

Scores rapportserie 2015:1

Akademiska omdömen under osäkerhet: en studie av kollektiva ankringseffekter i Vetenskapsrådets beredningsgrupper

Lambros Roumbanis

SCORE

STOCKHOLM CENTRE FOR
ORGANIZATIONAL RESEARCH

Akademiska omdömen under osäkerhet: en studie av kollektiva ankringseffekter i Vetenskapsrådets beredningsgrupper

Lambros Roumbanis

lambros.roumbanis@score.su.se

Scores rapportserier 2015:1
ISBN 978-91-89658-84-4

Stockholms centrum för forskning om offentlig sektor
106 91 Stockholm
www.score.su.se



Abstract

The main focus of this study concerns the existence of *anchoring effects* in the peer review process of research proposals. According to the psychological theory of anchoring effects, once an anchor (represented by a number) is set, other judgments are made by adjusting away from the anchor, and there is a common human tendency toward interpreting other information around the anchor. This effect is commonly seen as a general flaw in human judgment, a cognitive bias that stems from a specific heuristic that guides people when they involve their intuition in order to solve a problem. In the following study, this cognitive bias will be analysed from a sociological point of view, that is, as a *collective* phenomenon rather than an individual one. It will also be applied to a new empirical case, the peer review process of research proposals. The anchoring effects have here been investigated using a meeting ethnographic approach. The evidence comes from direct observations of ten panel groups at the Swedish Research Council (*Vetenskapsrådet*). The analysis suggest that collective anchoring effects emerge as a result of the combination of the evaluation techniques that are being used (grading scales and middle ranking) and the efforts of the evaluators to reach consensus. The author suggests that what many commentators and evaluators have earlier interpreted as a certain proportion of chance in the peer review process, might instead be better understood as anchoring effects.

Innehållsförteckning

Inledning	3
Ankringseffekter och kognitiv partikularism	4
<i>Vad är ankringseffekter?</i>	4
<i>Ankringseffekternas mångfald i vardagen: att söka nya angreppssätt</i>	6
<i>Kognitiv partikularism, kommunikation och epistemisk smak</i>	7
<i>Bedömningsprocessens organisering och om den reglerade metodens betydelse</i>	9
Metodologisk redogörelse	11
<i>Vad observerades?</i>	11
<i>Retrospektiv analys av ankringseffekter på gruppnivå</i>	12
Om bedömningsprocessen: en kort beskrivning	13
Förankringar och justeringar av de akademiska omdömena: iakttagelser och exempel från Vetenskapsrådets beredningsgrupper	14
<i>Vilka ankare väger tyngst?</i>	17
<i>Kollektiva ankringseffekter och tillfälligheternas makt</i>	20
Slutord	21
Litteraturförteckning	22

Inledning

“To reach a port we must sail, sometimes with the wind, and sometimes against it. But we must not drift or lie at anchor.”

Oliver Wendell Holmes, Sr.

Inom modern kognitionspsykologi är det idag en vedertagen uppfattning att det mänskliga omdömet ofta påverkas av det som kallas för *ankringseffekter*. Teorin om ankringseffekter avser en lutning eller snedvridning som uppstår i själva omdömesakten i förhållandet till ett givet siffervärde (”ankaret”). Det första siffervärdet som förankras i individen – till exempel utropspriset på ett hus som en mäklare angivit eller den fängelsedom som en åklagare yrkat på i antal månader i fängelse – tenderar påfallande ofta att inverka på bedömningen av det som skall värderas. Siffran bildar en kognitiv referenspunkt, kort sagt, ett mentalt ankare som i viss mån blir bestämmande för hur långt individen med hjälp av sin intuition kan komma att sträcka sina kommande uppskattningar. Detta betyder att när ankaret väl är satt, förefaller det nästintill naturligt för individen att justera sitt omdöme utifrån detta ankare. Denna företeelse har ansetts problematisk på grund av den godtycklighet som den medför och som smyger sig in i omdömet och i förlängningen påverkar avgörande beslut (Tversky & Kahneman 1974).

I följande undersökning skall jag ta hjälp av teorin om ankringseffekter i en sociologisk studie om hur beredningsgrupper inom Vetenskapsrådet bedömer forskningsansökningar. Eftersom det handlar om en organiserad bedömningskontext där individerna översätter sina kvalitativa omdömen till olika sifferbetyg, som de sedan jämför och justerar genom dialog, skall ankringseffekterna här i första hand studeras på gruppnivå. Studien tar fasta på ett stort antal förankringar och justeringar som på olika sätt går in i varandra i kraft av den sociala interaktionen och som leder till att ankringseffekterna uppstår på ett aggregerat plan. Ur en sociologisk synvinkel är det organisationens specifika bedömningsmetoder (betygskriterier, granskningsansvar, medelranking) samt problem som är kopplat till konsensuskapandet (osäkerhet, vaghet, oenighet) som sammantaget genererar de kollektiva ankringseffekterna som inverkar på bedömningsprocessens slutresultat. Bidraget pekar dessutom på det faktum att dessa oönskade och problematiska snedvridningar till viss del ligger bortom både bedömarens och Vetenskapsrådets direkta kontroll.

Ankringseffekter har påvisats i flera olika sammanhang. Men så vitt jag vet har de aldrig tidigare studerats i samband med den organiserade bedömningen av forskningsansökningar. Att teorin om ankringseffekter inte utgör en sociologiskt uttömmande förklaringsmodell för att förstå hela denna bedömningskontexts villkor, är enligt min mening självklart. Däremot bidrar teorin till en viktig del i förståelsen av dessa villkor, och äger också i sig ett vetenskapligt egenvärde. Här används teorin i första hand för att belysa vad som sker när Vetenskapsrådets bedömare ställs inför den många gånger utmanande uppgiften att värdera forskningsansökningar med hjälp av sifferbetyg som sedan skall justeras för att gruppen skall kunna nå en acceptabel konsensus. Att fokus läggs vid siffrornas makt i bedömningsprocessen innebär inte något förnekande av att själva grunden för alltsammans handlar om det akademiska omdömet förmåga att urskilja, värdera, jämföra och kommunicera vad vetenskaplig kvalitet är eller kan tänkas vara för någonting. Det akademiska omdömet utgör själva kärnan i den kollegiala bedömningsprocessen, men är på samma gång den springande punkten i situationer där det uppstår delade meningar och oenighet. Sammanförandet av den kognitionspsykologiska teorin om ankringseffekter och olika vetenskaps sociologiska teorier

om det akademiska omdömet villkor anser jag här leda fram till en kompletterande och fruktbar analys som även kan användas i kommande studier. Andra viktiga sociologiska dimensioner såsom exempelvis makt, hierarki, status, nätverk och moral, är också i allra högsta grad relevanta, men skall inte behandlas specifikt i den här undersökningen, utan berörs främst indirekt. Men jag anser på det hela taget att jag i och med blottläggandet av de kollektiva ankringseffekterna kan kasta nytt ljus över de snedvridningsmoment som uppstår i denna inomvetenskapliga bedömningsprocess – snedvridningsmoment som tidigare ofta har betraktas som resultat av slumpen.

Ankringseffekter och kognitiv partikularism

Vad är ankringseffekter?

I boken *Ankare* från 1988 berättar författaren Klas Östergren om en man som insett att han måste ändra kurs i livet för att inte kapsejsa och gå under. Men hans förflutna gör sig ständigt påmint, ”Den gamla tron ville inte ge vika, den var så ingrodd och förankrad och gjorde allt för att förvirra sin bärare”. Ankaret, vars praktiska syfte är att förhindra båtar från att driva iväg med vind och strömmar, har i olika kulturer fått symbolisera säkerhet och stabilitet. Östergren använder dock ankringsmetaforen för att peka på en inre osäkerhet i förhållandet till det förankrade. Detta visar på metaforens grundläggande styrka i att den både kan skapa en ny språklig inramning och samtidigt inskräpa fokus.¹ När ankringsmetaforen lyfts ur sin ursprungliga nautiska kontext kvarstår en viktig del av dess meningsbärande innehåll – tyngddimensionen – som bidrar till att den gamla förankrade tron som Östergren nämner får en helt annan innebörd. Denna litterära skildring pekar mot att vanans makt inte sällan kan vilseleda omdömet och leda till självbedrägeri. Även inom vetenskapen har ankaret fått tjäna som metaforisk hävarm i samband med analyser och studier av det mänskliga omdömet bryderier.

När de två psykologerna Amos Tversky och Daniel Kahneman presenterade sin upptäckt av ankringseffekter i artikeln ”Judgment under Uncertainty: Heuristics and Biases” (1974), blev detta något av ett startskott till det som skulle kunna betraktas som ett nytt paradigm inom modern kognitionspsykologi. Så här beskrev man ankringseffekten:

”In many situations, people make estimates by starting from an initial value that is adjusted to yield the final answer. The initial value, or starting point, may be suggested by the formulation of the problem, or it may be the result of a partial computation. In either case, adjustments are typically insufficient. That is, different starting points yields different estimates, which are biased toward the initial values. We call this phenomenon anchoring”. (Tversky & Kahneman 1974, s. 1128)

När ett ankringsvärde väl har satts tenderar det att bilda en tyngdpunkt som håller människors kommande omdömen inom ett visst räckhåll från detta ankringsvärde. Och detta gäller även när helt godtyckliga siffervärden har förankrats. Ankringseffekter kan alltså i någon mening tyckas angränsa till det rent slumpmässiga. Men enligt de två psykologerna bör de i själva verket förstås som en del i en sorts kognitiv snedvridning eller lutning som har sina rötter i en av flera olika heuristiska principer som människor naturligt använder sig av vid problemlösningar i vardagen. Dessa heuristiska principer kan helt enkelt betraktas som en sorts ”genvägar” som människor använder sig av när de gör bedömningar som mer eller

¹ Se särskilt Max Blacks *Models and Metaphors: Studies in language and philosophy* (1962), Paul Ricoeurs *The Rule of Metaphor* (1978), samt Richard Swedbergs *The Art of Social Theory* (2014).

mindre härrör ur intuitionen. All ny information, oavsett vilken typ, tolkas generellt sett i förhållande till det ursprungliga ankringsvärdet. Denna upptäckt har senare kommit att uppfattas som fundamental, "it is arguably one of the most important truths about human judgment" (Simmons, Nelson & LeBoeuf 2010). Tversky och Kahneman betraktade detta som ett grundexempel på hur rationalitet begränsas när människor tvingas göra improviserade uppskattningar av siffervärden i olika kontexter. Kahneman har senare understrukt att intuitionen visserligen är högst avgörande för människans förhållningssätt till verkligheten, och att den oftast fungerar alldeles utmärkt, men att detta i sig gör det extra intressant att studera den när den brister i sin förmåga att vägleda oss (Kahneman 2011). Upptäckten av ankringseffekterna och andra former av bias har inte bara fått stort betydelse inom psykologin och socialpsykologin, utan även inom den ekonomiska och politiska forskningen, till exempel om människors risktaganden och beslut som är knutet till vinst/förlust på olika marknader.²

Hur kan man då påvisa existensen av ankringseffekter? Tversky och Kahneman lät utföra flera olika sorters experiment för att illustrera hur effekterna manifesterar sig. I ett av sina experiment gav man ett antal försökspersoner en slumpmässigt vald siffra (mellan 0 och 100) som i procent fick representera hur många afrikanska länder som är medlemmar i Förenta Nationerna. Utifrån denna siffra skulle försökspersonerna sedan göra en egen uppskattning. I en grupp fick försökspersonerna ett medianvärde på 10 som ankringsvärde, vilket resulterade i en medianuppskattning runt 25. I en annan grupp fick försökspersonerna medianvärdet 45 som ankringsvärde, vilket istället resulterade i en medianuppskattning runt 65. Experimentet visade följaktligen hur ankringsvärdena hade en given effekt på försökspersonerna när de försökte göra sina bedömningar. Detta experiment avsåg visserligen lekmän och rörde dessutom smått triviala kunskapsfrågor. Men ankringseffekten påverkar även forskare och experter, vilket också påpekades i artikeln, "Experienced researchers are also prone to the same biases – when they think intuitively" (Tversky & Kahneman 1974, s. 1130).

Ett intressant exempel där just professionella bedömare utsätts för ankringseffekter går att finna i Birthe Englich och Thomas Mussweilers artikel "Sentencing under uncertainty: Anchoring effects in the courtroom" (2001). I denna studie, som grundar sig på ett experiment med erfarna domare inom rättsväsendet, fann man tydliga indikationer på ankringseffekternas inverkan inom ett område där specialkunskaper erfordras för den som skall fälla ett omdöme. Ankringseffekterna visade sig till och med få ganska dramatiska följder i utfallet. Englich och Mussweiler fann skillnader i domar på upp till 8 månader för exakt likvärdiga brott. Orsaken knöts till den siffra (månader i fängelse) som åklagaren först yrkade på – som lägst 12 månader och som högst 34 månader. Domarnas juridiska omdömen påverkades alltså helt tydligt av ankringseffekter när de fattade sina avgörande beslut, vilket i verkliga fall skulle ha resulterat i skeva och orättvisa konsekvenser för de dömda människornas liv.

Forskningen om ankringseffekter har påvisats i många olika kontexter, såsom till exempel vid uppskattning av priser (Mussweiler, Strack & Pfeiffer 2000; Nortcraft & Neale 1987), sannolikheten för kärnvapenkrig (Plous 1989), förhandlingar (Galinsky & Mussweiler 2001; Ritov 1996), och i samband med generella kunskapsfrågor (Mussweiler & Englich 2005;

² Herbert Simon (1957) utvecklade tidigt teorier om olika heuristiska principer som styr människors omdömen vid problemlösning. Han anförde bland annat de numera klassiska idéerna om "bounded rationality" och "satisficing". Tversky och Kahneman har utifrån sin egen så kallade *prospect theory* kommit att kritisera vissa av Simons antaganden i artikeln "The Framing of Decisions and the Psychology of Choice" med följande motivering: "Like other intellectual limitations, discussed by Simon under the heading of 'bounded rationality', the practice of acting on the most readily available frame can sometimes be justified by reference to the mental effort required to explore alternative frames and avoid potential inconsistencies. However, we propose that the details of the phenomena described in this article are better explained by prospect theory and by the analysis of framing than by ad hoc appeals to the notion of costs of thinking" (Tversky & Kahneman 1981, s. 458).

Jacowitz & Kahneman 1995). Många psykologer och socialpsykologer har, vilket jag redan har påpekat, kommit att betrakta ankringseffekterna som en av de viktigaste sanningarna om det mänskliga omdömet villkor (McElroy & Dowd 2007; Simmons, LeBoeuf & Nelson 2010). Likväl råder det pågående diskussioner om hur man skall förstå ankringseffekternas inflytande på samhällslivet. Ett mer mångfacetterat synsätt har sakta men säkert kommit att utvecklas under de senaste åren. Denna utveckling har av vissa forskare kommit att uppfattas som en ”ny våg”, såväl teoretiskt som metodologiskt (Epley & Gilovich 2010).

Ankringseffekternas mångfald i vardagen: att söka nya angreppssätt

Ankringseffekternas allmänna innebörd, generaliserbarhet och eventuella begränsningar har diskuterats allt sedan Tversky och Kahneman presenterade sina första resultat (Chapman & Johnson 1994; Mussweiler 2001; Oppenheimer, LeBoeuf & Brewer 2007). En del forskare har uppfattat dessa effekter som lite gåtfulla, samtidigt som de har visat sig vara tämligen enkla att bevisa med hjälp av experiment. Effekterna har i flera studier framstått som nästintill oundvikliga. I till exempel en studie, där individerna blivit tydligt förvarnade innan experimentet om ankringseffekterna, lyckades de likväl inte värja sig från dem. Snedvridningen i justeringen av omdömet skedde trots förkunskap och medvetenhet om effekten ifråga, vilket ger en signal om ankringseffekternas generella styrka (Wilson, Houston, Etling, Brekke 1996). Men vad är det egentligen som gör att ankringseffekterna uppstår och inverkar på det mänskliga omdömet? Vad finns det för underliggande mekanismer och faktorer; vilka förutsättningar krävs för att effekterna överhuvudtaget skall uppstå? Finns det variationer och gradskillnader, och varifrån härstammar dessa i så fall? Och kan vissa personer vara mer motståndskraftiga mot ankringseffekter än andra; kan tillfälliga sinnesstämningar eller bestämda personlighetsdrag spela någon roll (Englich & Soder 2009; McElroy & Dowd 2007)? Kan ett ”öppet sinne” motverka eller påskynda eventuella ankringseffekter? Och hur mycket påverkar en individs övertro (*overconfidence*) till det egna omdömet i huruvida effekten uppstår eller inte? Och vad skiljer ett ankringsvärde som en person har fått presenterat för sig och ett ankringsvärde som är självskapat (Simmons, LeBoeuf & Nelson 2010)? Detta är några exempel på sådant som psykologer har diskuterat under de senaste åren och som kan vara relevant att ha i åtanke när man undersöker ankringseffekter.

I artikeln ”Anchoring unbound” (2010) förklarar Nicholas Epley och Thomas Gilovich att ankringsparadigmet borde öppnas upp ordentligt för att man på allvar skall kunna börja beskriva och förklara alla de olika scenarier i det vardagliga livet där ankringseffekter faktiskt uppstår. De förtydligar sin syn på saken så här:

”The main benefit of this broader perspective is that it brings into view sources of variability in anchoring effects that have so far been largely ignored: variability in the kinds of anchors that influences judgment, variability in processes that give rise of anchoring effects, variability in the contexts in which anchoring effects are likely to be elicited, and variability in consequences of anchoring that go beyond an immediate influence on numerical judgment”. (Epley & Gilovich 2010, s. 20)

Här kan sociologin bidra till en breddning av denna kognitionspsykologiska upptäck genom att föra in ett dynamiskt interaktionsmoment som tar fast på ankringseffekter i grupp. Därmed skapas snarast ett fokus på kollektiva ankringseffekter som uppstår på en aggregerad nivå. Ankringseffekterna förstås därigenom som resultatet av samspel, där olika attityder, övertalningsförsök, makt och statusordningar kan få en avgörande betydelse för dess verkningsgrad. I den här studien är det dessutom av största vikt att ta hänsyn till den

organisatoriska kontexten, där särskilda regler, kontroll, metoder och teknologier används för bedömning och beslutsfattande. Människor som bedömer saker tillsammans i grupp ställs inför en *serie av anpassningar* av sina respektive omdömen, vilket enligt min mening medför att ankringseffekterna uppstår på ett mera komplicerat plan. Det förutsätter betydligt flera figurationer av bedömningar som kan skapa svåröverskådliga ankringseffekter på ett kollektivt plan. I Vetenskapsrådets beredningsgrupper, där det eftersträvas konsensus beträffande värdet på en stor mängd forskningsansökningar, blir förhandlingar och övertalningsförsök en naturlig del i det som skapar grogrunden för att ankringseffekter sprider sig till en aggregerad nivå. På sätt och vis går det att betrakta beredningsgruppernas arbete som ett organiserat försök att komma fram till rimliga överenskommelser. Det handlar om att närma sig varandras perspektiv: "People 'jump' some amount from their original egocentric anchor and evaluate whether this new perspective plausibly captures the other's perception" (Epley, Keysar, van Boven & Gilovich 2004, s. 328). Det speciella i fallet med Vetenskapsrådet – som beskrivs mer i detalj nedan – är den betydelse som den framräknade medelrankingen får som en aggregerad samling av ankringsvärden. Denna medelranking står ständigt i centrum under beredningsgruppernas möten och inverkar med nödvändighet på den konsensus som bedömarna gemensamt lyckas skapa. Ytterst sett är det mot bakgrund av medelrankingen som de kollektiva ankringseffekterna får sin kraft på den aggregerade nivån.

Kognitiv partikularism, kommunikation och epistemisk smak

För att förstå ankringseffekternas akademiska kontext är det viktigt att göra några väsentliga tillägg som berör själva hjärtat av bedömningsprocessen. Hur når man konsensus i en värld som i hög grad är specialiserad, differentierad och heterogen? Eller istället, om man vrider på frågeställningen och problematiserar konsensusskapandet som sådant: Hur kan nyskapande och annorlunda forskningsprojekt få gehör i en grupp av bedömare med allt för likartade synsätt och standarder för vad som räknas som god vetenskap? Tidigare forskning har vid upprepade tillfällen gjort gällande att det akademiska omdömet ofrånkomligen förmedlar olika former av kognitiv och institutionell bias (Bourdieu 1975; Chubin & Hackett 1990; Hemlin 1996; Huutoniemi 2012; Langfeldt 2006; Travis & Collins 1991). Samtidigt innebär den kollegiala bedömningsprocessen i praktiken att man som bedömare förväntas kunna stiga utanför sina egna preferensramar för att tillsammans kunna urskilja de starkaste och mest lovande ansökningarna genom kommunikation och skapandet av kompromisser (Huutoniemi 2012; Lamont 2009). Likväl är processen avhängig vilka individer som väljs ut för att bedöma ansökningarna, vilket ibland kan beror på rena tillfälligheter. När sociologen Stephen Cole och hans kollegor genomförde en experimentstudie där man testade skillnader i bedömningar av 150 olika ansökningar, fann man en anmärkningsvärd variation. Ansökningarna bedömdes både av National Science Foundations beredningsgrupp och av en kontrollgrupp. I artikeln "Chance and Consensus in Peer Review" som 1981 publicerades i tidskriften *Science*, summerade man sin upptäckt på följande sätt:

"The fate of a particular grant application is roughly half determined by the characteristics of the proposal and the principal investigator, and about half by apparently random elements which might be characterized as the 'luck of the reviewer draw'". (Cole et al. 1981, s. 885)

En viss slump tycktes alltså vara bekräftad. Cole och hans kollegor tillade dock att detta inte fick tolkas som att hela processen är slumpmässig. Istället bör man lägga fokus på spänningen i själva skapandet av konsensus i beredningsgrupperna, och särskilt den betydelse som oenighet mellan bedömarna kan få för slutresultaten. Slumpen ligger och lurar i marginalerna,

trots forskningsfinansiärernas tappra försök med att få bort den ur bedömningsprocessen.³ Drygt tio år senare hänvisar Cole till denna studie i sin bok *Making Science*, och påpekar då att det inte verkar finnas mer grund för konsensus när det gäller bedömning av nya vetenskapliga projekt än vid bedömningen av exempelvis konstverk eller Bordeauxviner (Cole 1992, s. 19). Frågan om ”smak” tycks alltså även göra sig gällande inom vetenskapens domäner – men i vilket avseende?

Många forskare som har intervjuats i olika studier, har ofta påpekat att det akademiska omdömet handlar om välgrundad kunskap, men också om en särskilt utvecklad ”känsla” (en intuitiv förmåga) som gör att man lär sig känna igen värdefulla frågeställningar, problem och angreppssätt. Man får enkelt uttryckt en viss vana att bedöma forskare och deras forskning. I boken *The Sociology of Science* är Robert Merton inne på detta spår då han refererar till ett flertal kända amerikanska Nobelpristagare som berättat om sina personliga erfarenheter, ”They uniformly express the strong conviction that what matters most in their work is a developing sense of taste, of judgment” (Merton 1973, s. 453). Michele Lamont har också beskrivit det akademiska omdömet genom att bland annat hänvisa till denna estetiska dimension. En av hennes intervjupersoner ger följande bild av den kollegiala bedömningsprocessen och om vad som sker när olika epistemiska standarder möts:

”There are competing narratives about what passes for being good. There are different standards of excellence, different kinds of excellence, and I’m certainly willing to entertain somebody else’s standard of excellence up to a point. I’m not sure that I could articulate what that point is, but I’m pretty confident that I’d know it when I see it. You develop a little bit of a nose for it. Particularly for what’s bad” (Lamont 2009, s. 1)

Både enighet och oenighet mellan bedömarna hänger ihop med det som har kommit att kallas för *kognitiv partikularism*, det vill säga, förekomsten av kognitiva gränser inom och mellan olika vetenskapliga specialiteter och discipliner (Travis & Collins 1991, s. 327). Att bedöma forskning är alltså delvis en fråga om ”epistemisk smak”, som på olika sätt är rotat i epistemiska kulturer (Knorr Cetina 1999). Detta sätter fingret på en hel del av det som sociologiskt betraktat möjliggör vetenskapen så som mellanmännisklig aktivitet överhuvudtaget. Det subjektiva elementet och dess estetiska resonansbotten har en inte ovidkommande betydelse i detta sammanhang. Går man tillbaka till filosofen Immanuel Kants verk *Kritik av omdömeskraften* från 1791, får man en förbluffande genial analys av hur det döljer sig en estetisk dimension mitt inne i det ”rena förnuftets” hjärta. När det kommer till frågan om att värdera något, kopplas omdömet till känslan. Michele de Certeau har apropå Kants teori påpekat att, ”acknowledging an *art* at the root of thought, makes judgment a ’middle term’...between theory and praxis” (de Certeau 1984, s. 72). Omdömet utgör följaktligen en avgörande länk mellan teori och praktik, och är i en väsentlig mening knuten till den subjektiva känslan och fantasin.

Kants perspektiv ger oss ett filosofiskt rotsystem som grund för de sociologiska idéerna om kognitiv partikularism och epistemiska kulturer, och han gör det för övrigt genom att framhålla kommunikerbarhetens, den gemensamma sinnlighetens (”*sensus communis*”) och objektivitetens betydelse.

”Med *sensus communis* måste man förstå idén om ett gemensamt sinne, det vill säga en förmåga att falla omdömen, som i sin reflektion tar hänsyn till vars och ens föreställningsart i tankarna (a priori), som om det jämförde sitt omdöme med det samlade mänskliga förnuftet, för att därigenom undgå

³ Robert Porter har hävdad att, ”luck has a great deal to do with the outcome. Despite the sponsor agencies’ best efforts, there is an element of randomness in the final decision” (2005, s. 347).

den illusion som stammar ur subjektiva privatbetingelser, vilka lätt kan hållas för objektiva och har ett skadligt inflytande på omdömet. Detta sker genom att man jämför sitt omdöme med andras omdömen, inte bara verkliga utan också möjliga sådana” (Kant 2007 [1791], s. 152).

Hans perspektiv (om än i vissa avseenden motsägelsefullt), bidrar enligt min uppfattning, till en högst intrikat förståelsegrund för både skillnaderna och likheterna inom den akademiska världen. Att bedöma olika slags kvaliteter, oavsett om det rör sig om unika konstföremål, gamla Bordeauxviner eller nya forskningsansökningar, förutsätter en särskild känslighet och en fantasiförmåga, kort sagt: en utvecklad estetisk förmåga inom ett visst område.

När Vetenskapsrådet organiserar sina beredningsgrupper är det en självklarhet att alla omdömen räknas, samtidigt som man aktivt vill motverka olika korrumpierande krafter som kan ge oönskade effekter på bedömningsprocessens slutresultat. Individuella värdeomdömen skall leda till kollektiva beslut. Kommunikationen är avgörande för att skapa tillit och legitimitet. Ett sätt att se på den kollegiala bedömningsprocessen och värdet av kommunikation har Michele Lamont presenterat:

”Debating plays a crucial role in creating trust: fair decisions emerge from dialogue among various types of experts, a dialogue that leaves room for discretion, uncertainty, and the weighing of a range of factors and competing forms of excellence” (Lamont 2009, s.7).

Att organisera möten där de sakkunniga bedömarna kan träda i dialog med varandra är avgörande för att olika typer av akademiska omdömen skall kunna komma till uttryck ordentligt och för att hela bedömningsprocessen skall vinna legitimitet. Men frågan är på vilket sätt dialogen verkligen kan ändra på den redan givna medelrankingen i gruppen och de latent a oenigheterna? Vad leder kompromisserna till för beslut? Finns det något verkligt utrymme för att omvärdera olika ansökningar från grunden om så behövs – en omvärdering av alla värden – eller blir medelvärdet i kombination med den ganska stora tidspresen allt för dominerande i arbetet? I en relativt nyligen publicerad artikel underströk författarna resultatet av sin studie med att, ”panel discussions per se did not improve the reliability of the evaluation” (Fogelholm et al. 2012). Teorin om ankringseffekter kan bidra till en utvidgad förståelse för dessa problem som bedömarna ofta kan uppleva när de utför sitt bedömningsarbete i grupp.

Bedömningsprocessens organisering och om den reglerade metodens betydelse

Förutsättningarna för bedömarna att nå konsensus bygger inte enbart på deras egna professionella kunskaper och i upprätthållandet av en god kollegialitet, utan också till stor del på hur forskningsfinansiären (i det här fallet Vetenskapsrådet) har organiserat alltsammans. Bedömningsprocessen omgärdas av en mängd olika regler och förordningar som måste följas. En sådan regel berör exempelvis jävssituationer, som man tar på mycket stort allvar och gör allt för att motverka från organisationens sida. Den som är jävig måste lämna rummet och skall absolut inte bedöma en ansökan om det råder någon som helst partiskhet. Två villkor som särskilt påverkar bedömningsarbetet har dels att göra med hur *bedömningsmetoderna* är utformade och dels på *budgetens* omfattning. I den här studien syntes det vara en allmän uppfattning bland bedömarna i samtliga tio beredningsgrupper att budgeten var alldeles för snål. Många uttryckte sitt direkta missnöje över hur lite pengar man hade att fördela i relation till det samlade arbetet som utfördes. När det kommer till bedömningsmetodernas utformning (betygsskalor, rankingförfaranden, granskningsansvar) kan de faktiskt ha en ganska påtaglig och intressant inverkan på processen såtillvida att de potentiellt sett kan forma utsikterna för projekt som till exempel uppfattas som säkra respektive riskfyllda (om än i vissa fall

vetenskapligt nyskapande). Vissa bedömningsmetoder kan resultera i att ”säkrare” ansökningar har större chanser att få anslag, och detta är även något som Vetenskapsrådet har identifierat som ett mer generellt problem: ”Det finns en risk att säker forskning favoriseras på bekostnad av nyskapande och att tvärvetenskaplig forskning kan ha svårt att få gehör.”⁴ I Norge har Liv Langfeldt (2006) analyserat betydelsen av beredningsgruppens sammansättning och olika slags bedömningsmetoders inverkan på bedömarnas handlingsmöjligheter. Forskningsfinansiärer skiljer sig ofta åt i sina arbetsmetoder. Om man till exempel använder sig av grovranking och ett öppet beslutsfattande tenderar det att kunna ge ett större utrymme för vetenskaplig mångfald och nyskapande projekt (entusiastiska bedömare kunde iaktta förändra sina kollegors åsikter om ansökningar som de först hade ansett vara för riskabla, perifera eller omogna). Langfeldt påpekar vidare, att ett större antal bedömare per grupp som sätter betyg på alla ansökningar i kombination med en fingradig betygsskala (1,0-4,0) och en medelranking, tenderar att bidra till ett mera grundligt och förutsägbart slutresultat. Och hon konstaterar även att, ”few reviewers per application and no scholars to compare and rank the whole portfolio of applications give ample room for randomness” (Langfeldt 2006, s. 832).

Spänningen ligger någonstans i den omfattande växelverkan mellan de individuella bedömningarna och den aggregerade medelranking som sammanställts (och som hela tiden ligger till grund för beredningsgrupperna i deras arbete). Utifrån en skisserad teori om kollektiva ankringseffekter skall jag göra antagandet att medelrankingen bör och kan betraktas som en *ansamling av ankringsvärden* som på olika sätt påverkar hur långt gruppen kan gå i sina gemensamma överläggningar. Medelrankingen har av förklarliga skäl en viktig funktion att fylla för gruppen – tanken är att den skall ge en ungefärlig bild av hur gruppen som kollektiv har bedömt ansökningarna, vilket bedömare sedan skall kunna förhålla sig till under mötet. Men det ger alltså även upphov till ett nytt läge av förankringar som på olika sätt influerar kommande gemensamma justeringar.

Vetenskapsrådets bedömningsmetoder kan på sätt och vis ses som en blandning av de typer som Langfeldt har tagit upp. Man har exempelvis en öppen diskussion mot bakgrund av en medelranking där varje ansökan har granskats av tre bedömare. Man använder sig av en mer grovskalig betygsskala (1-7), men i flera av de observerade grupperna använde man sig provisoriskt av decimaler för att kunna sortera och rangordna ansökningarna. De organisatoriska villkoren är på många sätt avgörande för den kollegiala bedömningen av forskningsansökningar. Reglerna för hur allting skall gå till i praktiken kan knappast underskattas.

⁴ www.vr.se/forskningsfinansiering/beredning/beredningsprocessen/forskaregranskarforskare/faeerreview/vilkanackdelarharpeerreview.5.5fa10c312ed4d5b90680001692.html

Metodologisk redogörelse⁵

Den empiriska undersökning som ligger till grund för den här studien genomfördes under sensommaren 2013 då tio av Vetenskapsrådets sammanlagt nitton beredningsgrupper inom ramen för natur- och teknikvetenskaper observerades under deras tvådagarsmöten. Valet av beredningsgrupperna gjordes dels med syfte att få en ämnesmässig spridning, och dels utifrån den varierande graden av internationella och inhemska bedömare. I några av grupperna var alla bedömarna från Norden, vilket gjorde att man då valde att tala på ”skandinaviska” istället för på engelska.⁶ Beredningsgruppernas möten ägde rum på ett större konferenshotell i Stockholm. Jag och två kollegor genomförde observationerna under den aktuella tiden. Resultatet blev sammanlagt tjugo dagars observationer. En direkt fördel med att jag hade ett litet observationsteam till min hjälp, förutom att materialet blev betydligt mer omfattande än om jag hade gjort observationerna själv, bestod i att jag på ett fruktbart sätt kunde jämföra och diskutera mina egna iakttagelser med mina kollegor efter mötenas slut. Detta stärkte på flera sätt bilden av vad som faktiskt är relevant och viktigt att förstå när det kommer till frågan om hur expertis och kunskap organiseras i denna specifika bedömningskontext.

Vad observerades?

Utgångspunkten för observationerna bestod i att med så öppna ögon som möjligt beskriva vad som skedde under beredningsgruppernas möten. Vad som genomfördes var helt enkelt en slags mötesetnografi.⁷ Huvudfokus lades vid hur interaktionen manifesterade sig mellan de involverade personerna och hur bedömningsarbetet gick till i praktiken. Den allmänna strategin som jag och mina kollegor enades om innan studien påbörjades var att försöka skriva ner så mycket som möjligt av det vi kunde se och uppfatta när vi var på plats. I linje med denna ambition försökte vi så obemärkt som möjligt utföra vårt arbete utan att på något nämnvärt sätt störa processen. Detta tycktes dock i regel ganska enkelt, för när mötena väl var igång blev tempot så högt att bedömarna knappt alls verkade bry sig om vår närvaro. Däremot var det flera bedömare som nyfiket ställde frågor under pauserna. Flera bedömare uttryckte dessutom ett visst nöje över att för en gångs skull själva vara studieobjekt. Det som skedde under mötena, då bedömningsarbetet var i full gång, var primärt sett det mest relevanta för studiens inriktning. Men även under luncher och kaffepauser förmedlades en hel del intressant information då bedömarna samtalade med varandra. Dessa samtal återanknöt många gånger direkt till sådant som rörde mötet och bedömningsarbetet, men också till andra helt vitt skilda frågor. Under dessa pauser fick även vi som observatörer chansen att delta i flera spännande samtal med bedömare och administratörer från Vetenskapsrådet. Dessa samtal utgör i sig en rik källa till kunskap om hur personerna ifråga tyckte och tänkte kring sitt arbete.

⁵ Jag vill här passa på att tacka den dåvarande huvudsekreteraren för Vetenskapsrådets natur- och teknikvetenskaper, Sven Stafström, som så vänligen gav mig möjligheten att genomföra denna observationsstudie. Även koordinatör Emma Olsson och övriga involverade skall ha stort tack. Jag vill i övrigt tacka Adrienne Sörbom och Staffan Furusten som vänligen hjälpte till att berika studien med sina observationer av vardera tre beredningsgrupper. Till kollegorna som deltog vid mitt seminarium på Score är jag också tack skyldig. Detta gäller särskilt Martin Gustafsson och Anette Nyqvist för deras goda synpunkter på det slutgiltiga manuskriptet. Sist men inte minst vill jag tacka Moa Bursell för många värdefulla kommentarer under arbetets gång.

⁶ Vetenskapsrådets regler gör gällande att det måste finnas minst 20 % utländska bedömare representerade i varje beredningsgrupp.

⁷ Om det metodologiska värdet av att studera formella möten, se Thedvall (2013), vari det bland annat går att läsa följande: ”Meetings reveal themselves to be condensed field sites for the examination of the workings of organisational logics” (Thedvall 2013, s. 117).

Retrospektiv analys av ankringseffekter på gruppnivå

De hittills vanligaste metoderna för att upptäcka och bevisa ankringseffekter har gått ut på att konstruera olika psykologiska experiment. Man har då systematiskt undersökt just befintligheten och graden av effekterna i olika sammanhang. Oftast har det gällt större sifferspann som till exempel 0-100, där effekterna tydligare kan testas. Fokus i dessa studier har från början varit inställt på ankringseffekterna som psykologisk mekanism. Men min utmaning är följande: Hur beskriver och belyser man ankringseffekter utifrån ett stort empiriskt material som från början inte alls varit direkt inställt på att undersöka just själva förekomsten av ankringseffekter i sig? Och hur förhåller man sig till de små sifferspannen som betygsskalorna för kvalitetskriterierna ger för handen; går det att se sådana effekter i dessa sammanhang? Låter det sig överhuvudtaget göras på ett acceptabelt sätt utan att det enbart leder till rena spekulationer? Detta är ingalunda obefogade frågor. Min uppfattning är emellertid att det visst går att undersöka ankringseffekter även med icke-experimentella metoder som grundar sig på samma vetenskapliga intuition, fast istället belyst genom en tolkande analys.

I den här studien blev ankringseffekterna uppenbara i retrospektiv, det vill säga, efter att själva observationerna hade genomförts – efter noggranna läsningar av de omfattande observationsanteckningarna och efter en diger genomgång av den litteratur som behandlar det mänskliga omdömet. Till skillnad från många kognitionspsykologer kom jag att lokalisera ankringseffekterna till en organiserad gruppnivå, och inte enbart på individnivå. Effekterna kom att identifieras som ett resultat av de samlade jämförelserna, justeringarna och rankingförfarandena, där siffrorna hela tiden var en ytterst viktig funktion för grupperna i deras arbete med att urskilja och avgöra kvalitetskillnader i de många ansökningarna. Även om de initiala ankringsvärdena i ett bestämt avseende skapades när individerna i enskildhet satte sina preliminära sifferbetyg, är det likväl organiserat kring gruppens medelranking, vilket i sig skapar en ny kollektiv ”ankringssituation”, som sätter gränser för och begränsas utav konsensuskapandet som reglerad bedömnings teknik.

Att förstå ankringseffekter utifrån den organiserade gruppinteraktionen innebär dock inte en negligering av individerna som gör själva arbetet, tvärtom, det viktiga är att se till både det ömsesidiga beroendet och de individuella variationerna i bedömningsprocesserna. Kognitiv partikularism och epistemisk smak (som inkluderar subjektiv känsla och fantasi) är, så som jag beskrev i det föregående stycket, högst avgörande för att förstå det akademiska omdömet. Men ankringseffekterna skall här likväl tolkas på den aggregerade nivån. Men för att kunna åstadkomma detta, utan några som helst experiment, krävs det ett rekonstruktionsarbete. Ankringseffekterna belyses här istället genom en sorts fenomenologisk-retrospektiv analys. Genom att beskriva sådant som sades och gjordes under beredningsgruppernas möten går det att på ett annat sätt frambringa evidens för att kollektiva ankringseffekter faktiskt existerar. Vetenskapliggörandet ligger således inte i att försöka mäta och bevisa, utan snarare i att belysa och tolka. Denna metod medför förvisso en del begränsningar i precisionsförmågan, för det är svårt att exakt klarlägga vart någonstans i bedömningsprocessen som ankringseffekterna får sin eventuella verkningsgrad.

Med stöd av Epley och Gilovichs (2010) uppmaning om en generell breddning av synen på ankringseffekter skall jag utgå ifrån att sifferskalans storlek i sig inte betyder så mycket i det här fallet som själva översättningen mellan mening och siffror, som potentiellt sett kan rymma en hel del osäkerhet, vaghet och ambivalens. Detta ser jag som en viktig fingervisning för framtida experiment om ankringseffekters inverkan och konsekvenser vid kollegial bedömning av forskningsansökningar. Min målsättning består i att presentera allmänna

iakttagelser baserat på vad som sades och gjordes under beredningsgruppernas möten, för att på så sätt finna stöd för att det faktiskt förekommer ankringseffekter i denna kunskapsintensiva bedömningsprocess.

Om bedömningsprocessen: en kort beskrivning⁸

Hela bedömningsprocessen går förenklat uttryckt till på följande vis: bedömarna läser först var och en för sig ansökningarna under sommaren. Det är under denna del av processen som bedömarna betygsätter och rangordnar ansökningarna som de fått sig tilldelade.⁹ Därefter sammanställs alla bedömnarnas betyg och en preliminär medelranking upprättas i dataprogrammet VR-Review. Denna medelranking (eller grovranking) är mycket viktig som utgångspunkt för det kommande bedömningsarbetet – den ger en bild av hur beredningsgruppen, så som kollektiv betraktat, har bedömt ansökningarna. Det är utifrån denna medelranking som mycket av diskussionerna sedan kretsar. Det man tillslut lyckas enas om i grupperna under dessa tvådagarsmöten utgör sedan underlag för de kommande bidragsbesluten som Vetenskapsrådets styrelse fattar senare under hösten.

När man sitter med under beredningsgruppernas möten slås man ganska snabbt av den ofta hövliga och vänliga tonen mellan bedömarna. Michele Lamont har beskrivit kollegialitetens konkreta betydelse när bedömarna skall fastställa konsensus, ”it is the oil that keeps the wheels of deliberation turning when panelists otherwise might not be willing to accommodate one another” (Lamont 2009, s.120). Den ömsesidiga respekten som bedömarna visar varandra, gör att konflikter i princip aldrig tar fart på riktigt, utan istället stannar vid enklare små dispyter eller tystnad. Tystnaden kan för övrigt ses som en viktig del i skapandet av konsensus – den lägger locket på oenighet, vilket kunde iaktas vid upprepade tillfällen under observationerna av de olika grupperna (Se Haug 2013).

Förhandlingarna sker ofta i en ganska rask takt på grund av den stora mängden ansökningar som gruppen skall gå igenom. Vissa av dessa ansökningar som befinner sig på gränsen för finansiering kräver oftast mer samtalstid. Mest tid lägger man ner på de ansökningar som man anser vara mest intressanta och som har en rimlig chans att faktiskt kunna få anslag. Då kan det i vissa fall dra ut på tiden ordentligt, särskilt när bedömarna inte är överens. Ibland kan diskussionerna skena iväg ordentligt, och då är det som regel beredningsgruppens ordförande som måste gå in och styra upp. Varje ansökan granskas i regel av tre bedömare (ibland händer det även att en extern bedömare har involverats). Man går effektivt igenom sina omdömen och redovisar sina sifferbetyg. Alla de ansökandes namn syns på en stor projektorsskärm med alla de samlade betygen som bedömarna har satt. Ordförande, som har en viktig ledarroll att fylla, påminner ofta om tidsbristen och att man måste skynda på för att hinna med allt som skall göras. Överläggningarna börjar först efter att de jäviga har lämnat rummet – detta bidrar

⁸ En enkel och bra beskrivning om hur bedömningsprocessen går till finns i Vetenskapsrådets *Beredningshandbok* (www.vr.se)

⁹ Bedömningen görs utifrån fyra grundkriterier: (a) *nytänkande och originalitet*, (b) *vetenskaplig kvalitet*, (c) *sökandes kompetens* (meriter), (d) *genomförbarhet*. Därutöver har man även en *sammanfattande bedömning av ansökans vetenskapliga kvalitet*. Betygsskalan är sjugradig för alla kriterierna utom den som avser genomförbarheten som istället är tregradig. Den sjugradiga skalan går från 7=”Enastående”, 6=”Utmärkt”, 5=”Mycket bra till utmärkt”, 4=”Mycket bra”, 3=”Bra”, 2=”Svag”, och 1=”Dålig”. (*Beredningshandboken* s.26-32)

till ett konstant springande in och ut.¹⁰ Jävsreglerna är strikt hållna, men bygger ändå i hög grad på att bedömarna själva är ärliga med sina relationer och intressen.

Ett typiskt förfarande kan se ut som följer: förslagsställaren börjar med att påpeka att hon tycker att ansökan är välskriven och spännande. Hon tillägger att personen som söker anslag vill undersöka särskilda proteiner genom en ny metod. Och hon tillägger att det rör sig om ett mycket innovativt projekt som kan leda fram till viktig och relevant kunskap om en viss specifik process. Den ansökande har publicerat 18 artiklar sedan 2009, flertalet i topptidskrifter. Förslagsställaren påpekar därpå att hon har rankat den aktuella ansökan som nummer 2 av de 17 ansökningarna som hon granskat. Och hon understryker därpå igen att det rör sig om ett mycket bra projekt som imponerar på henne. Hon redovisar betyget 6 för originalitet och nyskapande, 6 för projektets vetenskapliga kvalitet, 6 för kompetens, och 3 för genomförbarhet. Sammanfattande betyg sätter hon 6. Efter att förslagsställaren har gjort sitt anförande följer de andra två bedömarna med sina synpunkter och betyg. Den förste bedömaren håller med förslagsställaren. Han säger att han hade lite lägre på originalitet och nyskapande, men att han kan gå med på 6 ändå. Han tycker också att det är ett mycket spännande projekt. Den andra bedömaren, anser dock att betygen är för höga och skulle helst vilja sänka dem en aning. Hon var inte lika imponerad av ansökan och har därför rankat den betydligt lägre. Betygen för det första och det tredje betygskriteriet hade hon satt 4 på, istället för 6 som de andra två bedömarna hade satt. Sedan frågar beredningsgruppens ordförande vilka betyg man preliminärt skall skriva in i systemet. De kommer överens om 5, 6, 5 och 3. Som sammanfattande betyg sätter man en 5:a. I stora drag är det ungefär så här arbetet går till i grupperna när man skall nå konsensus. Men det är framför allt när man är oense om betygsättningen eller när det uppstår osäkerhet i hur en ansökan skall bedömas i relation till andra ansökningars värde som ankringseffekterna blir som tydligast.

Förankringar och justeringar av de akademiska omdömena: iakttagelser och exempel från Vetenskapsrådets beredningsgrupper

Om osäkerhet, vaghet och siffermagi

De samlade observationerna som gjordes av Vetenskapsrådets beredningsgrupper visade generellt sett på konstellationer av sakkunniga bedömare som tillsammans arbetade systematiskt för att skapa konsensus. Ingen av grupperna misslyckades i någon formell mening med uppgiften att få fram ett beslutsunderlag under den utsatta tiden. Likväl var bedömningsarbetet långt ifrån fritt från problem. Flera bedömare uttryckte exempelvis känslor av osäkerhet, vanmakt och ambivalens i samband med justeringarna av siffrorna och i jämförelserna av de olika ansökningarnas kvalitet. Möjligen hade de redan från början varit allt för försiktiga i sina betygsättningar, eller så blev de allt för villiga att släppa efter när diskussionerna väl tog fart ordentligt. Andra bedömare ansåg sig dock vara ganska säkra på sina omdömen, och där någonstans mellan säkerhet och osäkerhet, skulle anpassningarna av betygen och konsensus skapas. Ankringseffekter kan gå i olika riktningar beroende på den sociala interaktionen, där dragkamperna mellan de olika bedömarna ofta mynnar ut i en form av pragmatisk acceptans inför situationen. Att olika bedömare kan hysa olika grader av tilltro

¹⁰ I linje med vad March och Olsen har skrivit i *Ambiguity and Choice in Organizations*, kan min studie av Vetenskapsrådet indirekt även ses som, ”an attempt to understand how organizations deal with ambiguity – goals that are unclear, technologies that are imperfectly understood,...and participants who wander in and out” (March & Olsen 1976, s. 8).

till sina egna och/eller andras omdömen kan exempelvis bero på personlighet och känslor, något som vissa psykologiska studier har visat (Englich & Soder 2009; McElroy & Dowd 2007). Själva idén att organisera möten handlar om att bedömarna skall kunna ändra på det mekaniskt framräknade medelvärdet genom samtal och gemensamma reflexioner. På så vis kan de uttrycka sina uppfattningar och försöka övertyga varandra om de olika ansökningarnas vetenskapliga värde. Men möjligheten att kunna påverka och förändra, verkar inte alltid helt självklar, något som en bedömare påtalade inför sina kollegor i en kritisk situation: "I am worried that we are sticking with our initial ranking, and not changing things". Detta bekräftar den tröghet som finns i bedömningsprocessen och som är förknippad med rankingförfarandet och de reglerade bedömningsmetoderna som Vetenskapsrådet använder sig av och som inverkar på konsensuskapandets villkor (Jämför Langfeld 2006).

I ordväxlingen här nedan får man en inblick i hur det lät när några bedömare diskuterade saker som upplevdes som högst problematiska i deras arbete. Alltsammans utspelade sig under den andra mötesdagen då bedömarna fortsatte att finjustera rangordningen, för att på så vis nå en slutgiltig konsensus:

Bedömare 1: I couldn't sleep. It really says something about us when a person like that gets that low on scientific quality,...And I know why: because no one that has judged it really understands it. But not understanding the proposal is not the same as low scientific quality. I say this because I feel bad.

Bedömare 2: But everybody understands the rules of the game!

Ordförande: This is a dilemma. No one is an expert of everything, and the grades *fluctuate*. At this point we can't change anything. But I agree with you, there is always a bad taste after these meetings. But lets save the discussion to afterwards, because we have more to do.

Bedömare 3: We are "killing" people. They have no place to go without funding.

Ordförande: We are not finished; we need to have lunch now.

Konversationen ger en ödesmättad känsla – bedömarna verkar inte kunna göra så mycket åt de rådande omständigheternas makt. Trots att bedömarnas samarbete överlag tycktes fungera väl och kollegialiteten vanligtvis var god, uppstod det stundtals orosmoment som var direkt kopplat till själva uppgiften att värdera, jämföra, och sätta nya gemensamma betyg på de olika ansökningarna. Känslor av otillräcklighet, orättvisor och tidsbrist var några av problemen som kunde observeras under mötena. I konversationen ovan nämner gruppens ordförande att betygen "fluktuerar", vilket till viss del sätter fingret på en viktig aspekt som rör villkoren för hela denna bedömningsprocess. Betygen ändras ju onekligen under mötets gång, och så är det också tänkt att vara – bedömarna uppmuntras att ändra sina siffror för att nå så rimliga överenskommelser som möjligt. Man kalibrerar sina betyg, det vill säga, man justerar dem successivt.

Men bortser man nu för ett ögonblick från själva bedömningsmetoden i sig, så tycks siffrorna likväl påfallande ofta utöva en sorts makt över bedömarna, något som enligt min mening indikerar förekomsten av ankringseffekter. Flera personer vittnade om detta fenomen, som på sätt och vis verka överskrida individernas och organisationens omedelbara kontroll. En bedömare förklarade under en lunchpaus att "siffrorna tenderar att bli starka", och i en annan grupp påpekade tre av Vetenskapsrådets administratörer att de såg stora problem i att bedömarna blev som "trollbundna av siffrorna". Alla tre kunde de se ett direkt problem i att bedömarna ibland nästan mekaniskt hänvisade till siffrorna, trots att "alla vet hur missvisande siffrorna kan vara". Men när siffrorna väl finns där (när de väl har förankrats), utövar de enligt en av administratörerna en slags magisk kraft över deltagarna, en kraft som de ofta

tycks ha svårt att ignorera under förhandlingarna när de justerar sina betyg. Här har vi alltså viktig kvalitativ evidens som styrker vår grundtes – både de individuella omdömena och medelrankingen tycks på olika sätt begränsa hur långt bedömarna är redo att anpassa sina betyg – vilket pekar mot olika variationer av ankringseffekter som sker etappvis i den sociala interaktionen.

Ankringseffekterna kan och bör alltså inte enbart förstås på individnivå när det kommer till justeringarna av betyg, för bedömarna måste hela tiden anpassa sina bedömningar i relation till varandra. Ett väsentligt problem som dessutom uppstår är, att ingen har någon fullständig överblick, och att betygen kan variera ganska mycket ibland beroende på vem som satt dem. Vad jämför man med, och med vilka standarder? Detta var ett återkommande dilemma som bedömarna tog upp i sina diskussioner. I ett viktigt avseende läggs förstås själva grogrunden till ankringseffekterna individuellt då bedömarna i enskildhet läser och betygsätter ansökningarna. Det är i detta skede som bedömarna subjektivt förankrar sina omdömen och sätter sina preliminära sifferbetyg som sedan inverkar på medelrankingens utseende.¹¹ Men vad representerar siffrorna i förhållande till de uppskattade värdena, och hur inverkar detta på medelrankingen om den sedan skall anpassas i samband med diskussionerna? En sak som flera bedömare påpekade vid olika tillfällen var hur otroligt svårt det kan vara att värdera de olika kvalitetskriterierna i en ansökan; en bedömare påpekade bland annat att, ”det är lite vagt i vissa fall hur vi använder oss av betygen och vad de motsvarar i ord”. Denna vaghet kan ses som en del i ankringseffekternas uppkomst då bedömarna tvingas improvisera. Frågan är alltså vad för slags konsensus man lyckas skapa om betygen i grunden bär på en vaghet och osäkerhet. Min poäng är, att bedömarna ofta i sådana situationer måste förlita sig på sin intuition när de gör sina uppskattningar, och precis som den kognitionspsykologiska forskningen har visat är det då som ankringseffekterna också uppstår och inverkar på framtida beslut. Betygen blir till ett krux, något som en ordförande påpekade för sin grupp: ”Det är alldeles uppenbart att vi har lite svårt att hantera betygen.” Det allvarsamma i detta påstående är att det signalerar att ansökningarnas öde avgörs genom en grundläggande svårighet i att fastställa betygen (något som i princip kunde iakttas i samtliga tio beredningsgrupper).

Den osäkerhet som förekommer i bedömningsprocessen kan också få en förklaring i det som en annan ordförande påpekade i en paus, nämligen att det fanns alldeles för många olika specialiteter som skulle täckas upp i hennes grupp. Detta leder, enligt henne, till att det ibland kan vara väldigt svårt att till exempel förstå vissa algoritmer som anförs eller värdet av olika publikationer. Detta var något som nämdes i flera av grupperna (även om man även framhöll värdet av de överlappande kompetenserna). Även i grupper där detta inte uttryckligen påpekades, kunde man likväl höra flera bedömare som i förvånansvärt många fall underströk sin egen brist på kunskap i förhållande till vissa ansökningar. En bedömare påpekade till exempel att hon tyckte det kändes väldigt olustigt att behöva vara förslagsställare på en ansökan vars fält hon inte riktigt behärskade, och hon tillade: ”Hur bedömer man om man skall sätta 3 eller 5 i betyg?”. Och det var flera bedömare som uttryckte hur svårt det kunde vara att bedöma ansökningarnas originalitet eller vad som är nytt i en ansökan. Vissa menade till och med att det i vissa fall kändes fullständigt omöjligt. En bedömare ansåg att den ansökan han hade granskat inte hade något särskilt, men tillade därpå att han å andra sidan inte hade förstått allt. Ordet ”svårbedömt” användes för övrigt ganska flitigt i samtliga

¹¹ Hur mycket tid man som bedömare lägger ner på att granska en ansökan kan variera en hel del. En bedömare påpekade lite generat under lunchen att hon bara hade hunnit lägga tjugo minuter på vissa av ansökningarna, vilket förstås kan tänkas vara till en stor nackdel för den forskare som skrivit ansökan i hopp om att kommunicera en ny forskningsidé.

observerade beredningsgrupper. Även graden av genomförbarhet framhölls som vanskligt att granska om man som bedömare inte är fullkomligt insatt i metoderna och teknikerna som skall användas i ett tilltänkt projekt. Detta visade sig bland annat i en specifik situation då en av bedömarna opponerade sig mot sina kollegors sätt att se på genomförbarhet i en ansökan: "I disagree on the feasibility. But again, its not certain what feasibility is".

Ett annat problem som är kopplat till den stora mängden ansökningar som skall bedömas och rangordnas är den vanligt förekommande betygsinflationen. Flera bedömare uttalade sin frustration över problemet med betygsinflationen och dess inverkan på medelrankingens utseende. Många bedömare tenderar ofta att lägga sig någonstans i mitten av betygsskalan när de känner sig osäkra på sitt omdöme, vilket bland annat framkommer i följande bedömares utsaga: "I don't know so much about this, and if you don't know so much you tend to stay in the middle". Men i en annan beredningsgrupp syntes problemet snarare vara att man satte för många sexor, vilket fick ordförande att halvt på skämt tala om för gruppen att de måste sluta sätta så mycket höga betyg. Betygsinflationen kan följaktligen ta sig i uttryck på olika sätt, och konsekvenserna för bedömningsprocessen blir ibland därefter.

På ett allmänt plan tycks det vara svårt att fullständigt omvärdera alla värden – hänsyn tas hela tiden till den befintliga medelrankingen och till det nödvändiga behovet av att skapa konsensus i gruppen. En bedömare uttryckte på ett kritiskt sätt sin förundran över meningen med alltihop: "Men har jag förstått det hela rätt: är det meningen att jag skall kunna omvärdera hela min värdering här i gruppen? Om det inte är så, är det väl ingen mening att vi sitter här?". Bedömarna skall gärna omvärdera sin värden, men lagom mycket, och däri ligger också den stora utmaningen för beredningsgruppen som bedömarkollektiv. Konsten ligger i att finna en rimlig och acceptabel konsensus utan snedvridningar och jäv. Men osäkerhet, vaghet och oenighet, gör så att denna konst ibland ändå leder till snedvridningar när bedömarnas tolkningar går isär i jämförelserna av ansökningarnas kvalitet. Ankringseffekterna smyger sig in i bedömningsprocessen på ett ofta subtilt sätt när justeringarna skall göras – tyngdförhållandena är inte alltid enkla att genomskåda – men "siffermagin" förefaller hålla grupperna i ett visst grepp.

Vilka ankare väger tyngst?

Ankringseffekter är alltid något relativt kontextbundet. Hur mycket bedömarna är villiga att strida för sina betyg blir oftast en förhandlingsfråga som kan pendla mellan blixtrande improvisationer och det tillsynes rutinmässiga i argumentationen. Att de tidigare förankrade värdena i vissa fall får betydande effekter i relation till medelrankingen tycks på många sätt vara en ofrånkomlig ingrediens i det allmänna skapandet av konsensus. Ingen är helt immun mot det givna. En ordförande påpekade under en paus att alla i hans grupp tycktes vilja "hålla fast vid de siffror de satt hemma". Detta var kanske möjligtvis en överdrift, då de flesta bedömarna ofta visade prov på att vilja tillmötesgå varandras åsikter. Likväl får betygen en effekt som i vissa fall har att göra med att den ena bedömaren håller hårt om sina siffror eller att alla bedömarna håller hårt i sina siffror (ofta med medelrankingen som direkt eller indirekt stöd). Och även när bedömarna släpper efter på sina betyg leder det i förlängningen till att justeringen rent kollektivt sett kan resultera i snedvridning i relation till de andra ansökningarna i rankingen. Att vissa av bedömarna släpper efter på sina betyg medan andra håller i sina betyg gör att man i vissa fall tycks driva iväg från vissa ursprungsvärden och ibland istället håller fast vid andra. De förankrade värdenas effekter måste, på det kollektiva planet, också förstås i relation till de värdeomdömen som tillåts "glida" iväg, på ett i vissa situationer lite godtyckligt sätt. Som exempel kan nämnas en beredningsgrupp där en av

bedömarna hade värderat en ansökan mycket högt med sjuor och sexor. En kollega som då inte hade haft granskningsansvar för denna ansökan, tog då självmant på sig uppgiften att över natten göra en extragranskning av dess innehåll, vilket dagen efter resulterade i en betydligt lägre värdering som styrde om betygsättning. Bedömaren som hade satt denna ansökan högt i sin rangordning gick därefter istället med på att sänka den, bort från alla chanser till anslag.

En bedömare underströk under ett samtal i en av beredningsgrupperna att: ”Vi måste verkligen komma överens om hur vi kommer fram till ett medelbetyg, så att det inte blir slumpmässigt och en massa oegentligheter”. Men hur mycket kan man vrida och vända på siffrorna utan att skevheter och snedvridningar uppstår under förhandlingarna? Detta utgör ett dilemma som Vetenskapsrådets personal är väl medveten om. I varje grupp kommer det nästan alltid till en kritisk punkt då betygen bollas fram och tillbaka, och finansieringsfrågan står och väger. I dessa situationer uppstod det ofta problem att urskilja och enas om de ansökningar som är på gränsen för finansiering. Hur skulle justeringarna göras? Man jämförde ansökningarna och de ansökande personernas kompetens, och man vände upp och ner på varje kvalitetskriterium. Ibland var det omöjligt att se vad som egentligen skedde – långa stunder av tystnad och ihärdigt reflekterande. Man presenterade sina förslag på betyg, man vägde in varandras betyg, man gjorde framryckningar för att höja eller sänka. En bedömare påpekade i samband med ett kritiskt avgörande att ”Det är inte helt tillfredsställande när det är så jämt mellan olika ansökningar”. I många fall är improvisationerna ofrånkomliga, vilket inte är helt utan risk. I en situation då man tydligt började improvisera med siffrorna under den andra mötesdagen, påpekade en av Vetenskapsrådets administratörer att man inte fick gå in och börja ”manipulera” siffrorna. Och han tillade en stund senare, ”Det är verkligen jätteviktigt att ni håller er till det som sades under gårdagen, och inte börjar ändra för mycket. För då faller lite av gårdagens arbete”.

I vissa fall uppstod det långa diskussioner om hur de olika kriterierna skulle tolkas – det kunde till exempel låta så här: ”Men på den föregående ansökan bedömde ni detta som en styrka, men i den här ansökan bedömer ni det istället som en svaghet. Vi måste vara mer konsekventa”. Samtal om huruvida betygsskalan användes relativt eller absolut var inte ovanliga. I regel var man ense om att betygsskalan användes både relativt och absolut. Men hur varje enskild bedömare hade gjort tidigare, när de i enskildhet gjorde sina bedömningar, går det knappast att ge ett uttömmande svar på. Och det uppstod flera situationer där man uppenbarligen ansåg sig ha bedömt utefter helt olika standarder redan från början, vilket ansågs ge en skev och snedvriden medelranking. När det blev för många ansökningar som låg någonstans i mitten, det vill säga, på samlade betyget 5, började man i flera grupper använda sig av decimaler för att kunna urskilja dem sinsemellan (5,8 på en ansökan och 5,9 på en annan). Ibland hade man flera ansökningar som låg runt samma decimal, och i någon beredningsgrupperna röstade man med handuppräckning för att komma fram till beslut om rangordningen. Någonstans i allt detta spel om siffrorna fick man en känsla av att man delvis förlorade lite av kontrollen över vad alla justeringarna egentligen ledde någonstans. Hela ankringsproblematiken tycktes också mynna ut i frågan om tolkningen av meningen som lades i de olika värderingarna som gjordes av de olika bedömarna. En utländsk bedömare påpekade irriterat för sina kollegor (i ett fall då hon ansåg dem vara för slappa i användningen av de högre betygen) att: ”’Excellent’ must mean something!”. Detta är själva stötestenen – meningsinnehåll och sifferbetyg – hur sammanförs dessa två i en gemensam gruppbedömning? En siffra kan betyda olika saker beroende på vem som har satt den, något som också uttrycktes väl av en bedömare från en annan beredningsgrupp: ”I gave her a five, which means a lot in my case”. Skulle detta implicera att betyget i själva verket skulle kunna

tolkas som en sjuva om man översatte det till de andra bedömarnas sätt att definiera sina värdeomdömen i siffror? Häri föreligger det med andra ord ett generellt osäkerhetsmoment i bedömningsprocessen och i tolkningen om hur de olika förankrade värdena skall vägas in.

I en annan situation rådde det uppenbara skillnader i betygsättningen. Efter att man lagt fram sina argument skulle man försöka nå beslut. Den ena bedömaren hade satt 6,6,6,3, medan den andra bedömaren hade satt 2,3,2,3.¹² Ordförande för gruppen hade valt att sätta betygen 4,4,4,3. Det var uppenbart att ingen helt ville ge vika. Bedömaren som hade satt de lägsta betygen kunde under inga omständigheter gå med på den andra bedömarens entusiastiska uppfattning om den aktuella ansökan.¹³ Och på grund av att de inte kunde lösa oenigheten valde ordföranden tillslut efter en längre dispyt, att gå på sina egna siffror vid justeringen. Men ponera nu att ordförande hade satt femmor istället, vad hade hänt då – hade detta också resulterat i fyror? Eller femmor, eller kanske sexor? Ankringseffekterna sker på en något svårgripbar nivå i processen, därför att bedömarna hela tiden gör ”interna” jämförelser mellan de olika ansökningarna på ett sätt som inte riktigt visar sig för kollegorna (eller andra observatörer), men som uppenbarligen får justeringarna att luta åt ett visst håll.

I en av de observerade beredningsgrupperna uppstod det en situation där en av bedömarna hade värderat en viss ansökan som nummer ett i sin rangordning, vilket också blev ett av hennes argument under förhandlingarna. Hon påpekade frustrerat vid ett tillfälle att hon hade satt 7,7,6,3, och att en 4 på vetenskaplig kvalitet, som hennes kollegor ville sätta, var på tok för lågt. Efter en stund utbrast hon högt: ”Come on!”. En av de andra bedömarna svarade då, att man hela tiden bedömer och värderar de olika ansökningarna genom att använda sig av olika standarder och tolkningar av vad som definierar kvalitet. Hon tog därpå upp ett annat exempel från en tidigare diskussion där en viss ansökan hade värderats mycket lägre på ett av betygsriterierna, men som nu tydligt räknades på ett helt annat sätt för den aktuella ansökan. Siffrorna blir starka och påverkar hela tiden bedömningsprocessen på ett sätt som inte alltid är önskvärt eller rättvisande. När det råder fullständig enighet om vilka ansökningar som är bäst, upplevs allt som relativt problemfritt, men när oenighet uppstår blir det desto tydligare att ”siffermagin” gör sig gällande i den sociala interaktionen. Under en kaffepaus påpekade tre bedömare som jag talade med, att man förvånansvärt ofta är ense om vilka ansökningar som är i toppen och vilka som är i botten. Detta är en bild som även lyfts fram i tidigare forskning (Lamont 2009). Likväl behöver detta inte implicera att alla bedömare skulle vara fullständigt eniga ändå. En bedömare skämtade exempelvis med sina kollegor genom att hävda att det förekom topansökningar i botten på gruppens medelranking. Hans poäng var att någonting håller dessa ansökningar nere där på botten och gör att de aldrig kommer upp till ytan. Det som håller dem nere är medelrankingen så som ett kollektivt ankare. Varje bedömare kan visserligen be om att få ”höja” de ansökningar som hamnat långt ner på rankingen, vilket betyder att man tar denna ansökan till diskussion, trots att den inte kvalificerar sig för anslag. Men detta förfaringsätt leder i princip aldrig till något anslag, utan stannar i regel vid en finputsning av betygen, vilket visserligen betyder att positivare signaler sänds tillbaka till den ansökande personen ifråga. Vilket ankare som väger tyngst är alltid en fråga som går via gruppen och som sker genom otaliga jämförelser som inte alltid synliggörs.

¹² Bedömningsriterierna är: *nyttänkande och originalitet, vetenskaplig kvalitet, sökandes kompetens, och genomförbarhet.*

¹³ Hon påpekade vid annat tillfälle att hon hade varit ”in a bad mood” när hon läst vissa ansökningar, vilket eventuellt skulle kunna förklara hennes hårda betygsättning. Se särskilt Birte Englich och Kirsten Soders artikel ”Moody experts – How mood and expertise influence judgmental anchoring” (2009) om skillnaden mellan glada respektive nedstämda bedömare för uppkomsten av ankringseffekter.

Kollektiva ankringseffekter och tillfälligheternas makt

Användningen av medelranking är så som jag redan tidigare har påpekat högst väsentlig för gruppens bedömningsarbete – genom den erhåller gruppen en vägledande översikt som sätter de individuella betygen i nytt ljus. Utan medelrankingen skulle det bli betydligt svårare för bedömarna att få ett bra grepp om det allmänna utgångsläget när de skall jämföra de olika ansökningarnas kvalitet. På grund av att man i regel är begränsade till att vara tre granskare per ansökan, lämnar det dock av naturliga skäl vissa luckor i gruppens helhetssyn. Medelrankingen hjälper följaktligen till med att fylla igen vissa av luckorna. Likväl ger den upphov till en viss ambivalens. En ordförande underströk glatt när första mötesdagen började hur fantastiskt bra medelrankingen är som underlag, ”den ger en viktig gruppbild”. Lite senare samma dag påpekade han dock, vid upprepade tillfällen, behovet av ett ”bättre system” för att räkna ut denna medelranking. Under observationerna var det flera bedömare som talade om slumpens inverkan i arbetet. En lite extrem åsikt uttrycktes under en kaffepaus då en av bedömarna klargjorde att granskningen av särskilt unga forskares ansökningar är som ett lotteri som lika gärna skulle kunna lösas med hjälp av en dator. Några andra bedömare påpekade då att detta i viss mån också även gäller för de mer etablerade forskarnas ansökningar, vilket i viss mån bekräftar det som sociologen Stephen Cole och hans medarbetare kom fram till i sin studie från 1981 om slumpens effekter vid skapandet av konsensus. Ett exempel på ambivalens som kunde observeras i studien av Vetenskapsrådet, kan illustreras genom följande justeringsomgång:

Ordförande: ”So what should we say? You two have given him 4, and X was the only one that understood it, he gave it a 6”.

Bedömare 1: ”You had too many good ones”.

Bedömare 2: ”I can go either way”.

Bedömare 3: ”I am not an expert in this field. But I could easily bring this guy up. I just had to many in the middle, that’s why I ranked him a bit low”.

När bedömarna i vissa fall håller hårt i sina siffror, medan de i andra fall släpper efter (direkt eller indirekt i förhållande till medelrankingen), måste det få effekter på den aggregerade nivån. Det är dock inte så enkelt att exakt identifiera på vilka sätt dessa effekter påverkar det kollektiva slutresultatet. Men att någonting sker som får de involverade själva att reagera, det är alldeles tydligt i denna studie. Tidsbrist, osäkerhet, och oenighet är nyckelfaktorer som bidrar till att kollektiva ankringseffekter uppstår när man justerar sina omdömen. Flera bedömare vittnade, som jag redan tidigare har påpekat, om slumpens inverkan. Men även tröghet och förutsägbarhet var något som underströks i de flesta beredningsgrupperna, vilket hade sin upprinnelse i upplevelsen av att de ansökningar som beviljades anslag ofta kändes givna från början. I flera situationer gick det att se hur vissa bedömare förgäves kämpade för att förändra gruppens ståndpunkt. Denna blandning av förutsägbarhet och oförutsägbarhet går på sätt och vis hand i hand när man ser till hela bedömningsprocessen i sig och att lutningarna i bedömning faktiskt sker på gruppnivå, och inte bara på individnivå. För trots att alltsammans ursprungligen växer fram ur individuella omdömen, så omformas de av organisationsvillkoren och de specifika bedömningsmetoder som används. Det rör sig hela tiden om ett komplicerat samband mellan en stor mängd olika ankringsvärden som skall justeras och som ger upphov till att tillfälligheternas makt tränger in i det tillsynes självklara i rankingen. Hur långt ifrån de tidigare förankrade värdena man förflyttar sig som grupp, varierar oftast från fall till fall, och graderna av godtycklighet likaså. Och vad som styr grupperna när de skall skapa konsensus, förefaller inte alltid vara helt klart, något som vid ett tillfälle uppmärksammades av en bedömare som i sin förvåning uttryckte sig på följande vis: ”How did we reach this – considering the original grades given?”.

Slutord

I den här undersökningen har min främsta avsikt varit att belysa hur det akademiska omdömet påverkas av effekter som delvis ligger bortom organisationens och bedömarnas direkta kontroll. Jag har velat ge en bild av hur det akademiska omdömet organiseras, kvantifieras och omvandlas på ett sätt som i vissa avseenden leder till snedvridningar i samband med skapandet av konsensus. Jag har tagit min utgångspunkt i den kognitionspsykologiska teorin om ankringseffekter. Detta har dock skett i en mer sociologiskt orienterad riktning – ankringseffekterna har betraktats på en aggregerad kollektiv nivå. Min utgångspunkt har inte varit att försöka bevisa eller mäta dessa ankringseffekter, vilket skulle krävt helt andra metoder. Istället har jag i retrospektiv velat *visa* att ankringseffekterna faktiskt uppstår via gruppens medelranking och genom de ständiga förhandlingarna som sker beträffande ansökningarnas värde. I denna studie var det ett flertal personer i olika grupper som lät tala om ”siffermagi”, och i nästan varje grupp nämndes vid åtminstone något tillfälle slumpens inflytande på bedömningsprocessen. Min ambition har varit att visa att denna slump bättre kan förstås som ett resultat av *kollektiva ankringseffekter*, som uppstår genom den stora mängd förankringar, justeringar och omvärderingar som sker i förhållandena mellan bedömarna. För att få grepp om ankringseffekterna som kollektivt fenomen är det dessutom viktigt att ta fasta på den osäkerhet och de oenigheter som förekommer i princip varje beredningsgrupp. Bedömarna tampas många gånger både med sina egna och med varandras omdömen ifråga om vad som är nyskapande och spännande forskning. Vad som är en rimlig betygssättning (givet medelrankingen) kräver ofta ett visst mått av intuition och improvisation. Detta öppnar i sin tur upp för olika grader av ankringseffekter på det slutgiltiga resultatet av bedömningsprocessen. Ibland får de komplicerade relationerna mellan de olika förankringarna och omjusteringarna grupperna att driva iväg i sina samlade bedömningar, och ibland ligger grupperna helt stilla kring ett kollektivt ankare. Vad som är rimligt eller korrekt i den konsensus som man lyckas nå, trots oenigheter och osäkerhet i bedömningarna, skall jag lämna därhän. Sammanfattningsvis har min ambition varit att illustrerar en faktor som på ett lite svårgripbart sätt påverkar det akademiska omdömet och som i förlängningen får konsekvenser för vilka forskningsprojekt som ges de ekonomiska möjligheterna att utvecklas i framtiden.

Litteraturförteckning

Black, M. (1962) *Models and Metaphors: Studies in language and philosophy*, Ithaca: Cornell University Press.

Blau, P. M. (1964) *Exchange and Power in Social Life*, New Brunswick: Transaction Publishers.

Bourdieu, P. (1975) "The specificity of the scientific field and the social conditions of the progress of reason", *Social Science Information*, 14 (6), 19-47.

Certeau, M. de (1984) *The Practice of Everyday Life*, L.A.: University of California Press.

Chapman, G.B. & Johnson E.J. (1994) "The Limits of Anchoring", *Journal of Behavioral Decision Making*, 7, 223-242.

Chubin, D.E. & Hackett, E.J. (1990) *Peerless Science: Peer Review and U.S. Science Policy*, New York: State University of New York Press.

Cole, S., Cole, J.R. & Simon, G.A. (1981) "Chance and Consensus in Peer Review", *Science*, New Series, 214 (4523), 881-886.

Cole, S. (1992) *Making Science: Between Nature and Society*, Cambridge, Mass: Harvard University Press.

Englich, B. & Mussweiler, T. (2001) "Sentencing under Uncertainty: Anchoring effects in the courtroom", *Journal of Applied Social Psychology*, 31, 1535-1551.

Englich, B. & Soder, K. (2009) "Moody experts: How mood and expertise influence judgmental anchoring", *Judgment and Decision Making*, 4 (1), 41-50.

Epley, N., Keyser, B., Van Boven, L. & Gilovich, T. (2004) "Perspective Taking as Egocentric Anchoring and Adjustment", *Journal of Personality and Social Psychology*, 87 (3), 327-339.

Epley, N. & Gilovich, T. (2010) "Anchoring unbound", *Journal of Consumer Psychology*, 20, 20-24, ScienceDirect.

Fogelholm, M., Leppinen, S., Auvinen, A., Raitanen, J., Nuutinen, A., Väänänen, K. (2012) "Panel discussion does not improve reliability of peer review for medical research grant proposals", *Journal of Clinical Epidemiology*, 65(1), 47-52.

Galinsky, A.D. & Mussweiler, T. (2001) "First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus", *Journal of Personality and Social Psychology*, 81 (4), 657-669.

Graves, N., Barnett, A.G. & Clarke, P. (2011) "Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel", *BMJ* 343.

- Haug, C. (2013) "Consensus decision-making in meetings as an interactive accomplishment: silence without silencing?", Paper presenterat vid 29:e EGOS kollokvium, 4-6 July Montreal.
- Hemlin, S. (1996) "Research on research evaluation", *Social Epistemology*, 10 (2), 209-250.
- Huutoniemi, K. (2012) "Communicating and compromising on disciplinary expertise in the peer review of research proposals", *Social Studies of Science*, 42(6), 897-921.
- Jacowitz, K.E. & Kahneman, D. (1995) "Measures of Anchoring in Estimation Tasks", *Personality and Social Psychology Bulletin* 21, 1161-1165.
- Kahneman, D. (2011) *Thinking, Fast and Slow*, Farrar, Straus and Giroux.
- Kant, I. (1790) *Kritik av omdömeskraften*, Stockholm: Thales, 2007.
- Lamont, M. (2009) *How Professors Think: Inside the Curious World of Academic Judgment*, Cambridge: Harvard University Press.
- Langfeldt, L. (2006) "The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome", *Social Studies of Science*, 31 (6), 820-841.
- March, J.G. & Olsen, J.P. (1976) *Ambiguity and Choice in Organizations*, Oslo: Universitetsforlaget.
- McElroy, T. & Dowd, K. (2007) "Susceptibility to anchoring effects: How openness-to-experience influence responses to anchoring cues", *Judgment and Decision Making*, 2 (1), 48-53.
- Merton, R.K. (1973) *The Sociology of Science*, Chicago: University of Chicago Press.
- Mussweiler, T., Strack, F. & Pfeiffer, T. (2000) "Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility", *Personality and Social Psychology Bulletin*, 26, 1142-1150.
- Mussweiler, T. (2001) "The durability of anchoring effects", *European Journal of Social Psychology*, 31 (4), 431-442.
- Mussweiler, T. & Englich, B (2005) "Subliminal anchoring: Judgmental consequences and underlying mechanisms", *Organizational Behavior and Human Decision Processes*, 98, 133-143.
- Nortcraft, G.B. & Neale, M.A. (1987) "Experts, amateurs, and real estate: An anchoring-and adjustment perspective on property pricing decisions", *Organizational Behavior and Human Decision Processes*, 19, 84-97.
- Oppenheimer, D.M., LeBoeuf, R.A., & Brewer, N.T. (2007) "Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming", *Cognition*, ScienceDirect.

- Plous, S. (1989) "Thinking the unthinkable: The effects of anchoring on likelihood estimates of nuclear war", *Journal of Applied Social Psychology*, 19, 67-91.
- Porter, R. (2005). "What do grant reviewers really want, anyway?", *The Journal of Research Administration*, 36 (2), 47–55.
- Ricoeur, P. (1978) *The Rule of Metaphor: Multi-Disciplinary Studies in the Creation of Meaning in Language*, London: Routledge and Kegan Paul.
- Ritov, I. (1996) "Anchoring in simulated competitive market negotiation", *Organizational Behavior and Human Decision Processes*, 67, 16-25.
- Simmel, G. (1908) *Hur är samhället möjligt?*, Bokförlaget Korpen, 1995.
- Simmons, J., LeBoeuf, R. & Nelson, L. (2010) "The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors?", *Journal of Personality and Social Psychology*, 99 (6), 917–932.
- Simon, H. (1957) *Models of Man: Social and Rational*, New York: John Wiley and Sons.
- Swedberg, R. (2014) *The Art of Social Theory*, New Jersey: Princeton University Press.
- Thedvall, R. (2013) "Punctuated entries: Doing fieldworks in policy meetings in the European Union", i Garsten & Nyqvist (Red.), *Organisational Anthropology*, London: Pluto Press.
- Travis, G.D.L. & Collins, H.M. (1991) "New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System", *Science, Technology, & Human Values*, 16 (3), 322-341.
- Tversky, A. & Kahneman, D. (1974) "Judgment under Uncertainty: Heuristics and Biases", *Science*, New Series, 185 (4157), 1124-1131.
- Tversky, A. & Kahneman, D. (1981) "The Framing of Decisions and the Psychology of Choice", *Science*, New Series, 211 (4481), 453-458.
- Wilson, T., Houston, C., Etling, K., & Brekke, N. (1996) "A New Look at Anchoring Effects: Basic anchoring and its antecedents", *Journal of Experimental Psychology: General* 125 (4), 387-402.
- Zuckerman, H. (1977) *Scientific Elites: Nobel Laureates in the United States*, NY: Free Press.
- Östergren, K. (1988) *Ankare*, Bonniers förlag.