upplevelser utifrån en föränderlig blandning av förväntningar och faktiska sinnesintryck. Mina omedvetna förutsägelser om vad jag sannolikt skulle höra när jag vaknade gjorde, enligt den bilden, att min perceptuella upplevelse för en kort stund fördes i den riktningen, vilket skapade en kortlivad hallucination som snart korrigerades när mer information strömmade in genom mina sinnen. Denna nya information (som förmedlade avsaknaden av fågelsång) genererade »signaler om prediktionsfel« och det var – åtminstone i detta fall – allt som krävdes för att få min upplevelse att återigen stämma överens med verkligheten. Hallucinationen ersattes av en tydlig upplevelse av ett tyst rum. Men i andra fall kan, som vi kommer att få se, felaktiga förutsägelser bli fast förankrade och kontakten med verkligheten (i sig ett komplext och omtvistat begrepp) svårare att uppnå. Även när det inte förekommer några felaktigheter och vi ser saker och ting »som de är« spelar vår hjärnas förutsägelser fortfarande en central roll. Förutsägelser och felaktiga förutsägelser ses alltmer som den viktigaste valutan i den mänskliga hjärnan, och det är i deras skiftande jämvikt som alla mänskliga upplevelser tar form.

Boken handlar om denna jämvikt och en ny vetenskap som vänder upp och ner på mycket av det vi trodde att vi visste om hur vi uppfattar världen. Enligt denna vetenskap försöker hjärnan ständigt gissa hur tillståndet i världen (och i vår egen kropp) mest sannolikt är, utifrån vad den lärt sig från tidigare

erfarenheter. Allt jag ser, hör, vidrör och känner speglar – enligt denna nya vetenskap – dolda källor av förutsägelser. Om förväntningarna är tillräckligt starka eller (som i fågelalarmets tidiga kvitter) den sensoriska evidensen tillräckligt svårfångad, kan jag missuppfatta saker och ting och i praktiken ersätta delar av den verkliga sensoriska informationen med min hjärnas bästa gissning om hur saker och ting borde vara.

Detta betyder inte att framgångsrik varseblivning bara är en form av hallucination, även om mekanismerna är relaterade till dem som finns vid hallucinationer. Vi bör inte bagatellisera vikten av all den rika sensoriska informationen som når våra ögon, öron och andra sinnen. Men det ger en ny och annorlunda bild av processen att se – och att förnimma i största allmänhet. Det framställer den som en process som styrs av hjärnans egna bästa förutsägelser: förutsägelser som sedan kontrolleras och korrigeras med hjälp av sensoriska intryck som vägledning. När förutsägelsemaskineriet är igång blir perceptionen en process som inte bara struktureras av inkommande sensorisk information utan också av skillnaden – skillnaden mellan de faktiska sensoriska signalerna och de som hjärnan förväntade sig att ta emot.

Eftersom hjärnan aldrig bara »slås på« från noll – inte ens när jag vaknar på morgonen – är alltid förutsägelser och förväntningar igång och strukturerar proaktivt den mänskliga upplevelsen i varje ögonblick, varje dag. Enligt denna alternativa förklaring reagerar den varseblivande hjärnan aldrig passivt på omvärlden. Istället försöker den aktivt hallucinera världen, men kontrollerar hallucinationen mot de intryck som kommer in via sinnena. Med andra ord håller hjärnan ständigt på att måla en bild, och den sensoriska informationens roll är främst att justera penseldragen när de inte stämmer överens med inkommande signaler.

Denna nya förståelse av perceptionsprocessen har stor betydelse för våra liv. Den förändrar hur vi bör tänka om evidensen från våra egna sinnen. Den påverkar hur vi bör tänka om hur vi upplever våra egna kroppsliga tillstånd – smärta, hunger och andra upplevelser som ångest eller depression. För hur vi upplever våra kroppsliga tillstånd speglar också en komplex blandning av vad hjärnan förutsäger och vad de aktuella kroppssignalerna föreslår. Detta innebär att vi ibland kan förändra hur vi mår genom att förändra vad vi (medvetet eller omedvetet) förutsäger.

Detta betyder inte att vi helt enkelt kan »förutsäga oss själva bättre«, och inte heller att vi kan förändra våra egna upplevelser av smärta eller hunger hur vi vill. Men det tyder ändå på ett visst principiellt och kanske oväntat handlingsutrymme – ett utrymme som vi med omsorg och övning kan vända till vår fördel. Om vi hanterar det varsamt kan en bättre förståelse av förutsägelsens kraft förbättra hur vi tänker om våra egna medicinska symptom och presentera nya sätt att förstå psykisk hälsa, psykisk ohälsa och neurologisk mångfald.

Den smarta kamerans modell för seendet

Idén att hjärnan i grunden är en gigantisk förutsägelsemaskin är relativt ny. Tidigare trodde man att sensorisk information bearbetas på ett mestadels framåtriktat sätt – det vill säga att den tas emot av våra sinnen och dirigeras »framåt« in i hjärnan. För att ta det mest studerade exemplet registreras visuell information (som den äldre teorin antyder) först i ögonen och bearbetas sedan steg för steg djupare och djupare inuti hjärnan, som långsamt får fram mer och mer abstrakta typer av information. Om vi börjar med mönster av inkommande ljus

kan hjärnan först få fram information om enkla kännetecken som linjer, former och kanter, och sedan sätta ihop dessa till större och mer komplexa helheter. Jag kallar detta för den »smarta kamerans« sätt att se. Men det här är uppenbarligen ingen kamera, utan snarare ett mycket smart intelligent system. Likväl strömmar all information, precis som i en enkel kamera, mestadels inåt, från ögonen och vidare in i hjärnan. Först vid ett ganska sent skede i denna process kopplas minnen från livet och kunskap om världen in, vilket gör det möjligt för dig (den varseblivande) att förstå tillståndet i din värld.

Versioner av den smarta kamerans (framåtriktade) seende har haft stort inflytande inom filosofin, neurovetenskapen och AI. Ett sådant seende är intuitivt eftersom vi vanligtvis föreställer oss perception som ett flöde av information från omvärlden till sinnet. Denna bild återfinns till exempel i Descartes *Tractatus de homine* från 1662. Där beskriver Descartes perception som ett komplext öppnande och stängande av nätverk av inre rör som först präglar en bild av världen på sinnesorganen (som ögonen) och sedan djupare och djupare in i hjärnan via ett nätverk av små tunnlar. När intryck från omvärlden (och inifrån kroppen) strömmade framåt in i hjärnan, sades de ha bevarats i vårt sinne ungefär på samma sätt som information om fingrars form bevaras när man trycker in dem i vax.

Det blev aldrig klarlagt hur Descartes mekanism skulle fungera. Men även när mycket mer sofistikerad vetenskaplig kunskap utvecklats kvarstod grundtanken om den varseblivande hjärnan som ett relativt passivt organ som tar emot sensoriska intryck från omvärlden och sedan »bearbetar« dem på ett övervägande framåtriktat sätt (utifrån och in). Den idén var i stort sett standard inom kognitiv neurovetenskap i slutet av 1900-talet. Detta berodde troligen på att den kom att bli en ledande princip i David Marrs enormt inflytelserika datormodell av seendet.

Marr var en framstående person vars arbete inom neurovetenskap, datorseende och AI betraktas som några av de viktigaste bidragen som någonsin gjorts till kognitionsvetenskapen. Enligt Marrs beskrivning börjar den visuella bearbetningen med att grundläggande beståndsdelar i en inkommande signal detekteras – till exempel en ordnad matris av pixlar. Därifrån bygger den skiktade bearbetningen långsamt upp en mer komplex förståelse. I nästa steg kan man till exempel leta efter platser där pixelintensiteten uppvisar snabba förändringar jämfört med de omgivande pixlarna – vanligtvis en ledtråd till förekomsten av en gräns eller en kant därute i världen. Allteftersom bearbetningen fortskrider, steg för steg och allt djupare in i hjärnan, upptäcks ytterligare mönster, såsom de återkommande sekvenser som kännetecknar ränder. Seendet handlar här om att utsätta den obearbetade signalen för en serie operationer, såsom kant- eller randdetektering som långsamt avslöjar alltmer komplexa mönster i miljön – källan till den inkommande signalen. Så småningom kommer de komplexa detekterade mönstren i kontakt med kunskap och minne för att leverera (även om denna del av pusslet, avslöjande nog, aldrig har lösts på ett tillfredsställande sätt) ett slags 3D-bild av världen.

Marrs datormodell hade (liksom alla datormodeller) den tydliga fördelen att den specificerade de viktigaste beräkningarna som kunde vara inblandade i bearbetningens tidiga stadier, även om utformningen av de avgörande sista stegen förblev något av ett mysterium. Marrs modell var under många år standardbilden inte bara inom artificiellt seende utan även inom neurovetenskapen. Ända in på 2000-talet betraktades det visuella systemet främst som en framåtriktad analysator av inkommande sensorisk information i enlighet med Marrs beskrivning.

En annan påverkansriktning saknades emellertid i Marrs modell, nämligen en som gick bakåt, från djupet av hjärnan ner mot ögonen och andra sinnesorgan. Antalet nervcellskopplingar som på detta sätt överför signaler bakåt uppskattas överstiga antalet kopplingar som överför signaler framåt med en betydande marginal, på vissa ställen så mycket som fyra till ett. Vad gör alla dessa nedåtgående kopplingar som överför information från djupet av hjärnan till områden närmare de sensoriska periferierna? Dessa kopplingar går i motsatt riktning mot kopplingarna som behövs för att utföra de bearbetningsuppgifter som beskrivs i Marrs tidiga beräkningsmodell, men de når ändå ner till just de områdena.

Verkliga sådana neurala kopplingar är kostsamma att installera och underhålla. Hjärnan, som väger cirka 2 procent av människans kroppsvikt, uppskattas stå för cirka 20 procent av kroppens totala energiförbrukning. Den är vårt i särklass »dyraste« adaptiva tillbehör. Ändå vet man nu att en stor del av kostnaden går åt till att etablera och underhålla ett enormt nätverk av nedåtgående (och sidledes) kopplingar, som inte bara omfattar tidig visuell bearbetning utan också hela hjärnan. Detta är en gåta. Det var så gåtfullt att det fick den artificiella intelligensens pionjär Patrick Winston att så sent som 2012 fälla kommentaren att med så mycket information som uppenbarligen strömmar i den andra (nedåtgående) riktningen, ställs vi inför »en märklig arkitektur som vi nästan inte vet någonting om«. Det hela ser dock annorlunda ut när vi inser attraktionskraften i en djärv ny tes: att hjärnan inte är något annat än en storskalig förutsägelsemaskin.

Vända flödet

Den varseblivande hjärnans centrala funktionsprincip verkar vara raka motsatsen till den smarta kamerans seende. Istället för att ständigt förbruka stora mängder energi på att bearbeta inkommande sensoriska signaler, utgörs huvuddelen av hjärnans arbete av att lära sig och upprätthålla en sorts modell av kroppen och världen – en modell som ögonblick för ögonblick kan användas för att försöka förutsäga den sensoriska signalen. Förutsägelserna bidrar till att strukturera allt vi ser, hör, vidrör och känner. Det var de som var i arbete när jag hörde fågelsång som inte fanns på morgonen. Det var de som var i arbete när jag kände fantomvibrationer från en mobiltelefon som inte ens fanns i min ficka. Men som vi kommer att se är det också de som är i arbete när jag hör verklig fågelsång, känner verkliga mobilvibrationer och ser de olika föremålen som ligger utspridda på mitt skrivbord på universitetet.

En förutsägande hjärna är ett slags ständigt pågående simulering av världen omkring oss – eller åtminstone den värld som är viktig för oss. Inkommande sensorisk information används för att garantera modellens användbarhet – genom att jämföra förutsägelsen med det sensoriska intrycket och generera en felsignal när de två inte stämmer överens. Trots de stora kostnaderna för kopplingen innebär konstanta förutsägelser många effektivitetsvinster, som vi strax ska se. Ännu viktigare är kanske att det också gör oss flexibla och att vi kan anpassa våra reaktioner på ett sätt som speglar kraven i de uppgifter och sammanhang som är aktuella. Istället för att hela tiden ta fram en rik bild av världen ur en strid ström av sensoriska ledtrådar, är utgångspunkten den rika, föränderliga bilden av världen, och den sensoriska informationen används för att testa, undersöka och finjustera den bilden. Innan nya sensoriska

signaler anländer är den förutsägande hjärnan redan upptagen med att måla en rik bild av hur saker och ting sannolikt ser ut.

Detta förklarar i stora drag behovet av alla de där nedåtgående kopplingarna. De för med sig förutsägelser djupt inifrån hjärnan och skickar dem vidare till de sensoriska perifera områdena. Det förklarar också den enorma energiförbrukning som krävs enbart för att upprätthålla hjärnans inneboende aktivitet. Den aktiviteten är nödvändig för att bevara modellen som i varje ögonblick levererar förutsägelser. När hjärnan möter ny sensorisk information är dess uppgift att avgöra om det finns något i den inkommande signalen som ser ut som en viktig »nyhet« – oförutsedd sensorisk information som är betydelsefull för det vi försöker se eller göra. Det råder ökande enighet om att det i princip är så här våra hjärnor bearbetar sensorisk information. Under de senaste tio till femton åren har hypotesen analyserats och det har skett en explosion av forskning inom beräkningsorienterad och kognitiv neurovetenskap som nu ger en detaljerad och testbar förklaring till detta, och därmed löser mysteriet med Winstons »märkliga arkitektur«. Forskningen går under olika namn som »prediktiv bearbetning«, »hierarkisk prediktiv kodning« och »aktiv inferens«. Jag kommer främst att hålla mig till »prediktiv bearbetning« som en praktisk benämning på denna familj av teorier.

Enligt detta synsätt var den smarta kamerans bild av perception ett stort misstag. Trots dess intuitiva attraktionskraft är det rätta sättet (för det mesta) inte att betrakta perception som en process som främst går från ögonen och andra sinnesorgan inåt. Det är inte heller så att hjärnan bara tålmodigt väntar på att sensorisk information ska anlända. Däremot förutser den aktivt sensorisk information och använder allt den vet om mönster och föremål i världen – fåglars kvitter (och min partners alarm tidigt på morgonen), de alltför frekventa stör-

ningarna från mobilens vibrationer och hur de olika föremålen på mitt skrivbord brukar vara placerade. Hjärnan använder sig också hela tiden av den aktiva kroppen, huvudet som rör sig, ögonen och kroppsdelarna på sätt som samlar in ny och bättre information. Istället för att vara en passiv mottagare och bearbetare av sensorisk information är en sådan hjärna en outtröttlig siare (och, som vi senare ska se, en skicklig och aktiv utfrågare) beträffande sina egna sensoriska flöden.

Dåliga radioapparater och kontrollerade hallucinationer

Den moderna bilden av den förutsägande hjärnan har sina historiska rötter i idéer som den tyske fysikern och mångsidige vetenskapsmannen Hermann von Helmholtz framförde på 1800-talet. Helmholtz uppfann oftalmoskopet som används av optiker för att undersöka ögat och formulerade lagen om energins bevarande. Han var också intresserad av teorier om perception och hävdade att vi uppfattar världen endast tack vare ett slags omedvetna resonemang eller slutledningar där hjärnan frågar sig: »Med tanke på allt jag vet, hur måste världen se ut för att jag ska kunna ta emot det nu aktuella mönstret av signaler?« Det är den frågan som perceptionssystemen är byggda för att lösa.

Du kanske inte är medveten om hur vanligt detta är i vårt vardagsliv. Om du lyssnar på en välbekant låt på en radio med dålig mottagning hörs orden och rytmerna förvånansvärt tydligt. Men försök lyssna på en helt ny låt med samma mottagningskvalitet, så är ljuden mycket otydligare och sången svår att urskilja. I båda fallen använder din hjärna, precis som Helmholtz hävdade, det den vet för att försöka sluta sig till

vilka ord och ljud som troligast ligger bakom de något hackiga ljudsignaler som just nu tas emot av dina öron. Men hjärnans gissningar blir mycket bättre när det gäller den välbekanta låten, vilket gör att den hörs så mycket tydligare. Faktum är att gissningarna förändrar hjärnans reaktioner ända ner till de tidiga auditiva bearbetningsområdena, så att reaktionerna bättre stämmer överens med de förväntade ljuden. I praktiken spelar din hjärna nu upp en stor del av låten för sig själv, så att den dåliga inkommande signalen rättas till med hjälp av lagrad kunskap om omvärlden.

Detta är hjärnan i sitt esse, när den får ösa ur sig »bra hallucinationer« genom att fylla i och bygga ut det som saknas i signalen, i enlighet med vad den förväntar sig att höra. Vår hjärna känner till hur låten går och de olika nyanserna i just den sångarens tolkning, och den kan använda all tidigare kunskap för att aktivt förutsäga den mest troliga formen på de auditiva signalerna när låten spelas. Ifall ingen stark motsägande evidens kommer från omvärlden, formar förutsägelserna upplevelsen och gör att du hör låten tydligare.

Det är viktigt att betona att detta inte är ett spratt som minnet spelar, utan snarare ett fascinerande fönster mot hur perceptionen i sig fungerar. Hjärnans förutsägelser om den välbekanta låten hjälper den att mejsla fram signalen ur bruset, vilket gör att ljuden återges tydligare än vad den dåliga signalen annars skulle tillåta. Den sortens perception är mycket aktiv. Den medför att komplexa förutsägelser skickas nedåt i kedjan från högre bearbetningsområden till de sensoriska perifera områdena, vilket genererar felmeddelanden när en allvarlig avvikelse upptäcks. Detta bakåtriktade flöde kallar kognitionsforskare ibland för »top–down«-flöde av information. Medan allt detta pågår är den varseblivande människan också aktiv och försöker samla in viktiga bitar av sensorisk information

genom kroppsliga handlingar som att vrida på huvudet eller röra ögonen. Även dessa handlingar väljs och startas av den prediktiva mekanismen, vilket skapar en enhetlig väv av mental och kroppslig aktivitet. Vi kommer att ha mycket mer att säga om handlingens roll allteftersom vår berättelse fortskrider.

Att på detta sätt låta förutsägelser styra förvandlar den vanliga perceptionen till något som ibland färgstarkt beskrivits som »kontrollerad hallucination« – hjärnan gissar hur världen ser ut genom att använda sensoriska intryck främst som ett sätt att korrigera och förfina gissningarna. När de inre gissningarna helt och hållet styr hallucinerar vi, punkt slut. Men när de är tillräckligt känsliga för sensoriska stimuli – via signaler om prediktionsfel – kontrolleras gissningarna och sinnet känner igen världen. När vi hörde den välbekanta låten på radion med den dåliga mottagningen drog vi nytta av just en sådan »bra hallucination«. Mobiltelefonens fantomvibrationer som vi mötte i förordet var visserligen vilseledande men genererades på precis samma sätt. Om teorin om prediktiv bearbetning stämmer byggs alla mänskliga upplevelser upp på detta sätt. Vi ser världen genom att förutsäga den. Men när prediktionsfel uppstår måste hjärnan förutsäga på nytt.

Den sparsamma hjärnan

Att låta perceptionen styras av förutsägelser har också en annan viktig fördel. Det gör det möjligt för hjärnan att bearbeta inkommande sensorisk information på ett synnerligen effektivt sätt. Den disciplin som är mest känd för att undersöka frågan om kommunikationseffektivitet är informationsvetenskapen, som har spelat en viktig roll i utvecklingen av mycket sparsamma sätt att överföra signaler. Vid mitten av 1900-talet pressades

de globala telekommunikationssystemen av en ständigt ökande efterfrågan. Problemet för telekomjättarna var hur man skulle överföra allt större informationsmängder via de brusiga och begränsade kanaler som de gamla telefonkablarna erbjöd. Det var där informationsvetenskapen kom in med ett smart sätt att öka effektiviteten. Den kraftfulla teknik som utvecklades av teoretiker som arbetade med sådana gåtor kom så småningom att kallas linjär prediktiv kodning, och den används än idag.

Linjär prediktiv kodning har sina rötter i en uppsats som publicerades 1948 av Claude Shannon, en matematiker och kryptograf som arbetade på Bell Laboratories. Detta centrala arbete visade att engelsk text kunde kodas på ett mycket effektivt och komprimerat sätt genom att man utnyttjade förutsägbarheten hos ord och bokstäver. Om en viss bokstav nästan alltid följs av en annan, kan ett effektivt kodningsschema helt enkelt anta att så är fallet, såvida inte fallet markeras som ett undantag. Att endast markera dessa enstaka undantag är mycket effektivare och kräver mindre bandbredd än vad som skulle behövas för att koda varje bokstav.

Med rätt förväntningar kan även frånvaron av en signal innehålla en stor mängd information. Anta att du gör upp en plan genom att säga till någon att om du inte ringer, så är allt »som förväntat« och att personen därför ska möta dig på Miamis flygplats när du landar nästa onsdag klockan 9 på morgonen lokal tid. Om du låter bli att ringa den dagen är det liktydigt med en minimal signal, på endast en bit, som nu förmedlar en mycket stor mängd information: du kommer att anlända med flyg vid den tiden, till den platsen. Knepet är att byta mottagarens information och förkunskap mot kostnaderna för att koda och överföra all information.

Telekommunikationsforskningen har dragit nytta av många versioner av prediktionsbaserad komprimering. I princip kan

man använda allt som redan är känt om den mest sannolika skepnaden hos en viss signal för att lättare kunna förutsäga den signalen i andra änden, och via de dyrbara ledningarna och kablarna endast överföra det som visar sig *avvika* från de förutsagda mönstren. Mottagaren uppdaterar sedan enbart på basis av det så kallade residualfelet, det kvarvarande felet. Det fina med denna procedur är att något innehållsrikt (exempelvis en bild eller ett meddelande) kan rekonstrueras genom att endast några få felaktiga bitar överförs. Det innehållsrika skapas till största delen av förutsägelserna men förankras i verkligheten med hjälp av residualfelen.

Tack vare en initierad förutsägelse sparar komprimeringen på det sättet bandbredd genom att i praktiken »återföra« alla framgångsrikt förutsagda element. Det här är samma smarta knep som gör det möjligt för oss att ekonomiskt lagra och överföra bilder, ljud och videor med hjälp av format som JPEG och MP3. När det gäller bilder fungerar prediktiv kodning på så sätt att värdet på varje pixel antas kunna förutsägas utifrån värdet på dess olika grannar. När det stämmer – vilket faktiskt är ganska ofta – finns det helt enkelt inget behov av att överföra värdet för den pixeln. Allt som behöver kodas är avvikelserna från det som redan förutsagts. Men det där är bara en enkel regelbundenhet. Överallt där det finns någon sorts påvisbar regelbundenhet blir förutsägelser (och därmed denna form av datakompression) möjliga.

Ta rörelsekomprimerad kodning när det gäller video, till exempel. År 1959 introducerades världen för prediktiv interbild-kodning för video. För att få ett hum om detta kan man föreställa sig att videon visar en person som springer genom en korridor. Enda skillnaden mellan bildruta 4 och 5 i videon är en lätt framåtlutad rörelse hos löparen mot en oförändrad visuell bakgrund. Allt som behöver överföras för att fånga bildruta 5 är

dessa få skillnader (residualfelet) mot det som redan förutsagts. På detta sätt kan man behandla bildruta 5 som bara en mindre variant av bildruta 4. Kanske består den enda skillnaden i hur fötterna är placerade. Att skicka information om fotplaceringen genom systemet kostar betydligt mindre än att överföra ett nytt värde för varje pixel i bildruta 5. Det knepet används än idag.

Föreställ dig nu ett system som redan vet mycket mer – ett system som till exempel känner till alla sätt som olika löpstilar tenderar att fortsätta på. Ett sådant system skulle kunna använda den mer detaljerade informationen (informationen på högre nivå) för att också göra förutsägelser, så att endast oväntade fotrörelser skulle generera signaler om prediktionsfel. Om vi antar att fötterna rör sig precis som förväntat, finns det återigen inget behov av att uppdatera med ny information mellan bildrutorna. Detta ännu mer intelligenta system skulle kunna »hallucinera« den vanliga typen av pågående rörelse och bara uppdatera när något oväntat inträffar (kanske att löparen plötsligt snubblar till). Oavsett hur komplexa eller avancerade förutsägelserna är, så är det prediktionsfelen som måste förmedla nyheterna, signalera skillnader mot det förväntade och därmed hålla kontakt med en föränderlig och ibland överraskande värld.

Människans hjärna verkar dra nytta av just sådana intelligenta förutsägelsestrategier, och det på ett särskilt kraftfullt sätt tack vare användningen av flera »bearbetningsnivåer«. I en sådan kontext med flera nivåer bäddas enkla förutsägelser in under mindre enkla, mer abstrakta sådana – ungefär som i exemplet med löpstilar, där förväntningen på löpstilen är en förutsägelse på högre nivå som i sin tur ger upphov till förutsägelser om de faktiska fotplaceringarna (en förutsägelse på lägre nivå). Då skapas prediktionsfel som skjuts uppåt genom systemet. De nyanserar och förfinar gissningarna på varje nivå

– till exempel genom att avslöja att löpstilen trots allt inte är den vi förväntade oss och att vi bör välja ett alternativ som fungerar bättre.

I hjärnans arkitektur för prediktiv bearbetning anses olika nervcellspopulationer vara specialiserade på olika saker, så att varje »högre« nivå kan använda sin egen specialiserade kunskap och sina resurser för att försöka förutsäga tillstånden hos nivån direkt under den. Så, förenklat sett, en nivå som är specialiserad på att förutsäga hela ord kan använda sin kunskap för att bidra till att förutsäga tillstånd på en lägre nivå vars specialitet är att känna igen bokstäver. Men den högre nivån som förutsäger ord kan själv förutsägas av en ännu högre nivå som specialiserar sig på hela meningar. Ett förtydligande exempel på detta, tillsammans med en del andra förklaringar, finns i bilagan.

För tillfället är det viktigt att notera att i detta arrangemang med flera nivåer är allt som flödar framåt (från de sensoriska ytterområdena allt djupare in i hjärnan) *nyheter* – avvikelser från det förväntade. Detta är effektivt. Värdefull bandbredd används inte för att skicka mycket förutsägbara saker uppåt. Vem på företagets huvudkontor vill veta att Billys arbete fortskred precis som förväntat? På samma sätt signalerar prediktionsfel på varje nivå i det neurala förkroppsligandet endast det oväntade, det som rimligen kan kräva ytterligare tankearbete eller åtgärder.

Sådana system är fenomenalt sparsamma i sin användning av det inkommande informationsflödet. Istället för att försöka hantera allt från grunden sorterar och filtrerar de effektivt inkommande data genom att endast lyfta fram det oväntade. Detta är kärnan i uppfattningen att den mänskliga hjärnan hallucinerar verkligheten. Det betyder att den värld vi upplever till viss del är den värld vi förutsäger. Perceptionen i sig är långt ifrån ett enkelt fönster mot världen, utan genomsyras från första början av våra egna förutsägelser och förväntningar.

Den genomsyras inte bara i den meningen att våra egna idéer och förutfattade meningar påverkar hur vi senare bedömer saker och ting, utan också i en djupare, ursprungligare mening. Perceptionsprocessen, själva mekanismen som ser till att vi har kontakt med världen, drivs i sig av en rik källa av förutsägelser och förväntningar.

På de kommande sidorna ska jag försöka ge dig några förstahandsupplevelser av förutsägelsernas förmåga att förändra vad du ser och hör.

Förutsägelsernas kraft

Figur 1.1 utgör ett enkelt exempel. Läs den först uppifrån och ner och sedan från vänster till höger. Lägg märke till att formen på tecknet i mitten är densamma åt båda hållen. Men den visuella upplevelsen skiljer sig en smula beroende på vilket håll du läser ifrån. När du läser uppifrån och ner börjar din hjärna förvänta sig ett tal där (13), medan från vänster till höger väcker förväntningen om en bokstav (B). Dessa olika förutsägelser påverkar den visuella upplevelsen.



FIGUR 1.1. Ett rutnät med siffror/bokstäver.