



Vetenskapsrådet

HUR MYCKET CITERAS SVENSKA PUBLIKATIONER?

Bibliometrisk översikt över Sveriges vetenskapliga publicering mellan 1982 och 2004

HUR MYCKET CITERAS SVENSKA PUBLIKATIONER?

**Bibliometrisk översikt över Sveriges vetenskapliga
publicering mellan 1982 och 2004**

Staffan Karlsson & Daniel Wadskog

Analysenheten, Vetenskapsrådet

HUR MYCKET CITERAS SVENSKA PUBLIKATIONER?

Bibliometrisk översikt över Sveriges vetenskapliga publicering mellan 1982 och 2004

Rapporten kan beställas på www.vr.se

VETENSKAPSRÅDET

103 78 Stockholm

Vetenskapsrådet

ISSN 1651-7350

ISBN 91-7307-095-5

Grafisk Form och omslagsillustration: Erik Hagbard Couchér, Vetenskapsrådet

Layout och illustrationer: Maria Wågberg

Tryck: CM Digitaltryck, Bromma 2006

SAMMANFATTNING

I denna rapport presenterar vi statistik över hur svenska vetenskapliga publikationer citeras jämfört med andra länders publikationer och hur publiceringsvolym och citering varierar mellan lärosäten och ämnesområden. Analysen är baserad på fältnormaliserad medelcitering av originalartiklar och översiktsartiklar ("articles" och "reviews") i internationella tidskrifter där självciteringar är borttagna. Fältnormaliseringen innebär att antalet citat till en publikation relateras till världsmedelvärdet i det eller de fält som tidskriften tillhör.

I internationell jämförelse har Sverige legat på en relativt konstant citeringsnivå de senaste 20 åren, 10 ± 5 % över världsgenomsnittet. Citeringen av de svenska publikationerna har blivit jämnare, dvs. andelen högt respektive lågt citerade publikationer har minskat. Till skillnad från Sverige, har flera andra europeiska länder – såsom Danmark, Nederländerna och Tyskland – fått betydligt fler citeringar av sina publikationer under det senaste årtiondet. Framförallt har dessa länder fått en ökad andel högt citerade publikationer. Publikationer från Nederländerna och Danmark citerades 22 % över världsgenomsnittet åren 2000–2002 medan svenska publikationer citerades 13 % över världsgenomsnittet under samma period. Publikationerna från flera europeiska länder (förutom de som nämnts ovan bl.a. också Frankrike och Spanien) har blivit citerade i ökande utsträckning, vilket resulterat i att Europamedelvärdet ökat i förhållande till resten av världen. Jämfört med det europeiska medelvärdet har trenden därför varit negativ för Sverige under hela den studerade perioden.

Skogs- och jordbruksvetenskap, naturvetenskap och teknik är de områden där svenska publikationer citeras mest. Inom dessa områden ligger medelciteringen 25–30 % över världsgenomsnittet. Publikationer i medicinska tidskrifter citeras 3 % mindre än världsgenomsnittet. Ser man på hur mycket publikationer från olika lärosäten citeras finner man att variationen är relativt liten; av 18 lärosäten med minst 20 publikationer per år under perioden 1982–2002 hade 14 en medelcitering över världsgenomsnittet.

INNEHÅLL

1 INLEDNING	5
2 ALLMÄNT OM CITERINGSSTATISTIK	7
2.1 Citeringsanalys som utvärderingsverktyg	7
2.2 Bakgrund till urval av publikationer och metoder	II
3 INTERNATIONELL ÖVERSIKT	20
3.1 Utveckling mellan 1982 och 2003	20
3.2 Länderjämförelse baserad på medelvärden för 2000-2002	24
3.3 Varför har inte Sverige utvecklats lika bra som övriga Europa?	26
4 VARIATIONER MELLAN ÄMNESOMRÅDEN OCH LÄROSÄTEN	32
4.1 Vem publicerar var? Författarnas "fakultetstillhörighet" jämfört med tidskrifternas ämnestillhörighet	32
4.2 Variation inom vetenskapsområden	35
4.3 Variationer mellan lärosäten	38
5 SLUTKOMMENTAR	49
6 CITERAD LITTERATUR	51
METODBILAGA	53
Gruppering av svenska lärosäten	53
Databasen	53
Normalisering av citeringar	53
Andelarna högt och lågt citerade publikationer	56
Eliminering av självciteringar	57

1 INLEDNING

I den svenska forskningspolitiska diskussionen har det vid upprepade tillfällen de senaste åren påpekats att citeringsstatistiken för svenska vetenskapliga publikationer har visat en negativ trend. Även i den bibliometriska litteraturen har Sveriges tillbakagång diskuterats.¹ Samtidigt hävdar sig Sverige ofta relativt väl i internationella jämförelser av forskningskvalité. En orsak till att omdömena varierar är att det finns många sätt att bedöma vetenskaplig kvalité och rangordna länder eller lärosäten. I denna rapport sammanfattar vi det aktuella läget när det gäller hur svenska forskningsresultat publiceras och citeras.

Som svenska publikationer räknas de som har minst en författare med svensk adress. Om inget annat anges avses i denna rapport endast publikationer av typerna originalartiklar och översiktsartiklar ("articles" och "reviews"). Konferenssammanfattningar ("meeting abstracts"), brev ("letters")² etc. är alltså inte med i analysen.³

Rapporten består av tre delar: Först presenteras några allmänna egenskaper hos citeringsstatistik med huvudsakligen svenska publikationer som exempel (avsnitt 2), därefter gör vi en internationell översikt där Sverige jämförs med 32 andra länder (avsnitt 3). Till sist gör vi en översikt över mönster inom Sverige: vem producerar vad och hur uppmärksammas, citeras, olika discipliner och lärosäten (avsnitt 4).

Den analys som presenteras här är baserad på Vetenskapsrådets publikationsdatabas. Grundmaterialet i denna databas är inköpt från Thomson Scientific och är detsamma som finns tillgängligt via Web of Science. Vetenskapsrådets publikationsdatabas innehåller dessutom ytterligare ett antal variabler, såsom rättade svenska organisationsnamn, antal citeringar en viss tid efter publiceringen och den globala medelciteringen i olika forskningsfält (se bilaga 1 för en mer detaljerad beskrivning).

Som nämnts ovan används många olika kriterier när länders eller lärosätens forskning granskas och rangordnas. Detta gäller även för citeringsanalys. Det finns ett antal sätt att göra jämförelserna, alltifrån enkla summeringar av citat till relativt komplicerade index som tar hänsyn till exempelvis

¹ Ingversen & Wormell 1999, Glänzel 2000, Glänzel et al 2003.

² "Notes" fanns bara med fram till 1996, därefter kodas dessa som "articles".

³ Publikationerna i tidskrifter där "letters" ingår i titeln klassas huvudsakligen som "articles", t.ex. "Physical Review Letters" innehåller till 95 % publikationer av typen "articles" (inga "letters").

sampublicering och självcitering. I tidigare rapporter från Vetenskapsrådet har framför allt två olika index använts. I det ena jämförs Sverige och andra europeiska länder med Europagenomsnittet och i det andra görs en global fältnormalisering, dvs citeringen uttrycks i relation till världsmedelvärdet i det ämnesfält som tidskriften där en publikation är tryckt tillhör.⁴ Beroende på hur citeringarna summeras väger olika ämnesområden olika tungt vid beräkning av medelvärden och slutsatserna kan variera något beroende på vilket index som väljs. I denna rapport har vi valt att använda ett index som tar hänsyn till variationer i citeringstraditioner mellan olika områden och till den "inflation" i citeringsnivån som förekommit under många år. Vi kommer därför huvudsakligen att presentera medelvärden baserade på fältnormaliserade citeringar.

Tack till alla som givit kommentarer på manuskriptet; speciellt till Ulf Heyman, Gunnar Leman, Jonas Lundberg, Olle Persson, Bo Sundqvist och kollegorna på analysenheten.

Certain data included herein are derived from the Science Citation Index Expanded, prepared by Thomson Scientific®, Philadelphia, Pennsylvania, USA. © Copyright Thomson Scientific® 2006. All rights reserved.

⁴ Se metodikbilagan för detaljer.

2 ALLMÄNT OM CITERINGSSTATISTIK

Publicerings- och citeringsstatistik används allt oftare vid nationella och internationella jämförelser och utvärderingar av forskningens volym och kvalitet. En lång rad exempel på detta finns i en nyutkommen bok om citeringsanalys som utvärderingsverktyg, Moed (2005) "Citation analysis in research evaluation". Användningen av bibliometri för utvärderingar är omdiskuterad. I detta kapitel summerar vi först diskussionen om de problem som finns med citeringsanalys som utvärderingsverktyg (avsnitt 2.1). Vidare diskuterar vi några statistiska egenskaper hos citeringar och metoder vi använder (avsnitt 2.2) som en bakgrund till den jämförande statistiken i kapitel 3 och 4. I denna bakgrundsbeskrivning använder vi huvudsakligen svenska publikationer från 1998 t.o.m. 2002, dvs. den senaste 5-årsperioden där vi har ett komplett tvåårigt citeringsfönster (dvs. citeringarna är summerade för publiceringsåret och de två följande åren).

2.1 Citeringsanalys som utvärderingsverktyg

Antalet citeringar som görs till ett vetenskapligt arbete kan ses som ett mått på den uppmärksamhet som arbetet får bland andra forskare. Ofta tolkas uppmärksamheten i sin tur som ett mått på publikationens kvalitet eller "impact". Men sambandet mellan antalet citeringar och kvalitet är inte självklart. Det finns många olika skäl till varför ett arbete citeras, av vilka en del inte alls är relaterade till kvalitet.⁵ En invändning mot att använda citeringar som ett kvalitetsmått är att många av citeringarna kan vara "negativa", dvs. att vissa arbeten citeras för att de har brister snarare än att de är av hög kvalitet. I en studie av orsaker till varför arbeten citeras (Case & Higgins 2000) var den vanligaste orsaken att det citerade arbetet hade positiva egenskaper. Därefter kom "sociala skäl", dvs. viktiga arbeten inom fältet citerades för att författarna ville visa att de kände till dessa publikationer. På tredje plats bland skälen kom att författarna var kritiska till det citerade arbetets kvalitet. Beräkningar har gjorts som visar att andelen negativa citeringar ofta uppgår till omkring 10 % av alla citeringar (Moed 2005a, sidan 84). Case och Higgins (2000) refererar ett antal uppskattningar där andelen negativa citat varierar mellan 2 % och 14 %. Trots de skiftande orsakerna

⁵ Se t.ex. Kotiaho 1999, Case & Higgins 2000.

till att ett arbete citeras har bibliometrisk forskning visat att det finns ett positivt samband mellan antalet citeringar och forskares uppfattning om publikationers kvalitet.⁶

Flera inflytelserika bibliometriska forskare har vid upprepade tillfällen poängterat vikten av att bibliometri används med försiktighet och med kunskap om metodikens brister och problem.⁷ Här ska vi endast kort peka på två faktorer som är viktiga att ha i åtanke och beskriva hur vi anpassat metodiken i denna analys för att minimera problemen.

A) Citeringsmönstret

Många vetenskapliga publikationer får bara några få citeringar medan ett fåtal arbeten blir mycket citerade. Citeringarna fördelas dessutom ojämnt över tiden. Detta påverkar tolkningen av citeringsstatistiken på flera olika sätt.

Några få högt citerade publikationer kan ha ett relativt stort inflytande på medelvärdet även då medelvärdet är baserat på stora material. All citeringsstatistik i denna rapport är därför baserad på medelvärden för 3 eller 5 år. De tidsserier som presenteras är baserade på glidande 3-årsmedelvärden. Vidare har vi valt att i flera fall komplettera presentationen av medelvärden med statistik över andelarna lågt respektive högt citerade publikationer.

Citeringsmönstret gör att statistiken varierar betydligt mellan olika år, framför allt när medelvärden är baserade på relativt små material (100-tals publikationer och mindre). Jämförelser mellan små enheter (t.ex. svenska publikationer som uppdelats på ämnesområden och/eller lärosäten) kräver därför större försiktighet än jämförelser mellan större enheter. Vid jämförelser mellan svenska lärosäten och/eller ämnesområden har vi därför valt att sätta en nedre gräns vid 100 publikationer (minst 20 publikationer per år under en femårsperiod) för att ta med ett ämne eller lärosäte.

B) Databasens täckning

Det är viktigt att notera att databasens täckning varierar mellan olika områden. Täckningen är god i naturvetenskap och medicin. Inom teknikvetenskap anses den som ganska bra. Inom samhällsvetenskap är täckningen relativt god för psykologi och ekonomiska ämnen men sämre för övriga områden.⁸ Exempelvis återfanns endast ca 20 % av de samhällsvetenskapli-

⁶ Aksnes 2005 se Wallin 2005 sid 266 för en kritisk diskussion om detta.

⁷ Se t.ex. van Raan 2004 och Moed 2005a.

⁸ Se Moed (2005a) kapitel 7. En diskussion om databasens begränsningar för humaniora och samhällsvetenskap finns också i Archambault m.fl. 2006.

ga publikationerna från Uppsala universitet i databasen.⁹ Inom humaniora, där publicering ofta sker i form av monografier och bokkapitel, har databasen stora brister eftersom den nästan enbart innehåller tidskrifter. Vi är därför återhållsamma med att presentera analyser av humaniora. Man bör också komma ihåg att även inom samhällsvetenskap varierar databasens täckning mellan olika områden. Resultat för båda dessa områden bör alltså tolkas med försiktighet.

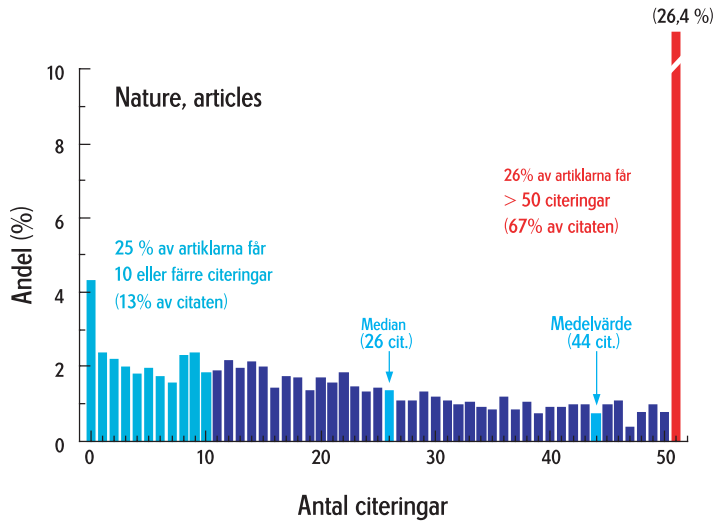
Moed (2005a, sidan 225) sammanfattar diskussionen om tolkningen av citeringar: "... analysing aggregates does not necessarily rule out all sorts of bias, but may, when properly conducted, reduce bias to a considerable extent". Även en annan inflytelserik bibliometriker, van Raan, gör en liknande bedömning (van Raan 2005, sidan 135).

"Journal impact factors" (JIF) är ett mått på medelciteringen av publikationerna i en viss tidskrift. I diskussioner om citeringsstatistik som förs i icke-bibliometriska vetenskapliga tidskrifter har JIF varit i fokus i åtskilliga redaktionella notiser, "brev till redaktören" etc. Ofta är det forskare och/eller tidskriftsredaktörer som ifrågasätter bibliometrin och framför allt användningen av JIF för kvalitetsbedömning av publikationer. Eftersom det tar flera år innan det går att få ett mått på hur en publikation citeras är det lockande att använda JIF-värden som ett (preliminärt) mått på en publikations kvalitet. Detta kan anses korrekt på det sättet att JIF återspeglar konkurrensen om utrymmet och därmed svårigheten att få ett manuskript accepterat i en viss tidskrift.

Ett problem med att värdera publikationer utifrån JIF för den tidskrift de är publicerade i är att citeringarna är mycket skevt fördelade och några relativt få högt citerade publikationer står för en stor del av alla citat. Detta är också fallet för publikationer i prestigetidskrifter, t.ex. Nature, vilket illustreras i figur 1. I Nature citeras 25 % av artiklarna 10 gånger eller mindre, dvs mindre än en fjärdedel av medelciteringen för tidskriften. Den högst citerade fjärdedelen får 67 % av alla citeringar.

⁹ Gottfries opublicerad undersökning citerad i Cavallin & Lindblad 2006.

Figur 1. Variation i antalet citeringar efter två år för artiklar i Nature publicerade mellan 2000 och 2002. Röda stapeln, visar andelen artiklar citerade mer än 50 gånger. Notera att urvalet här inte är begränsat till svenska publikationer.



De flesta forskare skickar säkerligen sina bästa manuskript till mer prestigefyllda tidskrifter. Segeln (1994) fann inget som tyder på att en publikation som är införd i en prestigetidskrift citeras mer än en liknande publikation i en mer ordinär tidskrift. Ett resultat som tyder på motsatsen är att en relativt vanlig orsak till citering är att den citerade författaren var välkänd eller att tidskriften hade hög prestige (Case & Higgins 2000).

Medvetenheten om att bibliometriska citeringsmått är vanligt förekommande utvärderingsverktyg har resulterat i att både författare och tidskrifter anpassat sig till detta.¹⁰ Det förekommer att redaktörer systematiskt försöker höja sina tidskrifters JIF-värden på olika sätt¹¹, t.ex. genom att ändra sin publiceringspolicy¹². Den ökande medvetenheten bland forskare har lett till att tidskrifter med höga JIF-värden överhoppas av manuskript vilket i sin tur resulterar i extremt höga andelar refuserade manuskript hos dessa tidskrifter¹³. En annan effekt är att ett kortsiktigt perspektiv får ökad betydelse; det kan uppfattas som viktigare (för både karriär och/eller för möjligheterna att få anslag) att få snabb uppmärksamhet (t.ex. genom publikationer

¹⁰ Weingart 2005.

¹¹ Agrawal 2005, Wallin 2005.

¹² Zetterström 2002.

¹³ Lawrence 2003.

i tidskrifter med höga JIF) än att föra forskningen långsiktigt framåt¹⁴. JIF är förmodligen det mest kortsiktiga av de bibliometriska måtten.

På grund av dessa och andra problem har det i olika sammanhang varnats för att använda JIF.¹⁵ Vi har här valt att inte ytterligare förstärka fokuseringen på JIF och redovisar ingen statistik baserad på JIF eller liknande mått.

2.2 Bakgrund till urval av publikationer och metoder

A) Typer av publikationer

Databasen innehåller publikationer av flera olika typer. Den i särklass vanligaste typen är vetenskapliga originalartiklar ("articles"). Därefter kommer sammanfattningar från konferenser ("meeting abstracts") och översiktsartiklar ("reviews"). Det finns stora skillnader i hur dessa typer av publikationer citeras; från 0,2 citat per publikation de första två åren för konferenssammanfattningar och "övrigt" till 12,0 för översiktsartiklar.

Tabell 1. Svenska publikationer 1998-2002 fördelade på typ av publikation. Tabellen är baserad på samtliga citeringar, även självcit. Publikationerna är inte fraktioniserade.

Typ av Publikation	Andel (%)	Antal citeringar efter 2 år		
		Medelvärde	Median	Maximum
Articles	83,5	4,9	2	1656
Meeting abstracts*	8,4	0,2	0	16
Reviews	2,9	12,0	6	1541
Editorial material	1,8	2,2	0	126
Letters	1,8	2,1	0	164
Övrigt	1,5	0,2	0	32

* Thomson Scientific samlar "proceedings" i en annan databas ("ISI Proceedings").

Den relativt stora skillnaden mellan medelvärdet och medianvärde när det gäller antalet citeringar i tabell 1 visar att citeringarna är skevt fördelade. Detta är ett välkänt mönster; många publikationer får inga eller bara några

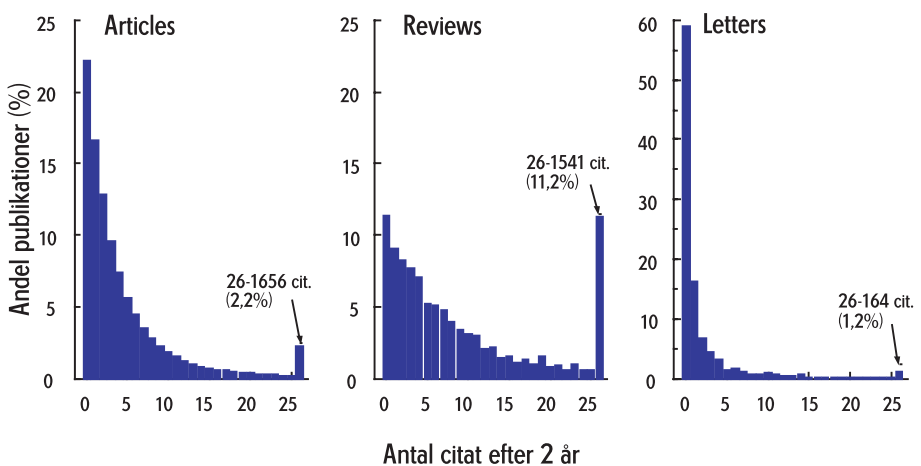
¹⁴ Lawrence 2003.

¹⁵Se t.ex. diskussionerna i Wallin 2005 och Moed 2005b.

få citeringar medan ett fåtal artiklar blir mycket citerade (figur 2). Två år efter publiceringen hade 22 % av artiklarna noll citeringar. Drygt hälften (52 %) av artiklarna citerades 2 gånger eller färre medan bara 0,6 % av dem citerades fler än 50 gånger.

Brev ("letters") citeras minst av de tre typerna av publikationer som illustreras i figur 2. Nästan 60 % av breven citerades inte under de första två åren efter publiceringen och endast 1 % av dem fick mer än 25 citeringar. Översiktsartiklar ("reviews") visar det motsatta mönstret med hög medelcitering, en relativt liten andel med noll citat och en stor andel högt citerade. I exemplet med Nature ovan visades att citeringarna är skevt fördelade på så sätt att en relativt stor andel artiklar är högt citerade. För mer "normala" tidskrifter är skevhet mot ett stort antal lågt citerade publikationer mer markerad (större andel lågt citerade). Detta innebär att medelvärdet inte är representativt för flertalet av publikationerna. Några få högt publicerade arbeten kan påverka medelciteringen för alla svenska publikationer ett visst år relativt mycket. Om man tar ut de högst citerade publikationerna, som motsvarar 1 % av det totala antalet publikationer, står dessa för ca 16 % av alla citeringar. Eftersom antalet mycket högt citerade publikationer är litet och de kan förväntas vara ojämnt fördelade över tiden kan alltså spridda förekomster av dessa högt citerade artiklar göra att medelvärden för länder eller lärosäten varierar relativt mycket från år till år.

Figur 2. Fördelning av antalet citat per publikation två år efter publiceringen. Figuren är baserad på publikationer mellan 1998 och 2002 med minst en svensk författare. Medelvärden och median för samma material redovisas i tabell 1. Figuren är baserad på samtliga citeringar, även självcitater. Publikationer är inte fraktioniserade.*



* Se metodbilagan för detaljer

I tabell 2 redovisas hur mycket de högt citerade publikationerna bidrar till det totala antalet citeringar till svenska publikationer. Exempelvis får ”topp 5 %” av de högst citerade publikationerna ungefär en tredjedel (29-38 %) av alla citat. Om materialet bryts upp på mindre enheter, såsom ämnesområden eller lärosäten, kan betydelsen av några få högt citerade publikationer bli avsevärt större.¹⁶

Tabell 2. Andelen högt citerade svenska publikationer och det relativa bidraget från dessa till det totala antalet citeringar (båda som procent). Andelen har beräknats för två definitioner av högt citerade; de med 25 eller fler citeringar (25+) och alla publikationer som citeras lika mycket eller mer än den 95e percentilen av världsproduktionen (topp 5 %). Tabellen är baserad på samtliga citeringar, även självcitater.

Definition av högt citerad	Articles		Reviews		Letters	
	Andel av publ.	Andel av citeringarna	Andel av publ.	Andel av citeringarna	Andel av publ.	Andel av citeringarna
25+	2,2	24,1	11,2	53,1	1,20	29,5
Topp 5 %	5,1*	35,2	5,1*	28,9	5,1 ^A	38,0

*Anledningen till att dessa andelar är större än 5 % är att det finns flera publikationer som citeras lika mycket som den 95e percentilen.

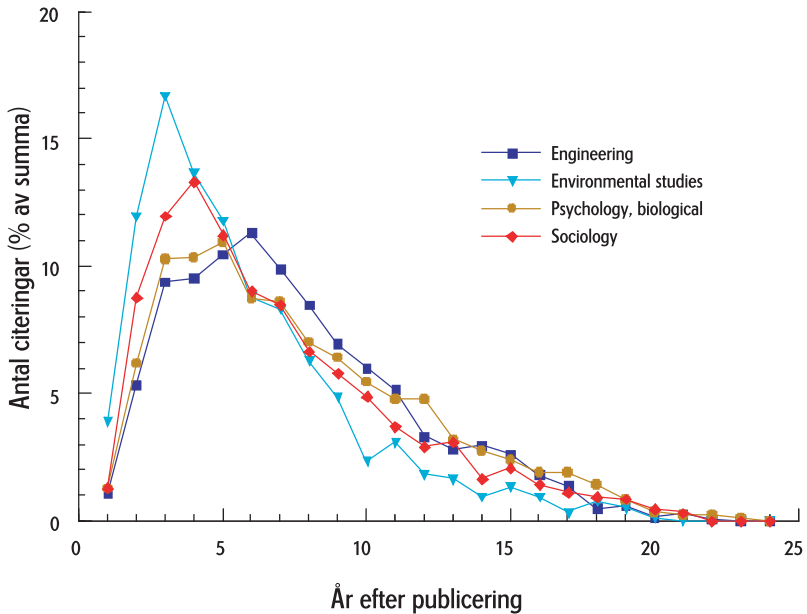
I fortsättningen är den redovisade statistiken baserad enbart på artiklar och översiktsartiklar.

B) Citeringarnas spridning över tiden

En annan aspekt av citeringsmönstret är tidsdynamiken i när publikationerna citeras. Antalet citeringar per år ökar under de första 3 till 5 åren efter publiceringen och minskar därefter (figur 3). Det tar olika lång tid inom olika fält innan en publikation citeras som mest. För de fält som exemplifieras i figur 2 tar det längst tid för publikationer inom teknikvetenskap att bli citerade medan miljöforskning citeras efter kortare tid.

¹⁶ Se Aksnes & Sivertsen 2004.

Figur 3. Tidsdynamiken i när publikationer från några olika fält blir citerade.

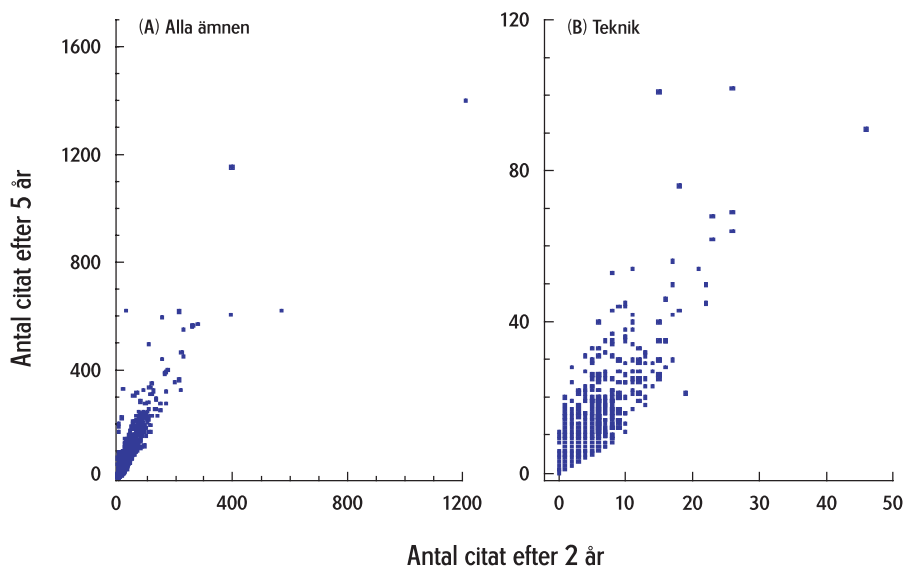


För att kunna analysera utvecklingen så långt fram i tiden som möjligt har vi här valt att använda oss av ett tvåårigt citeringsfönster, dvs. antalet citeringar två år efter publiceringsåret eller i genomsnitt 2,5 år efter publiceringen. Trots att tidsfönstret är relativt kort får man en bra indikation på antalet citeringar även efter längre tidsperioder; högt citerade arbeten upptäcks och citeras nästan alltid tidigt.¹⁷ Andelen publikationer som citeringsmässigt upptäcks sent ("sleeping beauties" enligt van Raan 2004a) är alltså mycket låg. Som ett exempel kan nämnas att av 3,9 miljoner artiklar ("articles") i databasen som publicerade mellan 1982 och 1989 hade 261 citerats lågt (5 gånger eller mindre) under de första 5 åren men senare blivit högt citerade (totalt minst 100 citeringar tom. 2004), dvs. mindre än en av 10 000 artiklar är en "sleeping beauty" med denna avgränsning.

Samvariationen mellan antalet citeringar efter två respektive fem år för samtliga svenska publikationer under de senaste åren illustreras i figur 4A. Som exempel på ett område där merparten av alla citeringar kommer relativt långt efter publiceringen, jämförs ett tvåårigt och ett femårigt citeringsfönster för teknikområdet i figur 4B. Även här ger ett tvåårigt fönster en relativt god uppskattning av vilka publikationer som blir högt och lågt citerade.

¹⁷ Glänzel & Garfield 2004, Adams 2005

Figur 4. Jämförelse mellan antalet citat efter 2 respektive 5 år för svenska artiklar publicerade 1998 och 1999. Korrelationskoefficient Alla ämnen; Spearman rankkorrelation $r = 0,90$ ($P < 0,001$, $N = 30960$), teknik $r = 0,82$, ($P < 0,001$, $N = 3771$). Figuren är baserad på samtliga citeringar, även självcitat. Varje punkt representerar i princip en publikation, men flera punkter kan ligga "ovanpå" varandra.

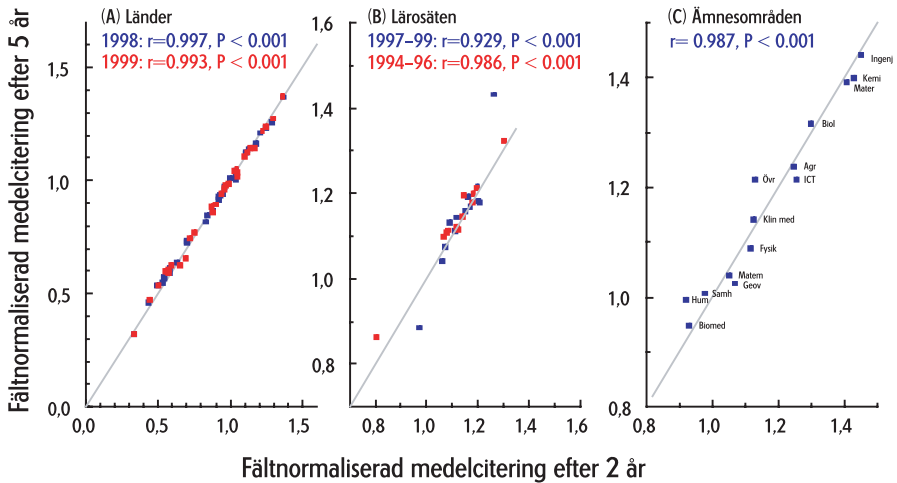


I figur 4 finns enstaka punkter (publikationer) som ligger ovanför punktsvärmen. Dessa är fall där antalet citeringar efter fem år är större än förväntat enligt antalet efter två år. För enskilda publikationer kan det tvååriga citeringsfönstret alltså ge en dålig indikation för antalet citeringar efter en längre period. Men för större material ger det tvååriga fönstret en relativt bra indikation på hur publikationer kommer att citeras även efter en längre tidsperiod. I denna rapport har vi därför satt den nedre gränsen till 100 publikationer för att presentera citeringsstatistik för en viss kombination av ämne och/eller lärosäte.

Även när man jämför fältnormaliserade medelvärden¹⁸ baserade på två eller fem år är överensstämmelsen mellan det två- och det femåriga citeringsfönstret mycket god (figur 5). För enskilda år är samstämmigheten naturligtvis sämre (rankkorrelationskoefficienten varierade mellan 0,81–0,99 för sex olika år). Jämförs den fältnormaliserade medelciteringen för svenska publikationer uppdelade på 14 ämnesområden (figur 5C) visar sig även här rangordningen vara robust. I figur 5B och 5C är det lärosäten eller ämnesområden med låga publiceringsvolymerna som varierar mest mellan år och avviker mest från 1:1-linjen.

¹⁸ Se metodbilaga för en förklaring

Figur 5. Jämförelse av fältnormaliserad medelcitering två och fem år efter publiceringen. I delfigur A jämförs årsmedelvärden för 32 länder, i B 3-årsmedelvärden för svenska lärosäten (med en volym om minst 100 publikationer under 3 år) och i C jämförs svenska publikationer uppdelade på vetenskapsområden. ICT står för "Information and communication technology". För (A) och (B) ges två exempel (markerade med blå respektive röda symboler). Rankkorrelationskoefficienten (r) för respektive jämförelse anges också. De grå linjerna indikerar ett 1:1 förhållande mellan de två måtten. Figuren är baserad på samtliga citeringar, även självcitater.

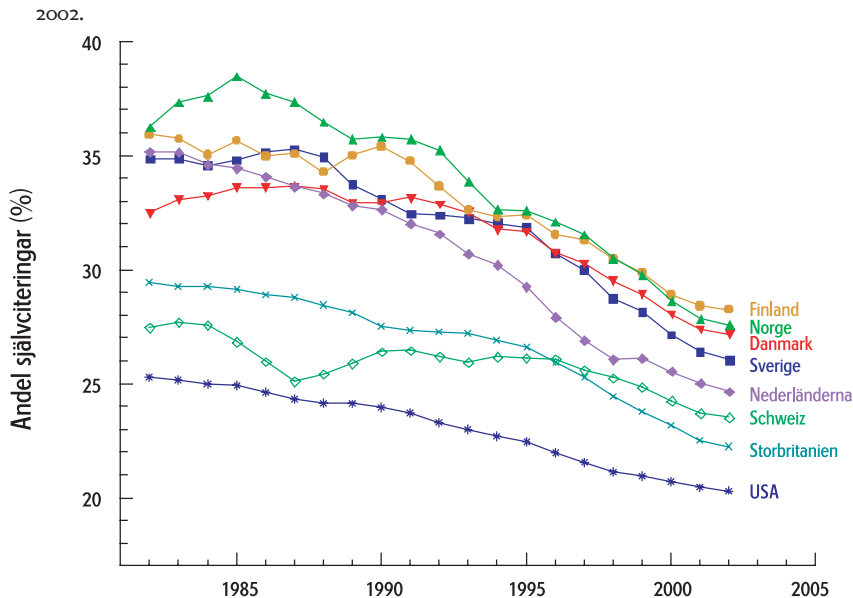


C) Självcitering

De citeringsdata som presenterats hittills har baserats på samtliga citeringar. Eftersom analyser av antalet citeringar avser att uppskatta den uppmärksamhet en artikel erhåller från omvärlden får man ett bättre mått om självciteringar tas bort. Det finns flera sätt att definiera självciteringar. Den vanligaste metoden är att inte räkna citeringar där det finns ett överlapp bland författarnamnen mellan den citerade och den citerande publikationen. All borttagning av självcitater i denna rapport är gjord på detta sätt. Vissa forskningsfält där många författare normalt samarbetar kan missgynnas av elimineringen av självcitater.

Förändringen i andelen av alla citeringar som är självcitater visas för några länder i figur 6. För Sveriges del har andelen minskat från 35 % under första halvan av 1980-talet till 27 % under början av 2000-talet.

Figur 6. Andel självcitering för svenska och några andra länders publikationer mellan 1982 och



D) Beräkning av antalet publikationer och citeringar - fraktionisering

I denna rapport är alla publikationer fraktioniserade med avseende på antalet författaradresser som en publikation har. Det innebär att när mer än en adress är angiven på en publikation delas publikationen mellan de angivna adresserna. Om det t.ex., finns en tysk och två svenska adresser, krediteras Tyskland med $1/3$ och Sverige med $2/3$ publikation. Samma förfarande används för beräkningen av lärosätenas publikationer. Om de två svenska adresserna i exemplet finns vid olika lärosäten krediteras respektive lärosäte $1/3$ publikation. På samma sätt delas publikationer mellan ämnesområden då en tidskrift tillhör flera områden.

Eftersom den vetenskapliga publiceringen i ökande grad är resultatet av samarbete skulle det ske en betydande dubbelräkning om alla lärosäten eller länder fullt ut tillgodoräknades alla publikationer där någon av författarna tillhör respektive lärosäte eller land.

Högt citerade publikationer är relativt ofta resultatet av internationellt samarbete med många författare.¹⁹ En effekt av fraktioniseringen är att sådana publikationer väger mindre vid medelvärdesberäkningen för enskilda lärosäten eller länder jämfört med publikationer med få författare. Effekten av detta kan variera mellan ämnesområden: områden där publi-

¹⁹ Se t.ex. Moed 2005 kap 23.

kationerna ofta är resultatet av mer omfattande samarbete påverkas mera än områden där antalet författare normalt är lägre. Till exempel påverkas medelvärdet för antalet citeringar för Lunds och Göteborgs universitet och Karolinska institutets mera av fraktioneringen än medelvärdet för de tekniska högskolorna (Kungl. Tekniska högskolan, Chalmers och Luleå tekniska universitet). Högt citerade publikationer från Lund, Göteborg och Karolinska institutet är alltså oftare resultatet av samarbete med forskare utanför dessa lärosäten än vad som är fallet för de tekniska högskolorna.

Fördelen med fraktioneringen är att en publikation bara räknas "en gång" oavsett antalet författaradresser (summan av alla fraktioner för en publikation är alltid 1). Eftersom högt citerade publikationer ofta har många författare skulle betydelsen av dessa överdrivas om inte publikationerna fraktioniserades. Utan fraktionisering skulle till exempel medelvärdet för alla lärosäten i ett land kunna ligga klart över landets medelvärde. På samma sätt skulle, utan fraktionisering, ett (vägt) medelvärde för alla länder ligga klart över medelvärdet för alla världens publikationer.

E) Metoder för normalisering av citeringarna

Ett problem med jämförelser mellan antalet citeringar inom olika forskningsfält är att citeringstraditionerna varierar relativt mycket; vad som är hög citering i ett fält kan vara låg i ett annat. Vidare är medelciteringen två år efter publiceringen högre nu än den var för 10 år sedan. Ett bra citeringsindex ska inte påverkas av denna inflation i citeringsnivån och den ska möjliggöra "rättvisa" jämförelser mellan olika forskningsfält. Den bibliometriska litteraturen innehåller många olika sätt att normalisera citeringarna för att öka jämförbarheten. Här använder vi framför allt en global fältnormalisering. I avsnitt 3.1 jämför vi denna metod med en Europainormalisering.

- 1 Fältnormaliseringen baseras på att publikationerna delas upp i 252 forskningsfält av Thomson Scientific. För varje kombination av fält, år och typ av publikation (här enbart "articles och "reviews") beräknar vi ett globalt fältmedelvärde (field citation score, FCS). Antalet citat (citations per paper CPP) divideras med fältmedelvärdet för att få det fältnormaliserade medelvärdet: CPP/FCS. Både CPP och FCS beräknas för en tvåårsperiod efter publiceringen. En snarlikt beräknad kvot kallas "the crown indicator" av dess upphovsmän eftersom den anses överlägsen andra mått²⁰ (se mer detaljer i metodikbilagan).

²⁰van Raan 2004b.

- 2 En annan metod är att använda en begränsad grupp länders medelcitering som referens, t.ex. ett urval EU-länder. Här använder vi EU₁₅²¹ gruppen plus Norge och Schweiz som referensgrupp (kallad "EU₁₇"). Till skillnad från fältnormaliseringen, tar denna typ av normalisering tar inte hänsyn till skillnader mellan olika forskningsfält vilket innebär att områden som normalt har höga citeringsnivåer, t.ex. medicin, väger tyngre än områden med låg medelcitering.

I kapitel 3 gör vi en jämförelse mellan dessa två sätt att normalisera i samband med att Sverige jämförs med några europeiska länder. All citeringsstatistik som redovisas nedan (avsnitt 3 och 4) är baserad på artiklar och översiktsartiklar ("articles" och "reviews"). Andra typer av publikationer är inte medtagna i beräkningarna.

²¹ Belgien, Danmark, Finland, Frankrike, Grekland, Irland, Italien, Luxemburg, Nederländerna, Portugal, Spanien, Storbritannien, Sverige, Tyskland, Österrike.

3 INTERNATIONELL ÖVERSIKT

Den internationella översikten består av tre delar. I avsnitt 3.1 ges en översikt av variationer i publiceringsvolym och citering över tid (1982 till 2003) för Sverige och ett urval andra europeiska länder samt i något fall även USA. Citeringsjämförelsen görs med två olika normaliseringsmetoder och med två sätt att räkna citeringar, med och utan självciteringar. I avsnitt 3.2 jämförs 33 länders medelvärden för den senaste treårsperioden varvid vi har ett komplett tvåårigt citeringsfönster i databasen (dvs. åren 2000-2002). Vi avslutar i avsnitt 3.3 med att redovisa en mer detaljerad uppdelning av statistiken för att kunna diskutera skillnader i utvecklingen mellan Sverige och andra länder.

3.1 Utveckling mellan 1982 och 2003

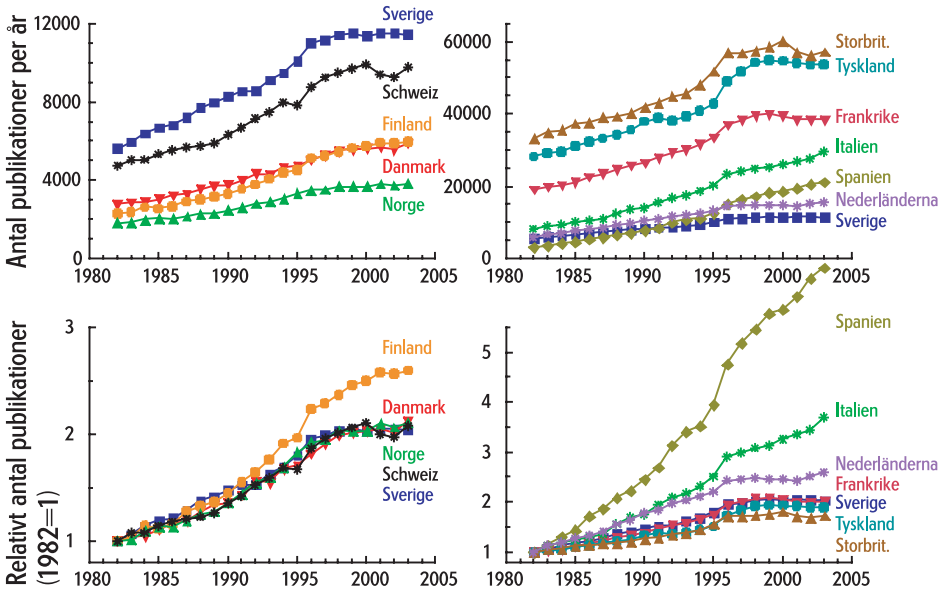
A) Antal publikationer

Figur 7 visar ökningen i antalet publikationer från några europeiska länder sedan 1982. Denna ökning över tiden är en kombinerad effekt av att publiceringen ökat och att databasen successivt täckt in fler tidskrifter. De flesta länder som redovisas i figur 7 har haft en snarlik utveckling av antalet publikationer sedan tidigt 1980-tal. Utvecklingen i Danmark, Frankrike, Tyskland, Norge och Schweiz är mycket lik den svenska. Finland och Nederländerna har haft en något kraftigare volymökning än Sverige, medan Italiens och framför allt Spaniens publicering har ökat betydligt mer än Sveriges.

Flera länder hade relativt stora ökningar i antalet publikationer mellan 1995 och 1997, detta beror troligen till största delen på att databasen täckning förbättrades (fler tidskrifter togs med) under mitten av 1990-talet.

Figur 7. Utvecklingen av antalet svenska publikationer i jämförelse med några europeiska länder.

De två övre graferna visar antalet publikationer per år och de nedre den relativa utvecklingen (volymen 1982=1). Publikationer med författare från flera länder är fraktioniserade på land (se metodikbilagan).



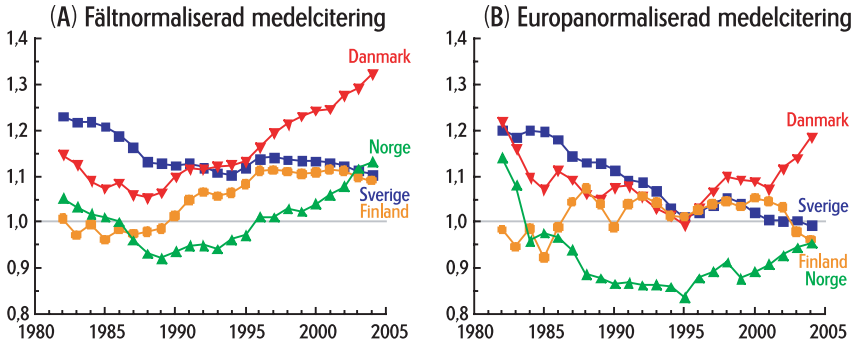
B) Förändringar i citeringsstatistiken

Citeringsnivåer med självciteringar

Både den fältnormaliserade medelciteringen (figur 8A) och den Europainormaliserade medelciteringen (figur 8B) visar en nedåtgående trend för Sverige de senaste 25 åren. Tillbakagången för Sverige är emellertid kraftigare om normaliseringen görs med hjälp av Europagenomsnittet än när globala medelvärden för olika forskningsfält används. Enligt den fältnormaliserade medelciteringen backade Sverige fram till mitten 1980-talet. Sedan 1987 har förändringarna varit relativt små och medelvärdet har legat drygt 10 % över världsgenomsnittet.

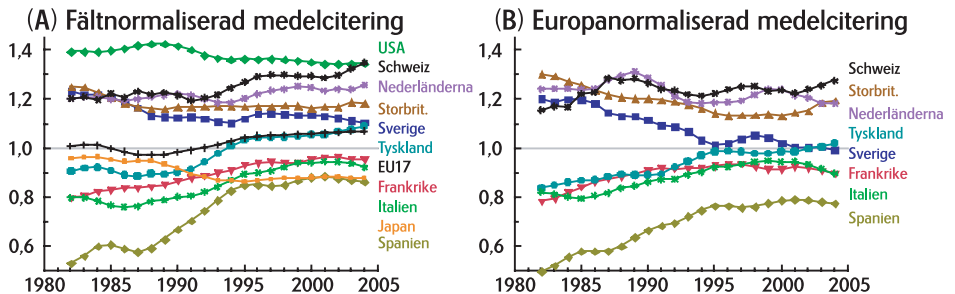
Både Norge och Danmark har haft en bättre utveckling än Sverige. Eftersom Danmarks uppgång började från en högre nivå ligger Danmark betydligt över Sverige i medelcitering (28 % över världsgenomsnittet att jämföra med Sveriges 11 % för åren 2001-2002). Finska och norska publikationer låg under slutet av 1980-talet under världsgenomsnittet men har de senaste åren citerats i ungefär samma grad som svenska publikationer.

Figur 8. Medelcitering av svenska artiklar jämfört med de nordiska gränsländerna. (A) visar globalt fältnormaliserade medelvärden och (B) medelciteringen i förhållande till medelvärdet för 17 europeiska länder. Alla kurvor visar glidande 3-årsmedelvärden. Värdena för 2003 och 2004 är preliminära eftersom publikationer från dessa år ännu inte hunnit citeras under två år. Självciteringar är inte borttagna.



Figur 9. Citering av svenska artiklar jämfört med några europeiska länder samt USA och Japan.

Figur (A) visar globalt fältnormaliserade medelvärden och (B) medelcitering i förhållande till medelvärdet för 17 europeiska länder²² (USA och Japan finns inte med i denna delfigur). Alla kurvor är visar glidande 3-årsmedelvärden. Värdena för 2003 och 2004 är preliminära eftersom publikationer från dessa år ännu inte hunnit citeras under två år. Självciteringar är inte borttagna.



Om Sverige jämförs med de största länderna i Europa när det gäller publiceringsvolym (figur 7) och citering (figur 9) får man en förklaring till skillnaderna mellan fält- och Europeanormaliseringen i figur 8. Framför allt

²² Belgien, Danmark, Finland, Frankrike, Grekland, Irland, Italien, Luxemburg, Nederländerna, Norge, Portugal, Schweiz, Spanien, Storbritannien, Sverige, Tyskland, Österrike.

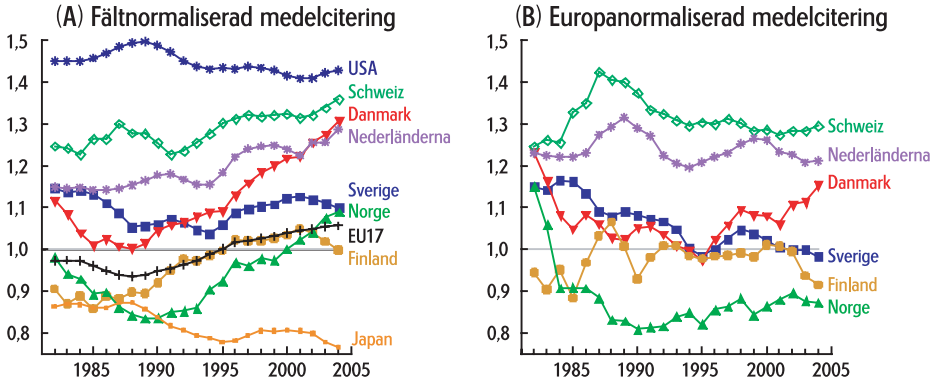
Tyskland, Frankrike och Spanien, men också andra länder som inte visas i figuren, har ökat sin publiceringsvolym i tidskrifter som indexeras i databasen och dessutom har dessa publikationer blivit alltmer citerade. Nedgången för svensk del under den senaste 15-årsperioden enligt figur 8B beror alltså på att viktiga europeiska länder förbättrat sina prestationer snarare än på att svensk forskning sjunkit i kvalitet med globala mått mätt. Att Sverige inte hållit jämn takt med grannländerna när det gäller att förbättra kvalitén kan inte förklaras med att Sverige publicerar för mycket. De flesta av våra grannländer har haft en liknande utveckling eller större relativ volymökning än Sverige.

Framför allt den starka utvecklingen hos stora länder som Tyskland och Frankrike har resulterat i att Europa som region har stärkt sin position från att ha legat i närheten av världsgenomsnittet under 1980-talet till 6 % över världsgenomsnittet de senaste åren (2003–2004, figur 9). Europa ligger dock klart efter USA som i stort sett lyckats vidmakthålla ett mycket högt medelvärde. USA har backat något, från att ha legat ca 40 % över världsgenomsnittet under 1980-talet till 34 % över världsgenomsnittet under de första åren av 2000-talet. Förändringarna för flera av de europeiska länderna kan till viss del bero på att Thomson Scientific i ökande grad inkluderat europeiska vetenskapliga tidskrifter men också att dessa länder i en ökande grad publicerar i internationella tidskrifter.

Citeringsnivåer utan självciteringar

Hittills har vi i detta avsnitt redovisat statistik baserad på alla citeringar. Om självciteringarna tas bort får vi en mer bättre uppskattning av hur omvärlden uppfattar publikationerna och bilden ändras något (figur 10). När statistiken baseras på alla citeringar (figur 8A) får Sverige i början av 1980-talet en mycket hög medelcitering (25 % över världsgenomsnittet 1982–84) och en andra plats efter USA. Det visar sig att en relativt stor andel av dessa citeringar var självciteringar. Utan dessa sjunker medelciteringen till 14 % över världsgenomsnittet och vi får en sjätte plats i en världsranking. I början av 2000-talet är det svenska medelvärdet 13 % över världsgenomsnittet, dvs, bara marginellt lägre än under början av 1980-talet och Sverige placerar sig fortfarande på sjätte plats i en världsranking (se figur 12). Danmark och Nederländerna har utvecklats betydligt bättre än Sverige de senaste 15 åren även när självciteringarna är borttagna.

Figur 10. (A) Fältnormaliserad och (B) Europanormaliserad medelcitering där självciteringarna är borttagna. (USA och Japan finns inte med i delfigur B). Alla kurvor är baserade på glidande 3-årsmedelvärden. Värdena för 2003 och 2004 är preliminära eftersom publikationer från dessa år ännu inte hunnit citeras under två år. Motsvarande data med alla citeringar (även självcitater) finns i figur 8 och 9.



I det följande kommer all citeringsstatistik att vara baserad på fältnormaliserad medelcitering med självciteringarna borttagna.

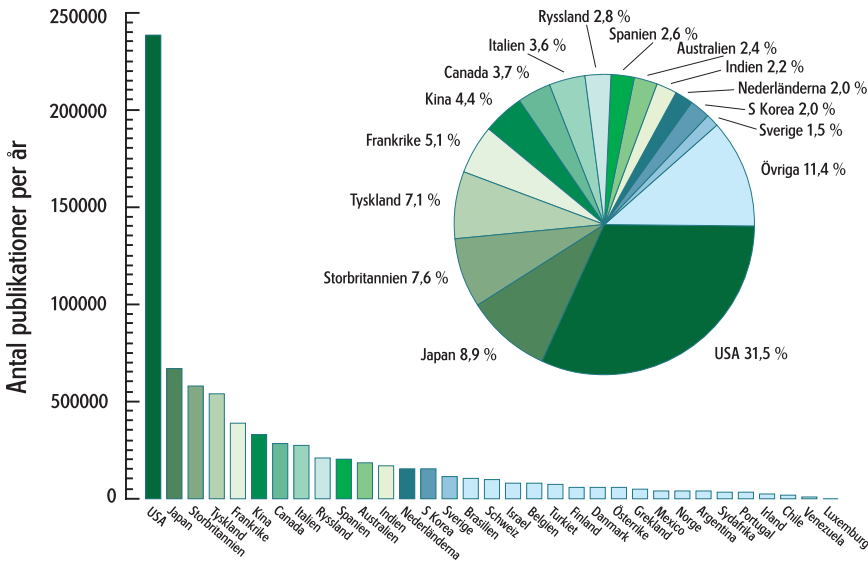
3.2 Länderjämförelse baserad på medelvärden för 2000-2002

A) Volym

Beträffande antalet publikationer (av "articles" och "reviews") dominerar USA kraftigt med nästan en tredjedel av världsproduktionen (figur 11)²³. Japan, Storbritannien och Tyskland följer efter USA med 7–9 % vardera medan Sverige bidrar med 1,5 % till världsproduktionen. De 33 länder som finns med i figur 11 står för 99 % av världsproduktionen.

²³En vanlig invändning mot citeringsanalys baserad på Thomson Scientific's databas är att eftersom företaget är USA-baserat är amerikanska tidskrifter representerade i oproportionerligt stor utsträckning. Det är därför troligt att databasen överdriver USA:s dominans något.

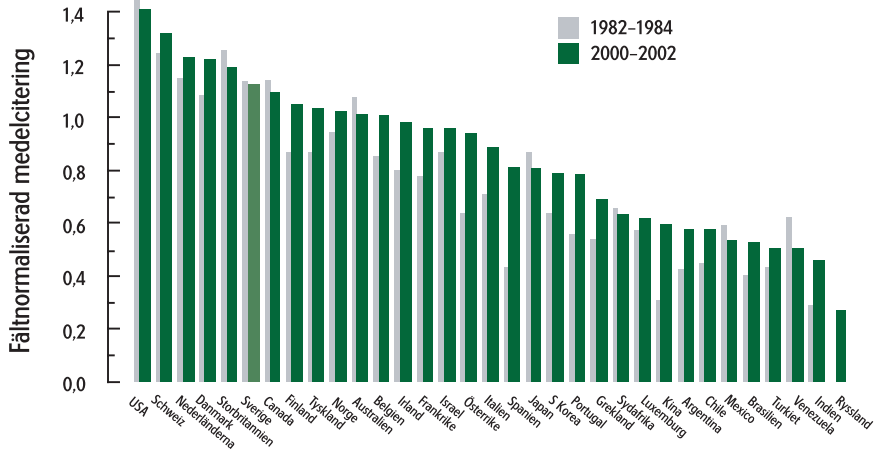
Figur 11. Antal publikationer per år från 33 länder (medelvärden för 2000–2003 av "articles" och "reviews"). Publikationerna är fraktioniserade.



B) Citering

Hur mycket citeras publikationerna från dessa länder? Medelvärdet för den fältnormaliserade medelciteringen (själv citeringar borttagna) för de tre senaste åren (2000–2002) redovisas i figur 12. Även när det gäller citering är USA klart världsledande, följt av Schweiz, Nederländerna och Danmark. I denna jämförelse hamnar Sverige på sjätte plats, följt av Kanada och Finland. För flertalet länder i figur 12 har citeringarna av deras publikationer ökat sedan början av 1980-talet. Detta är emellertid inte fallet för Sverige där citeringen minskat något. Exempel på andra länder som visat samma neråtgående trend är USA, Storbritannien, Kanada och Japan.

Figur 12. Fältnormaliserad medelcitering för 33 länder två år efter publiceringen. Gröna staplar visar medelvärdet för åren 2000-2002 och grå staplar för åren 1982-1984. Figuren är baserad på publikationer "articles och reviews" där självciteringar är borttagna (se metodikbilagan för mer detaljer).



3.3 Varför har inte Sverige utvecklats lika bra som övriga Europa?

I den forskningspolitiska debatten har det ofta framhållits att Sveriges konkurrenskraft har minskat mätt i form av citering av vetenskapliga publikationer. Även om Sverige i stort sett bibehållit citeringsnivån sedan 1980-talet när man ser på den fältnormaliserade medelciteringen kvarstår det faktum att jämfört med övriga Europa har Sverige haft en negativ utveckling under hela den studerade perioden. Flera europeiska länder, som t.ex. Danmark, Nederländerna och Tyskland har utvecklats betydligt bättre.

Något direkt orsakssamband som förklarar skillnaderna i utvecklingen av citeringsmönstret mellan Sverige och andra länder är svårt att få fram ur publikationsdatabasen. Men genom att granska den bibliometriska statistiken närmare går det att få några indikationer. En möjlig förklaring, att Sverige skulle ha ökat volymen på bekostnad av kvalitén, har vi redan kunnat avfärda (se avsnitt 3.1 ovan). Nedan redovisar vi några andra faktorer som kan ge indikationer på möjliga orsaker till Sveriges relativa tillbakagång:

- A) Publiceringsprofil: volym och citering inom olika områden
 B) Andel publikationer som inte citeras under de första åren
 C) Andelen högt citerade arbeten

A) Publiceringsprofil: volym och citering inom olika vetenskapsområden

En orsak till att länder har utvecklats på olika sätt kan vara att de har olika ämnesmässig inriktning. Om Sveriges publicering inom olika områden jämförs med våra nordiska grannländer och Nederländerna ser mönstren relativt lika ut (tabell 3). Medicin dominerar med nära 50 % av alla publikationer överallt utom i Norge där andelen är 43 %. Norge har en något högre andel inom skogs- och jordbruksvetenskap och samhällsvetenskap än Sverige. Finland, följt av Sverige, har högst andel teknik (12,7 % respektive 11,7 %)

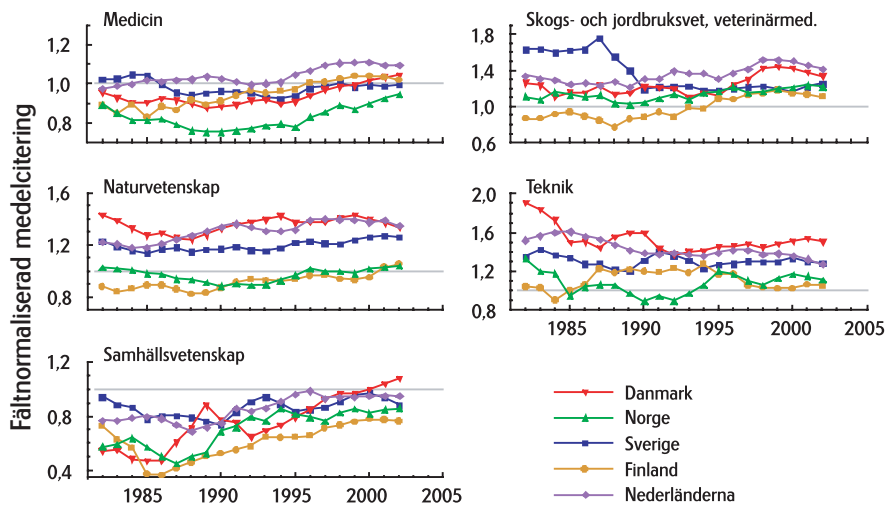
Inte heller om man ser på hur områdenas relativa storlek har förändrats sedan början av 1980-talet finns några dramatiska skillnader mellan dessa länder. Samtliga har ökat andelen publikationer inom teknik (med 3-5 %). I Sverige och Nederländerna har andelen naturvetenskapliga publikationer minskat, medan den har ökat något i Danmark, Finland och Norge. I de senare länderna har andelen medicin minskat (tabell 3).

Tabell 3. De nordiska länderna samt Nederländernas ämnesprofiler. Siffrorna visar den procentuella fördelningen av alla fraktioniserade publikationer uppdelade på vetenskapsområden (medelvärden för åren 2000–2002 och 1982–1984). SJV står för skog- och jordbruksvetenskap samt veterinärmedicin (ämneskategorin ”agronomy” i databasen).

<u>Vetenskapsområde</u>	<u>Sverige</u>	<u>Danmark</u>	<u>Finland</u>	<u>Norge</u>	<u>Nederländerna</u>
2000-2002					
Humaniora	1,1	1,7	1,5	2,2	2,1
Medicin	50,5	49,0	48,4	42,7	47,7
Naturvetenskap	25,3	26,0	23,6	26,4	24,2
Samhällsvetenskap	4,1	3,9	4,2	7,6	7,4
SJV	6,4	10,5	9,2	10,5	6,7
Teknik	11,7	8,2	12,7	9,9	11,2
Övriga	0,8	0,7	0,4	0,7	0,7
1982-1984					
Humaniora	2,1	1,6	1,6	2,2	2,4
Medicin	49,6	63,1	59,8	53,6	44,8
Naturvetenskap	30,8	22,1	20,2	23,4	29,9
Samhällsvetenskap	4,1	2,9	3,6	7,0	5,9
SJV	4,7	5,2	6,6	6,6	8,5
Teknik	6,6	4,4	7,9	6,5	7,7
Övriga	2,1	0,6	0,4	0,6	0,9

Om citeringsstatistiken delas upp på samma sätt, dvs. på vetenskapsområden (figur 13), visar framför allt svensk publicering inom medicin ungefär samma utveckling som Sverige-medelvärdet i figur 10, med en relativt jämn citeringsnivå under hela den studerade perioden. Eftersom en stor del av all publicering sker inom medicin (se tabell 3) påverkar detta område landsmedelvärdet relativt mycket. Till skillnad från Sverige har Danmark haft en positiv utveckling de senaste 10 åren i de flesta ämnesområden som visas i figur 13 och Danmark citeras mer än Sverige inom samtliga områden.

Figur 13. Fältnormaliserad citering för Norden och Nederländerna uppdelad på vetenskapsområde. Själv citeringar är borttagna.



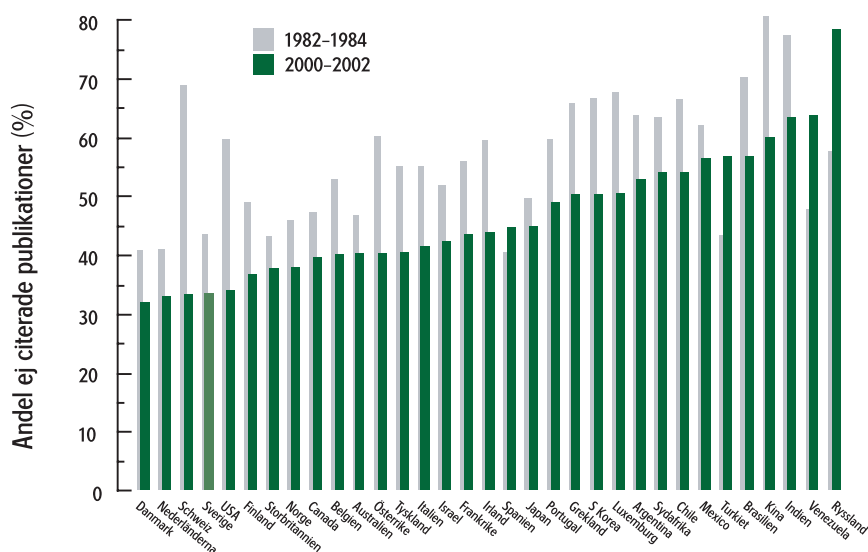
Skillnaden i citering mellan publikationer från Sverige och t.ex. Danmark eller Nederländerna kan alltså inte förklaras med ämnesmässiga olikheter mellan länderna; proportionerna mellan publiceringen inom vetenskapsområdena är relativt lika. Skillnaderna är inte heller beroende av ett enstaka område; publikationerna från Danmark citeras relativt mycket i både medicin, teknik och naturvetenskap, medan det höga holländska medelvärdet baseras på höga citeringsnivåer inom medicin, naturvetenskap och skogs- och jordbruksrelaterade ämnen.

B) Andel ej citerade publikationer

Andelen av alla publikationer som inte citerats under de två första åren varierar för de flesta länderna mellan 30 % och 60 %. Endast Ryssland av-

viker med 78 % ej citerade. Andelen ej citerade har minskat sedan början av 1980-talet för de flesta (29 av 33) av de länder som redovisas i figur 14. Danmark, följt av Nederländerna, Schweiz och Sverige, ligger bäst till när det gäller andelen publikationer som inte citeras. Det finns inte några stora skillnader mellan Sverige och dessa länder som förklarar skillnaderna i citeringens utveckling över tiden (enligt figur 8, 9 eller 10). För Schweiz har dock den positiva utvecklingen stärkts av att andelen ej citerade publikationer minskat kraftigt (från 69 % till 33 %).

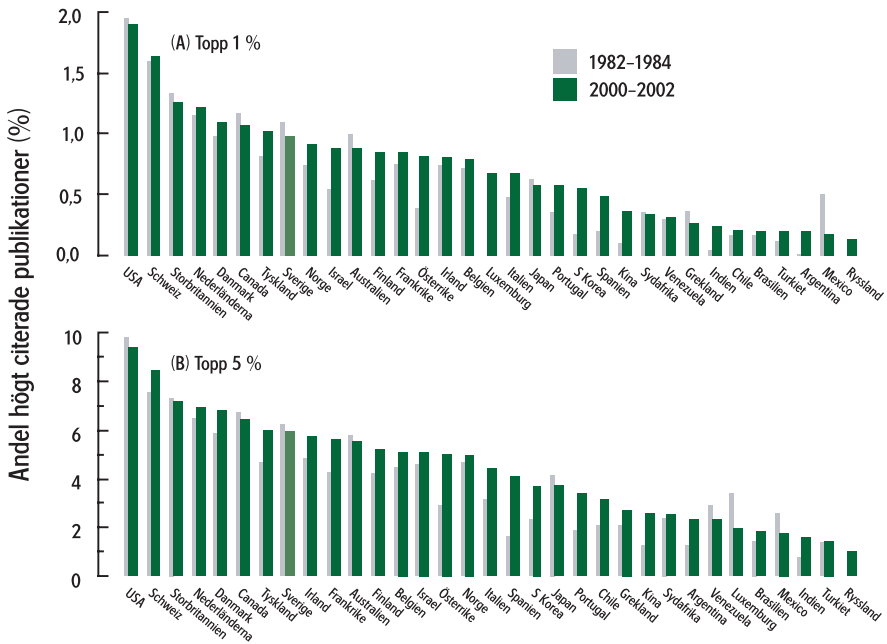
Figur 14. Andel ej citerade publikationer två år efter publiceringen. Medelvärden för två perioder 1982–1984 och 2000–2002. Själv citeringar är borttagna.



C) Andel högt citerade publikationer

USA, följt av Schweiz, Storbritannien och Nederländerna, är världsledande när det gäller andelen högt citerade publikationer (figur 15). En orsak till skillnaderna mellan Danmarks, Nederländerna och Schweiz utvecklats bättre än Sverige finns i andelen högt citerade publikationer. Dessa tre länder har ökat sin andel högt citerade publikationer sedan början av 1980-talet (med +0,5 till +0,9 % för topp 5 %-publikationerna) medan Sveriges andel har minskat från 6,3 % till 5,9 % (dvs -0,4 %).

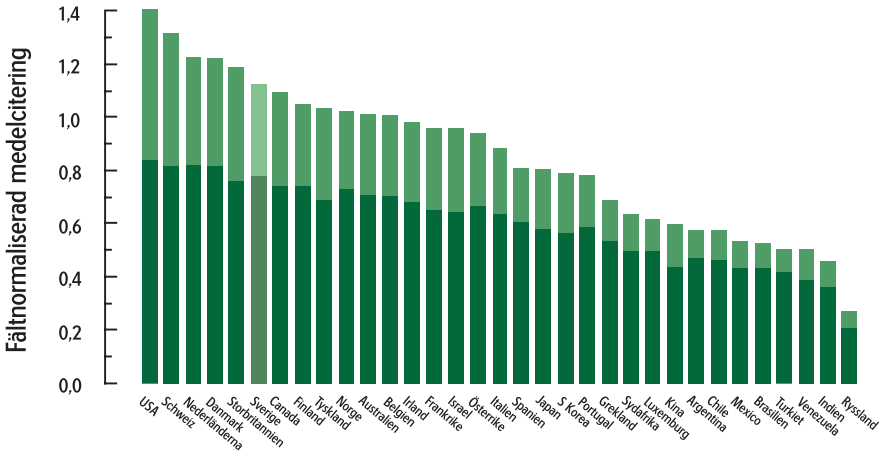
Figur 15. Andel högt citerade publikationer enligt två olika definitioner på högt citerade publikationer. (A) Andelen bland den högst citerade procenten (citerade som den 99:e percentilen eller högre globalt sett). (B) Andelen bland de 5 % högst citerade (citerade som den 95:e percentilen eller högre). Alla citeringar är fältnormaliserade och självciteringar är borttagna.



Beräknar man topp 5 %-publikationernas bidrag till landsmedelvärdet (figur 16) finner man att för länder med höga medelvärden bidrar de bästa publikationerna med en stor andel av alla citeringar. För USA står topp 5 %-publikationerna för 40 % av de fältnormaliserade citeringarna. Andelen av citaten som kommer från högt citerade publikationer minskar sedan successivt och är ca 20 % eller lägre för länder med lägst medelvärde. Bidraget av topp 5 %-publikationerna för Sverige är 31 % jämfört med t.ex. Schweiz (38 %) och Storbritannien (36 %). Även Danmark och Nederländerna ligger något högre än Sverige (33 % för båda).

Även om topp 5 %-publikationerna har en stor betydelse för landsmedelvärdet för många av länderna i figur 16, så blir rangordningen mellan länder i stort sett den samma även utan dessa högt citerade publikationer (den nedre mörkare delen av staplarna i figur 16). Skillnaderna mellan länderna med de högsta medelvärdena blir dock betydligt mindre.

Figur 16. Fältnormaliserad medelcitering för ett antal länder baserat på publikationer (articles och reviews) mellan 2000 och 2002. Den övre ljusare delen av staplarna visar de högst citerade (topp 5 %) publikationernas bidrag till landsmedelvärdet. Självciteringar är borttagna.



Sammanfattningsvis visar det sig alltså att Sverige har en låg andel publikationer som inte citeras alls under de två första åren efter publiceringen. Däremot har andelen högt citerade publikationer minskat något. Eftersom de högt citerade publikationerna påverkar landsmedelvärdet relativt mycket finns här en del av förklaringen till att Sverige inte hållit jämnt takt med t.ex. Danmark, Nederländerna och Schweiz.

4 VARIATIONER MELLAN ÄMNES- OMRÅDEN OCH LÄROSÄTEN I SVERIGE

I detta kapitel presenterar vi statistik för svenska publikationer och jämför ämnesområden och/eller lärosäten. Först delar vi upp de svenska publikationerna på vetenskapsområde enligt tidskrifternas ämnestillhörighet respektive enligt författarnas adresser (avsnitt 4.1). Därefter redovisar vi hur de svenska publikationerna fördelar sig bland 14 ämnesområden och hur dessa citeras (avsnitt 4.2). I det sista avsnittet (4.3) redovisar vi publiceringsvolym och citering uppdelat på lärosäten.

4.1 Vem publicerar var?

Författarnas "fakultetstillhörighet" jämfört med tidskrifternas ämnestillhörighet

Thomson Scientific kategoriserar alla tidskrifter inom ett eller flera ämnesområden. För att dessutom kunna sortera publikationerna efter författarnas adress har vi klassat alla svenska universitets- och högskoleadresser till en fakultet eller vetenskapsområde (humaniora och samhällsvetenskap redovisas var för sig). I detta avsnitt jämför vi publiceringens volym och citering när materialet delas upp dels efter tidskrifternas ämnestillhörighet, dels efter författarnas adress.

Publiceringens volym inom olika ämnesområden

Nästan 51 % av alla svenska publikationer 2000–2002 fanns i medicinska tidskrifter. Det näst största området, naturvetenskap, stod för 26 % av publikationerna medan teknik kom på tredje plats med knappt 11 %.

Delar man i stället upp statistiken på fakulteter enligt författarnas adresser var den medicinska dominansen något mindre; 44 % av publikationerna var publicerade av forskare vid medicinsk fakultet, 23 % kom från naturvetenskaplig fakultet och 20 % från teknisk (tabell 4).

Forskare från alla fakulteter utom teknisk publicerar framför allt i tidskrifter som tillhör fakultetens "egna ämnesområden" (de grå fälten i tabell

4). Medicinska forskare var mest trogna det egna området publiceringsmässigt; 91 % av publikationerna från medicinsk fakultet fanns i medicinska tidskrifter. Forskare vid tekniska fakulteter publicerade i högre utsträckning i naturvetenskapliga tidskrifter än i tekniska. Även forskare vid Sveriges lantbruksuniversitet (som alla klassas till "fakulteten" SLU i tabell 4) publicerar i stor utsträckning i tidskrifter utanför lantbruksuniversitetens klassiska ämnen (53 % av SLU-publikationerna var i skogs- och jordbruksrelaterade tidskrifter²⁴).

Tabell 4. Sveriges publikationer uppdelade på tidskrifternas ämnestillhörighet och författarnas fakultet (vetenskapsområdestillhörighet). SJV står för skog- och jordbruksvetenskap samt veterinärmedicin (ämneskategorin "agronomi" i databasen). Alla siffror är procent av Sveriges produktion av fraktioniserade publikationer. Figuren är baserad på svenska publikationer från 2000-2002. Tvärvetenskapliga områden är framför allt institutioner gemensamma för tekniskt och naturvetenskapligt område i Lund och medicinskt-naturvetenskapligt i Uppsala.

Författarnas adress	Tidskriftens ämnestillhörighet							Summa
	Humaniora	Medicin	Naturvet.	Samhällsv.	SJV	Teknik	Övriga	
Humanistisk	0,80	0,03	0,02	0,07	0,01	0,02	0,00	1,0
Medicinsk	0,05	40,56	1,47	0,73	0,64	0,54	0,44	44,4
Naturvet.	0,02	5,86	13,22	0,18	1,93	1,66	0,27	23,2
Samhällsvet.	0,10	1,05	0,23	2,83	0,08	0,28	0,01	4,6
SLU	0,01	1,00	1,12	0,07	2,74	0,18	0,03	5,2
Teknisk	0,07	1,88	9,35	0,31	0,56	7,83	0,08	20,1
Tvärvet.	0,00	0,25	1,12	0,03	0,06	0,17	0,01	1,6
Summa	1,1	50,6	26,5	4,2	6,0	10,7	0,8	100,0

Citering inom olika ämnesområden

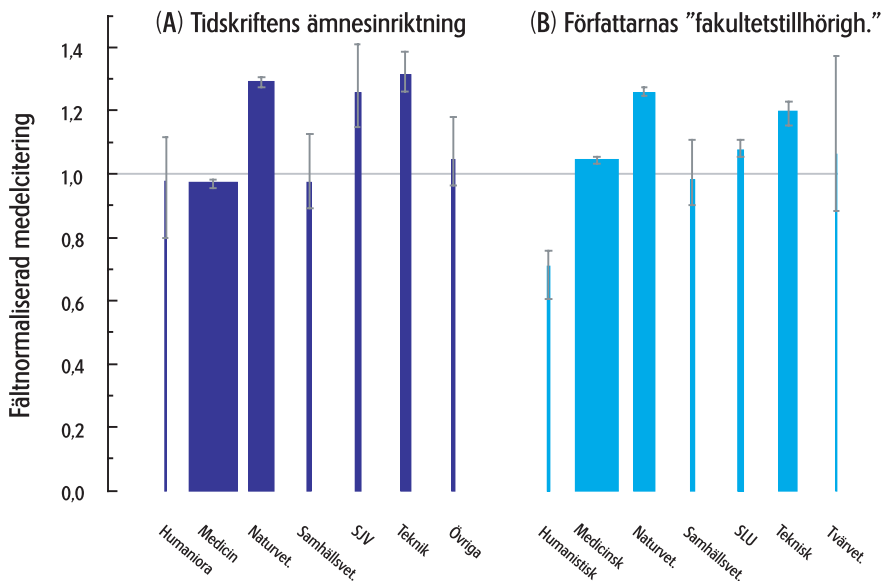
I figur 17 redovisas medelciteringen för olika ämnesområden uppdelade efter tidskrifternas ämnesklassning eller författarnas adresser (fakultets- eller vetenskapsområdestillhörighet).

Medelciteringen per ämnesområde är högst inom teknik, följt av naturvetenskap och skogs- och jordbruksvetenskap (31 till 26 % mer än världsgenomsnittet figur 17A). Humaniora, medicin och samhällsvetenskap citeras strax under världsmedelvärdet (0,97-0,98).

²⁴ Enligt databasens tidskriftklassning tillhör alla tidskrifter inom skogs- och jordbruksvetenskap samt veterinärmedicin gruppen "agronomi".

Delas publikationerna i stället upp efter författarnas adress (fakultets-tillhörighet) ser citeringsmönstret lite annorlunda ut (figur 17B). Publikationer från naturvetenskapliga fakulteter citeras högst (26 % över världsgenomsnittet), följt av teknik (20 % över genomsnittet). SLU:s genomsnitt ligger 8 % över världsmedelvärde men klart under genomsnittet för svenska publikationer inom skogs- och jordbruksvetenskap vilka citeras 26 % över världsmedelvärde. Motsvarande förhållande gäller för publikationer i tekniska tidskrifter som citeras något mera än publikationer från tekniska fakulteter (26 respektive 20 % över världsmedelvärde). Medicin visar det motsatta mönstret; publikationer från medicinska fakulteter citeras något mer (5 % över världsmedelvärde) än alla svenska publikationer i medicinska tidskrifter (3 % under världsgenomsnittet).

Figur 17. Fältnormaliserad medelcitering för svenska publikationer uppdelade på vetenskapsområden enligt (A) tidskriftens ämnesklassning och (B) författarnas fakultetstillhörighet. Medelvärde för åren 2000–2002. SJV står för skogs- och jordbruksvetenskap samt veterinärmedicin (ämneskategorin "agronomi" i databasen). Staplarnas tjocklek representerar antalet publikationer bakom respektive medelvärde (dock är tjockleken för staplarna som visar humaniora, humanistisk fakultet, övriga och tvärvetenskap överdrivna, se tabell 5 för det exakta antalet).



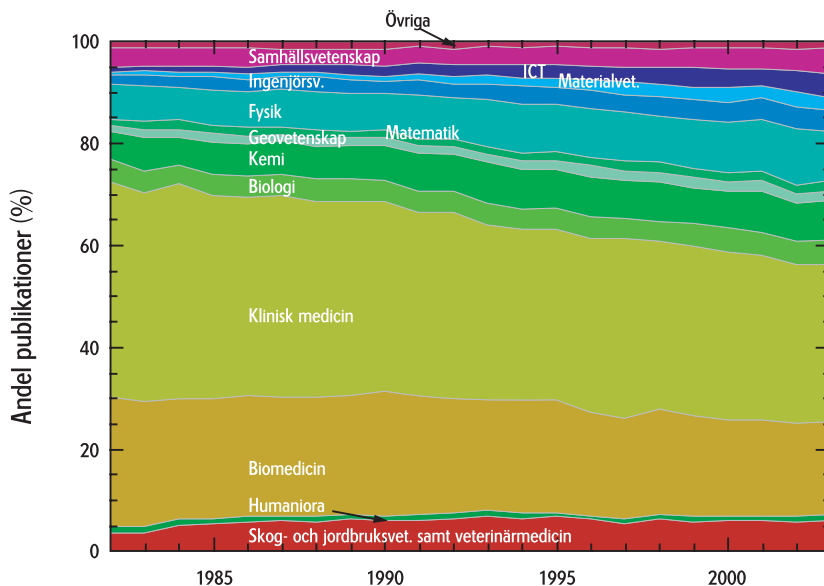
Tabell 5. Antal publikationer som medelvärdena i figur 17 är baserade på.

Uppdeln.	Hum.	Med.	Naturvet.	Samhällsv.	SJV	Teknisk	Övr./Tvär
Fakultet	94	4492	2285	444	517	1983	173
Tidskrift	103	5059	2592	411	584	1046	84

4.2 Variation inom vetenskapsområden

Figur 18 visar hur den svenska ämnesprofilen – publikationerna uppdelade på 14 ämnesområden - varierat sedan början av 1980-talet. Den största förändringen i Sveriges publiceringsmässiga ämnesprofil är att medicinområdet minskat sin andel av publiceringen betydligt sedan början av 1980-talet. Utvecklingen gäller framför allt klinisk medicin, där andelen har minskat från 42 % till 31 %. Alla andra områden utom humaniora har ökat sina andelar av antalet publikationer. De största relativa ökningarna står teknikvetenskap för; framför allt volymen inom ICT ("Information and communication technology" dvs vad som vanligen kallas för IT) och materialvetenskap ökade kraftigt. Även fysikområdet ökade relativt kraftigt (från 6,9 till 10,2 %)

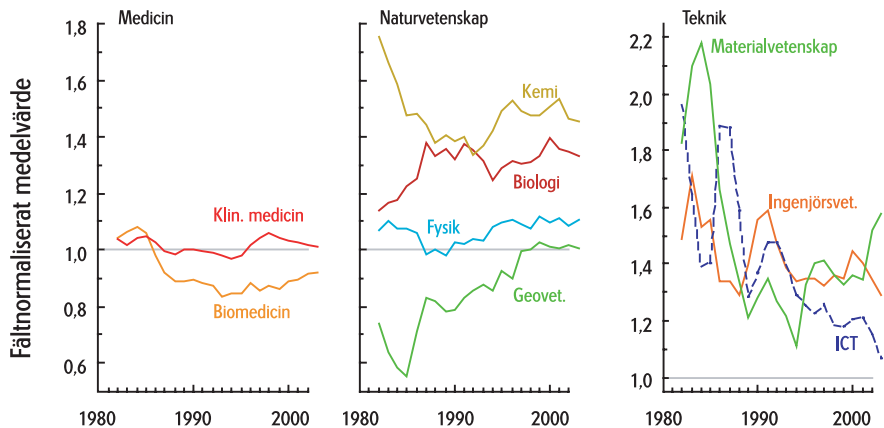
Figur 18. Förändringar i ämnesprofilen för svenska publikationer mellan 1982 och 2003. ICT står för "Information and communication technology"



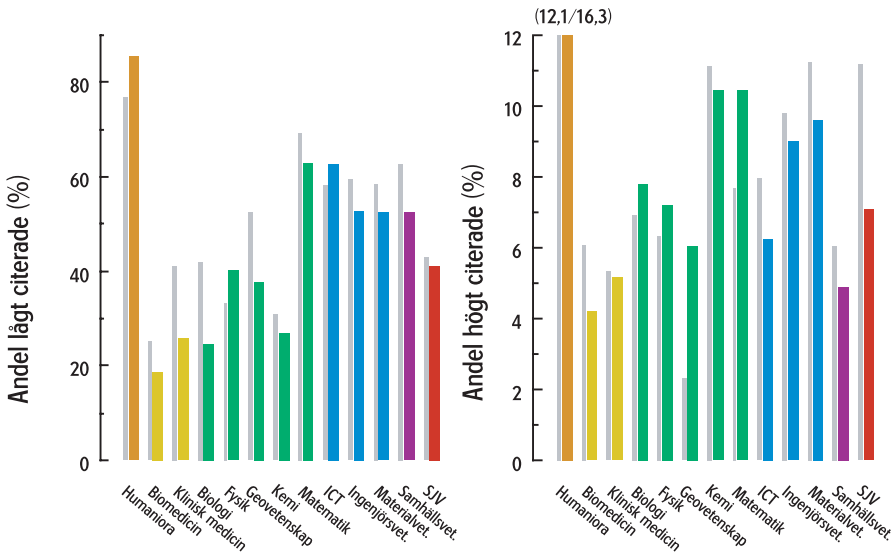
Det finns stora skillnader i citering mellan olika delar av vetenskapsområdena; figur 19 och 20 ger några exempel på detta. Inom medicin har biomedicinska ämnen utvecklats sämre (10 % under världsgenomsnittet sedan år 2000) än klinisk medicin (2 % över världsgenomsnittet). Bland de naturvetenskapliga områdena är variationen större. Kemi har backat från 60 % över världsgenomsnittet under början av 1980-talet till ca 40 % över världsgenomsnittet de senaste åren. Geovetenskap har samtidigt gått från att citeras betydligt sämre än genomsnittet till att de senaste åren ha citerats strax över världsgenomsnittet. Bland teknikvetenskap har citeringen av svenska publikationer inom materialvetenskap nästan halverats, dock från en extremt hög nivå. Alla teknikområden ligger för närvarande på 25-40 % över världsgenomsnittet.

Inom de flesta områden har andelen lågt citerade publikationer minskat något. Viktigare för förändringar i medelvärdet är att andelen högt citerade publikationer har minskat inom alla områden utom humaniora, geovetenskap och matematik. Inom dessa områden har andelen högt citerade publikationer ökat mellan tidigt 1980-tal och början av 2000-talet (figur 20). Notera att humaniora är ett problematiskt område att analysera bibliometriskt, låga fältmedelvärde för antalet citeringar resulterar i att nästa alla publikationer klassas som lågt eller högt citerade i figur 20.

Figur 19. Variation över tiden i fältnormaliserad medelcitering för tre vetenskapsområden uppdelade i mindre områden. ICT står för "Information and communication technology". Glidande 3-årsmedelvärdet. För Skogs- och jordbruksvetenskap och samhällsvetenskap se figur 15. Självciteringar är borttagna.



Figur 20. Förändringar i andelarna lågt och högt citerade svenska publikationer mellan 1982–84 (grå staplar) och 2000–2002 (färgade staplar). SJV står för skog- och jordbruksvetenskap samt veterinärmedicin. Andelen högt citerade är andelen bland de 5 % mest citerade publikationerna (citerade lika mycket eller mer än den 95e percentilen). Notera att staplarna för andelen högt citerade publikationer inom humaniora är förkortade. Själv citeringar är borttagna. Antalet publikationer per område redovisas i tabell 6.



Tabell 6. Antal fraktioniserade publikationer per ämnesområde (medelvärde för åren 2000–2002).

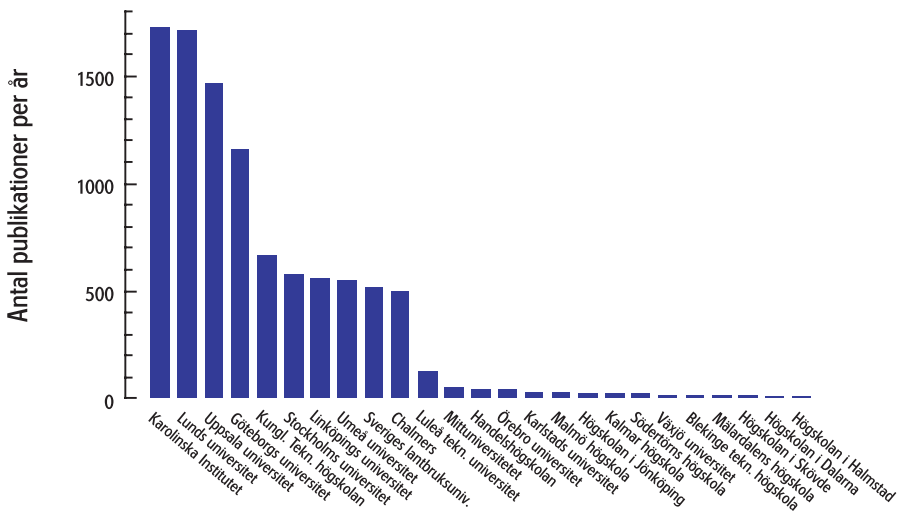
Område	Antal publikationer per år
Humaniora	103
Biomedicin	1836
Klinisk medicin	3137
Biologi	453
Fysik	1003
Geovetenskap	196
Kemi	741
Matematik	183
ICT	369
Ingenjörsvetenskap	406
Materialvetenskap	265
Samhällsvetenskap	411
Skog- och jordbruksvetenskap samt veterinärmedicin	570

4.3 Variationer mellan lärosäten

A) Antal publikationer

Karolinska institutet, följt av Lunds universitet, Uppsala universitet och Göteborgs universitet är störst bland svenska lärosäten när man mäter det genomsnittliga antalet fraktionaliserade publikationer åren 2000 till 2002 (figur 21). Alla dessa lärosäten producerar mer än 1000 publikationer per år. Därefter kommer en grupp om ytterligare sex lärosäten som producerar mellan 500 och 700 publikationer per år. Luleå tekniska universitet hade dessa år 120 publikationer i genomsnitt, medan ytterligare 14 lärosäten (högskolor och nya universitet) hade en volym mellan 10 och 50 publikationer per år.

Figur 21. Antal publikationer från svenska universitet och högskolor. Publikationerna är fraktionaliserade så att varje lärosäte får andelar i proportion till antalet adresser på respektive publikation²⁵. Medelvärde för åren 2000–2002. Endast lärosäten med ett genomsnitt om minst 10 publikationer per år är inkluderade i figuren. Publikationsvolymen för de lärosäten med mindre än 50 publikationer per år redovisas också i tabell 7.



²⁵Se metodbilaga för detaljer.

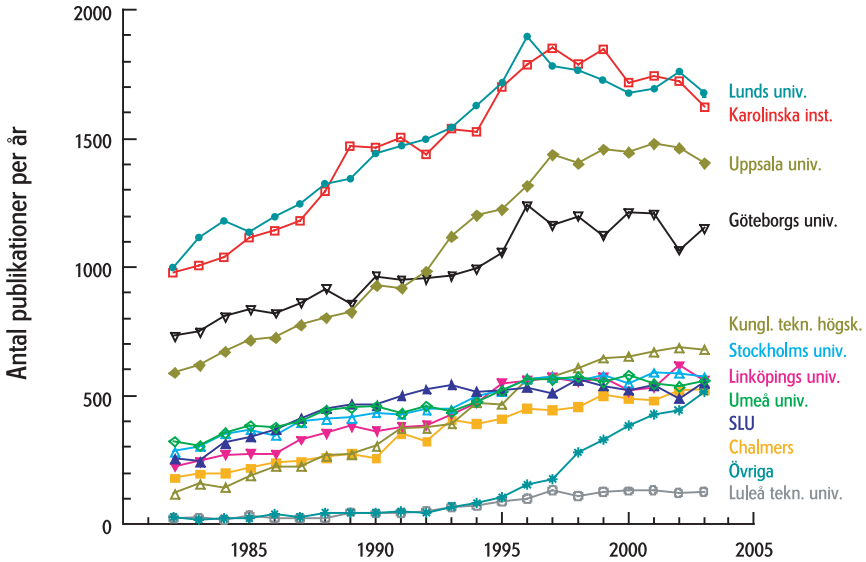
Tabell 7. Publikationsvolym för små lärosäten (mindre än 100 publikationer per år), Dessa är hopslagna till gruppen "Övriga" i figur 22 och 23, Genomsnitt av fraktioniserade publikationer för åren 2000–2002 ,

Lärosäte	Publikations- volym (frak. publ./år)	Lärosäte	Publikations- volym (frak. publ./år)
Mittuniversitetet	47,5	Mälardalens högskola	16,0
Handelshögskolan	46,9	Skövde högskola	13,6
Örebro universitet	46,1	Högskolan i Dalarna	10,9
Karlstads universitet	29,4	Halmstads högskola	9,9
Malmö högskola	28,0	Högskolan i Gävle	9,1
Högskolan i Jönköping	22,9	Högskolan i Kristianstad	7,3
Högskolan i Kalmar	21,0	Högskolan Väst	6,4
Växjö universitet	19,9	Högskolan i Borås	5,8
Södertörn högskola	19,7	Högskolan på Gotland	1,3
Blekinge tekn. högskola	18,6		

De största lärosätena, dvs. de som för närvarande har en publikationsvolym på minst 1000 publikationer per år, ökade sin volym markant fram till mitten av 1990-talet och har sedan dess legat på en relativt konstant nivå (figur 24). Stabiliseringen av antalet fraktioniserade publikationer under mitten av 1990-talet beror till stor del på att publikationerna i ökade grad är resultatet av samarbete. Det totala antalet publikationer som de större lärosätena bidrog till ökade fram till millennieskiftet, sedan dess har även det totala antalet publikationer legat stabilt.

Flera av de mindre universiteterna och högskolorna, framför allt de med ett genomsnitt under 50 publikationer per år i figur 21, har visat en starkare tillväxt det senaste årtiondet (sammanslagna till gruppen "Övriga" i figur 22 och 23). Gruppen Övriga har ökat sin volym drygt 7-faldigt mellan 1993 och 2003. Av figuren framgår det att också Uppsala universitet ökat sin publiceringsvolym kraftigare än Göteborgs universitet. Bland de äldre lärosätena har Kungliga Tekniska högskolan ökat publiceringsvolymen mest. I början av 1980-talet låg Kungliga Tekniska högskolan på 10:e plats bland svenska lärosäten, medan lärosätet de senaste åren hamnat på en femte plats.

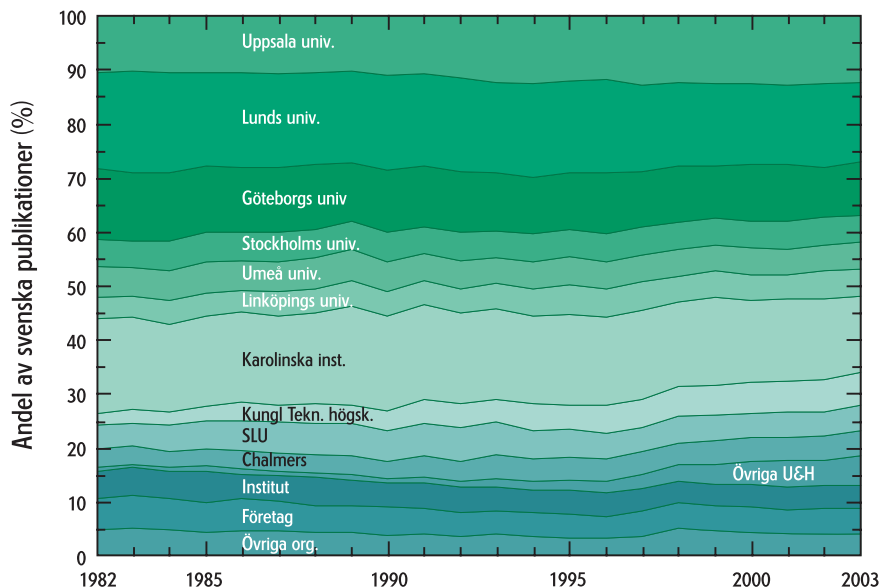
Figur 22. Variation i publikationsvolym för svenska universitet och högskolor. Alla publikationer är fraktioniserade på de adresser som finns angivna på respektive publikation. Alla lärosäten med en genomsnittlig volym under 50 publikationer per år sedan år 2000 är sammanslagna till "Övriga".



Universitet och högskolor står för knappt 90 % av alla svenska publikationer; andelen har varierat mellan 84 och 88 % sedan 1982 (figur 23). För närvarande (åren 2000–2002) står företag, institut och övriga organisationstyper för drygt 4 % vardera av de svenska publikationerna.

De största förändringarna bland universitet och högskolor mellan 1982–85 och 2001–03 var att publikationerna från tre lärosäten minskade sin andel: Lunds universitet och Göteborgs universitet med ca 3 % vardera och Karolinska institutet med 2 %. Kungliga tekniska högskolan och gruppen Övriga universitet och högskolor ökade båda med ca 4 % och Uppsala universitet med drygt 2 %. För alla andra lärosäten var förändringen mindre än 2 %.

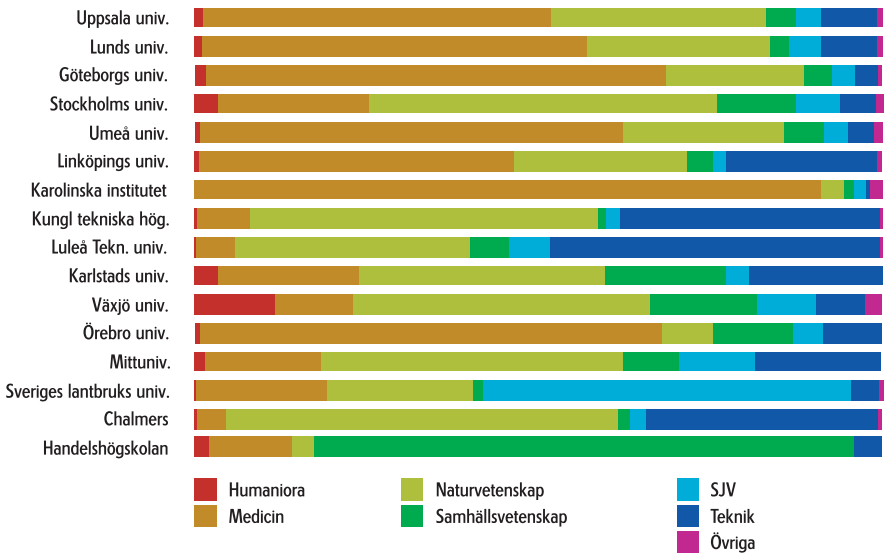
Figur 23. De relativa bidragen från lärosäten, institut, företag och övriga organisationer till de svenska publikationerna mellan 1982 och 2004. "Övriga U&H" är lärosäten med en volym på mindre än 150 publikationer per år. Publikationerna från gruppen "Övriga org." består till ca 75 % av sjukhus, andra som ingår här är kommuner och museer (med 8 resp. 3 % av gruppens volym).



Lärosätenas ämnesprofil

Lärosätenas ämnesmässiga publiceringsprofil redovisas i figur 24. För 7 av de 15 lärosäten som redovisas i figuren är medicin det största ämnet, mätt med antal publikationer. Den mest markerade specialiseringen finns hos Karolinska institutet med 91 % av publikationerna i medicinska tidskrifter. Andra lärosäten med en stor andel medicin är Göteborgs universitet och Umeå universitet (66 % respektive 61 %). För sex av lärosätena är naturvetenskap det största området. Bland lärosätena med teknisk inriktning är det bara Luleå tekniska universitet där teknik är det dominerade ämnet publiceringsmässigt (45 %). Forskarna vid Chalmers och Kungliga tekniska högskolan publicerar en större andel i naturvetenskapliga tidskrifter än i tekniska (51–54 % i naturvetenskap jämfört med 36–38 % i teknik).

Figur 24. Universitetens samt Chalmers och Handelshögskolans ämnesprofiler; publikationer uppdelade på vetenskapsområden (humaniora och samhällsvetenskap separerade). SJV står för skog- och jordbruksvetenskap samt veterinärmedicin (ämneskategorin "agronomi" i databasen). Figuren är baserad på fraktioniserade publikationer mellan 2000 och 2002. Gruppen Övriga innehåller framför allt tvärvetenskapliga tidskrifter. Ämnesklassningen är baserad på respektive tidskrifts ämnesklassning i databasen.



B) Citering

Fältnormaliserad medelcitering på lärosätetsnivå

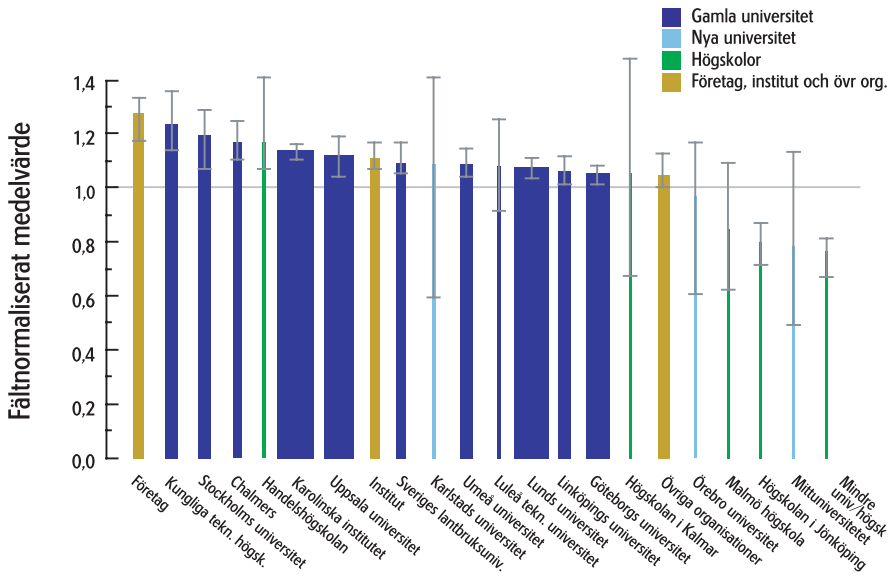
Den fältnormaliserade medelciteringen summerad per lärosäte för åren 1998–2002 redovisas i figur 25. Fjorton lärosäten hade ett medelvärde över världsgenomsnittet (1,0). Skillnaden mellan lärosätena är relativt liten jämfört med variationen mellan åren (illustrerad med vertikala grå liner). För jämförelse redovisas också medelciteringen för tre icke-akademiska grupperingar; företag, institut och övriga icke-akademiska organisationer (huvudsakligen sjukhus, notera att universitetssjukhusen ingår i respektive universitets publikationer). Det kan tyckas anmärkningsvärt att företagspublikationerna har en högre medelcitering än något av universiteten medan publikationer från institut och övriga organisation citeras i ungefär samma utsträckning som merparten av universiteten.

Nedan har vi delat upp lärosätena i tre grupper; (1) gamla universitet, dvs. de som erhållit examensrätt 1992 eller tidigare (hit räknar vi också Chalmers) (2) nya universitet, dvs. de som erhållit sin examinationsrätt efter 1992 och (3) högskolor. Till högskolorna räknar vi även Handelshögskolan. Dessutom redovisas medelvärden för publikationer från svenska företag, institut och övriga organisationer (huvudsakligen sjukhus).

Publikationerna från Kungliga tekniska högskolan citerades mest, följt av Handelshögskolan och Stockholms universitet (23 respektive 20 % och 19 % över världsgenomsnittet). De gamla universiteten dominerar bland de högst citerade lärosätena. Små skillnader i medelvärdet mellan många lärosäten i kombination med stora variationer mellan åren gör att förekomsten av ett enskilt högt citerat arbete kan leda till relativt stora förändringar i rangordningen, framför allt för lärosäten med ca 50 publikationer per år eller mindre. De långa spridningsstaplarna i figur 25 illustrerar denna variabilitet. Ett exempel på detta är Karlstads universitets medelvärde som är mycket beroende på ett starkt år - utan detta år skulle Karlstad flyttas ner 6 positioner.

Rangordningen mellan lärosäten är relativt oberoende av om självciteringarna tas bort eller ej. De två lärosäten som påverkades mest var Karlstads och Göteborgs universitet. Karlstad hade en låg andel självcitat och flyttas ner 5 positioner om beräkningarna baseras på alla citat medan Göteborgs universitet hade en relativt hög grad av självciteringar och skulle flyttas upp tre positioner om självciteringarna tas med. För alla andra lärosäten var rangordningen stabil (maximal förändring ± 2 positioner).

Figur 25. Medelcitering för svenska lärosäten (medelvärde för åren 1998–2002). Den horisontella grå linjen visar världsgenomsnittet (1,0) och den vertikala linjen i varje stapel visar lägsta och högsta årsmedelvärde för respektive lärosäte under de fem åren. Lärosäten mindre än 100 publikationer under perioden, dvs. mindre än 20 publikationer per år, redovisas i figuren sammanslagna till "Mindre universitet och högskolor" (se tabell 6). För lärosäten med 100 eller mer publikationer per år är staplarnas bredd proportionerlig mot respektive lärosätes publikationsvolym (staplarna för mindre lärosäten är överdrivna). Nya universitet är här definierade som de som fått examinationsrätt 1993 eller senare²⁶. Medelvärden redovisas även för svenska företag²⁷, institut och övriga organisationer (huvudsakligen sjukhus). Själv citeringar är borttagna.



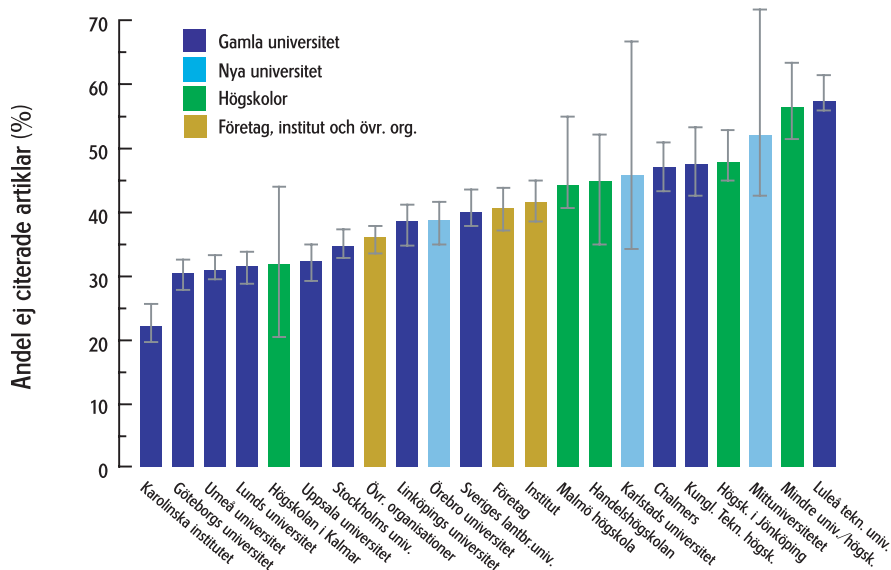
Andel lågt och högt citerade publikationer

Karolinska institutet hade lägst andel lågt citerade artiklar (21 %, figur 26). Sedan följer en grupp bestående av Göteborgs, Umeå, Lund och Uppsala universitet samt Högskolan i Kalmar med ca 30 % av artiklarna utan citeringar efter två år.

²⁶Här räknar vi Chalmers till de gamla universiteten och Handelshögskolan till högskolorna.

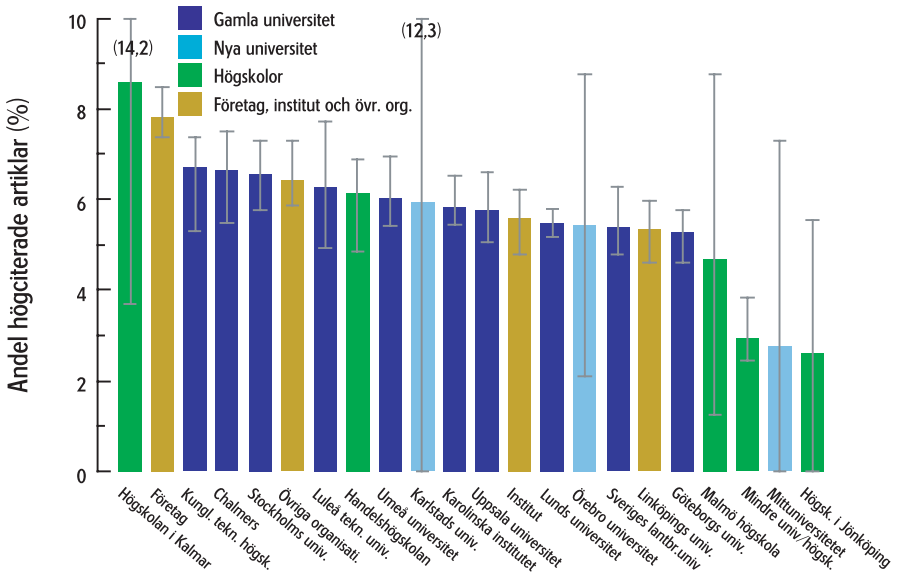
²⁷Knappt 60 % av företagens publicering sker i samarbete med högskolor. Företagets publicering redovisas i en separat kommande rapport från Vetenskapsrådets analysenhet.

Figur 26. Andel publikationer från åren 1998–2002 som ej citerats under de första två åren efter publiceringen. Den vertikala linjen i varje stapel visar lägsta och högsta årsmedelvärde för respektive lärosäte under de fem åren. Lärosäten mindre än 100 publikationer under perioden redovisas i figuren sammanslagna till "Mindre universitet och högskolor" (se tabell 6). Själv citeringar är borttagna.



Högskolan i Kalmar hade högst andel högt citerade publikationer (figur 27), men som framgår av variationen av spridningsstapeln är denna högskolas värde starkt påverkat av ett enskilt år. Bland de större lärosätena har Kungl. Tekniska högskolan högst andel högt citerade publikationer. Även företagens publikationer har en hög andel högt citerade.

Figur 27. Andel av respektive lärosätes publikationer som finns bland de högst citerade 5 % av världsproduktionen, dvs. de som citeras lika mycket eller mer än den 95:e percentilen. Baserat på publikationer från åren 1998-2002. Dessutom redovisas medelvärden för publikationer från svenska företag, institut och övriga organisationer (huvudsakligen sjukhus). Den horisontella grå linjen visar världsgenomsnittet (5,0) och den vertikala linjen i varje stapel visar lägsta och högsta årsmedelvärde för respektive lärosäte under de fem åren. Lärosäten mindre än 100 publikationer under perioden redovisas i figuren sammanslagna till "Mindre universitet och högskolor" (se tabell 6). Själv citeringar är borttagna.



Det finns ingen samvariation mellan medelcitering (figur 25) och andelen lågt citerade (figur 26) men däremot mellan medelcitering och andelen högt citerade publikationer (figur 27, Spearman rankkorrelation = 0,70, $P = 0,0004$). Framförallt var det högskolan i Kalmar som avvek från detta mönster med en stor skillnad i rangordningarna för medelcitering och andel högt citerade.

C) De största svenska producenterna av artiklar uppdelade på vetenskapsområde

I tabell 8 redovisas citeringsstatistiken för de fem största lärosäten inom respektive vetenskapsområde. Ämnesspecialiserade universitet dominerar inom sina områden: Sveriges lantbruksuniversitet dominerar skog- och jordbruksrelaterade ämnen, Karolinska institutet medicinpubliceringen och de tekniska högskolorna publiceringen inom teknik. Stockholms uni-

versitet är störst inom humaniora och samhällsvetenskap medan Uppsala universitet står för den största volymen inom naturvetenskap. Dock är det oftast inte den största aktören som får högst genomsnittlig citering per publikation. Medicin och naturvetenskap är de enda områdena där lärosätet med det största antalet publikationer också citeras mest.

Tabell 8. De fem största lärosätena i termer av publiceringsvolym inom respektive vetenskapsområde (enligt tidskrifternas ämnesklassning). Andelen lågt citerade är andelen som inte fått några citat under de två första åren och andelen högt citerade är de 5 % mest citerade (de som citerats lika mycket eller mer än den 95:e percentilen). Baserad på publikationer 1998–2002, självciteringar är borttagna. Ingen publiceringsstatistik redovisas för lärosäten med mindre än 20 publikationer per år (minst 100 publikationer under perioden vilket vi satt som minivolum för att redovisa citeringsstatistik). Notera att enbart ett lärosäte nådde denna gräns inom humaniora medan ytterligare tre hade något lägre volym men redovisas ändå. Antalet publikationer är medelvärdet av fraktioniserade publikationer under de fem åren.

Vetenskapsområde	Lärosäte	Fält-normaliserat medelvärde	Andel högt citerade	Andel lågt citerade	Antal publikationer per år
Humaniora	Stockholms universitet	0,83	24,0	85,3	20
	Göteborgs universitet	0,22	1,9	93,2	19
	Lunds universitet	0,83	17,1	84,9	19
	Uppsala universitet	1,65	16,4	85,2	18
	Umeå universitet				4
Medicin	Karolinska institutet	1,10	6,1	21,4	1581
	Lunds universitet	0,91	3,9	25,0	958
	Göteborgs universitet	1,00	5,0	25,7	763
	Uppsala universitet	0,95	4,6	24,4	736
	Umeå universitet	0,93	4,7	25,9	338
Naturvetenskap	Lunds universitet	1,42	9,7	31,3	447
	Uppsala universitet	1,27	7,8	35,6	436
	Kungl. Tekniska högskolan	1,27	8,2	42,4	326
	Stockholms universitet	1,28	7,7	31,5	273
	Chalmers	1,12	7,6	44,4	263
Samhällsvetenskap	Stockholms universitet	1,14	5,2	49,0	65
	Uppsala universitet	1,12	5,4	53,0	61
	Göteborgs universitet	0,70	3,8	57,7	44
	Lunds universitet	0,87	4,1	52,3	43
	Handelshögskolan	1,26	7,3	47,9	38

Tabellen fortsätter på nästa sida.

Skogs- och jord- bruk samt veterinärmedicin	Sveriges lantbruksuniversitet	1,21	7,5	45,4	284
	Lunds universitet	1,17	6,5	40,4	79
	Uppsala universitet	1,00	5,6	41,9	51
	Stockholms universitet	1,96	9,7	34,5	36
	Göteborgs universitet	1,15	7,1	36,3	32
Teknik	Kungl. Tekniska högskolan	1,32	7,8	59,4	245
	Chalmers	1,35	8,0	56,0	164
	Lunds universitet	1,26	7,2	59,1	140
	Linköpings universitet	1,30	9,3	55,4	118
	Uppsala universitet	1,60	10,8	51,7	113

Tabellen illustrerar också problemet med att analysera humanistisk forskning bibliometriskt. Mycket få publikationer i humaniora finns med i databasen; enbart ett lärosäte hade mer än 20 publikationer per år. Övriga lärosäten ligger nedanför den gräns på 100 publikationer under den summerade perioden (1998–2002) som vi valt som minimum för att presentera citeringsstatistik.

5 SLUTKOMMENTAR

Det mått som vi huvudsakligen använt oss av här, fältnormaliserad medelcitering utan självciteringar, visar att Sverige legat på en relativt konstant nivå, 10 ± 5 % över världsgenomsnittet sedan början av 1980-talet (figur 10A). Detta har inneburit att Sverige legat på sjätte plats i världsrankingen både i början och i slutet av perioden. Väljer man i stället att jämföra Sverige med Europagenomsnittet har Sverige haft en nedåtgående trend under i stort sett hela perioden (figur 10B).

Man kan alltså ifrågasätta påståendet som förekommer i den forskningspolitiska debatten att vetenskapliga publikationer från Sverige har sjunkit i kvalitet. Däremot finns flera europeiska länder, som t.ex. Danmark, Nederländerna, Norge och Tyskland, där citeringsnivåerna har ökat markant under det senaste årtiondet.

Orsaken till Sveriges försämrade position i förhållande till övriga Europa kan ses på två olika sätt:

- 1 Andelen av de svenska publikationerna som blivit högt citerade har minskat sedan 1980-talet. De högt citerade publikationerna påverkar landsmedelvärdet relativt mycket. Den negativa effekten av detta har balanserats av att kvalitén på de "sämsta" publikationerna har förbättrats, vilket avspeglas i att andelen lågt citerade svenska publikationer har minskat. Om man jämför med publikationerna från USA eller Storbritannien, som citeras mer än de svenska, har dessa länder en något högre andel av lågt citerade publikationer än vad Sverige har, men de har också en högre andel högt citerade publikationer.
- 2 Ett annat perspektiv på Sveriges relativt dåliga utveckling jämfört med övriga europeiska länder kan man få genom att bryta upp statistiken på vetenskapsområden. Hälften av alla svenska publikationer finns inom medicin. Klinisk medicin, som svarar för 32 % av alla svenska publikationer, har haft en relativt stabil citeringsnivå i närheten av världsmedelvärdet medan biomedicin (19 % av publikationerna) har visat en minskande medelcitering och citeras för närvarande 10 % under världsmedelvärdet (fältnormaliserat medelvärde 0,90). Den medelmåttiga citeringen av medicinska publikationer kan ses som en viktig förklaring till att Sveriges medelvärde inte är bättre. Andra stora områden (naturvetenskap, skogs- och jordbruks relaterade ämnen och teknik) har visserligen också visat minskande citeringssiffror men utifrån mycket bra utgångslägen. Dessa områden citeras fortfarande ca 25–30 % över världsgenomsnittet.

Om medicinpublikationerna hade citerats i samma uträkning som publikationerna inom naturvetenskap eller teknik hade Sverige legat några positioner högre i världsrankingen. Det verkar dock generellt vara svårare att uppnå höga medelvärden inom medicinområdet; bara två länder (USA och Schweiz) har medelvärden inom medicin över 1,1 medan sju länder ligger över denna nivå om medelvärdet beräknas för publikationer inom samtliga områden.

Vi vill också upprepa att humaniora är svårt att analysera bibliometriskt på grund av att relativt få publikationer inom detta område finns med i databasen. Av samma anledning varierar medelvärdet för små lärosäten betydligt mellan år.

6 CITERAD LITTERATUR

- Adams, J. 2005. Early citations counts correlate with accumulated impact. - *Scientometrics* 63: 567-581.
- Agrawal, A.A. 2005. Corruption of journal impact factors. - *Trends in Ecology and Evolution* 20: 157.
- Aksnes D.W. 2003. A macro study on self-citation. - *Scientometrics* 56: 235-246.
- Aksnes D.W. 2005. Citation rates and perceptions of scientific contribution. - *J. American Soc for Information Sci and Technol.* 57: 169-185.
- Aksnes D.W. & Sivertsen G 2004. The effect of highly cited papers on national citation indicators. - *Scientometrics* 59: 213-224.
- Archambault É., Vignola-Gange É, Côte G., Larivière V. & Gringras Y. 2006. Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. - *Scientometrics* 68: 329-342.
- Case D.O., & Higgins G.M. 2000. How can we investigate citation behavior? A study of reasons for citing literature in communication. - *J. Am Soc Inf. Sci.* 51: 635-645.
- Cavallin M. & Lindblad S. 2006. Världsmästerskap i vetenskap? - Göteborgs universitet. (Dnr GII 530/06)
- Glänzel W 2000. Science in Scandinavia: A bibliometric approach. - *Scientometrics* 48:121-150.
- Glänzel W., Danell R. & Persson O. 2003. The decline of Swedish neuroscience: Decomposing a bibliometric national science indicator. - *Scientometrics* 57:197-213.
- Glänzel W. & Garfield E. 2004. The myth of delayed recognition. - *The Scientist* 18: 8.
- Glänzel W., Thijs B. & Schubert A. 2005. The weight of author self-citations. - In: *Proceedings of ISSI 2005*, sid. 382-389.
- Heyman U. 2003. Sveriges vetenskapliga publiceing – en analys av NI. PM från Vetenskaprådet. <http://www.vr.se/publikationer/sida.jsp?resourceId=256>
- Ingwersen P & Wormell I 1999. Publication behaviour and international impact: Scandinavian clinical and social medicine 1988-1996. *Scientometrics* 46: 487-499.
- Kotiaho J.S. 1999. Papers vanish in the mis-citation black hole. - *Nature* 398: 19.
- Lawrence P. 2003. The politics of publication. - *Nature* 422: 259-261
- van Leeuwen T.N., Visser M.S., Moed H.F., Nederhoof T.J. & van Raan A.F.J. 2003. The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. - *Scientometrics* 57: 257-280.
- Moed H.F. 2005a. *Citation analysis in research evaluation*. Springer.
- Moed H.F. 2005b. Citation analysis of scientific journals and journal impact measures. - *Current Science* 89: 1990-1996.
- van Raan A.F.J. 1996. Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics*, 36(3):397-420.
- van Raan A.F.J. 2004a. Sleeping beauties in science. - *Scientometrics* 59: 467-472.

- van Raan A.F.J. 2004b. Measuring science. - In: Moed H.F. et al (eds) Handbook of Quantitative Science and Technology Research, Kluwer Academic Publ., pp 19-50.
- van Raan, A.J.F. 2005. Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. - *Scientometrics* 62: 133-143.
- Seglen P.O. 1994. Causal relationship between article citedness and journal impact. - *J. Am Soc Inf. Sci* 45: 1-11.
- Wallin J.A. 2005. Bibliometric methods: pitfalls and possibilities. - *Basic & Clinical Pharmacology and Toxicology* 97: 261-275.
- Weingart P. 2005. Impact of bibliometrics upon the science system: inadvertent consequences? - *Scientometrics* 62: 117-131.
- Zetterström R. 2002. Bibliometric data: a disaster for many non-American biomedical journals. - *Acata Paediatrica* 91: 1020-2024.

Gruppering av svenska lärosäten

Vid beräkning av publicering, citering etc. för svenska lärosäten har universitetssjukhusen slagits ihop med respektive universitet; Karolinska institutet inkluderar Karolinska sjukhusets²⁸ publikationer, Örebro universitet inkluderar Örebro universitetssjukhus, Göteborgs universitet inkluderar Sahlgrenska universitetssjukhuset etc.

Databasen

Grundmaterialet i Vetenskapsrådets publikationsdatabas är inköpt från Thomson Scientific och är det samma som finns tillgängligt via Web of Science²⁹. Databasen innehåller för närvarande ca 23 miljoner publikationer från 1982 till 2004 och uppdateras årligen. Databasen innehåller dessutom ett antal variabler som vi beräknat ur grundmaterialet som t.ex. antal citeringar en publikation fått 2 eller 5 år efter publiceringen och medelcitering i olika forskningsfält. Vidare är svenska organisationsnamn (lärosäten, företag m.m.) rättade och alla svenska adresser är klassade med avseende på organisationstyp (universitet, högskola, företag etc.). Dessutom har vi gjort en vetenskapsområdes- (fakultets-) klassning baserad på svenska författares institutionstillhörighet via adressfälten.

Normalisering av citeringar

Det enklaste sättet att uppskatta den uppmärksamhet som en publikation erhållit är att räkna hur många gånger den citerats ("citations per paper", CPP). Antalet citeringar för en artikel ökar självklart med tiden, det är därför vanligt att mäta antalet citeringar efter en viss tidsperiod efter publi-

²⁸Till Karolinska sjukhuset räknas också Huddinge och Danderyds sjukhus samt Södersjukhuset.

²⁹Web of Science inkluderar Thomson Scientific:s alla databaser: Science Citation Index Expanded, the Social Science Citation Index, och the Arts & Humanities Citation Index.

ceringen. Vi använder oftast antalet citeringar två år efter publiceringsåret (CPP₂), dvs. i genomsnitt 2,5 år efter publiceringen. Detta mått är nära korrelerat med antalet citeringarna efter längre tidsperioder, högt citerade artiklar upptäcks och citeras oftast mycket även tidigt efter publiceringen (se avsnitt 2.2 för en analys och diskussion om detta).

Ett problem med jämförelser av antalet citeringar per publikation mellan olika forskningsfält är att citeringstraditionerna varierar mellan olika områden. En metod för att korrigera för detta är att använda "fältnormaliserad citering". Alla tidskrifter som indexerats i databasen är klassade till ett eller flera av 252 forskningsfält t.ex. "rheumatology" eller "behavioral sciences".

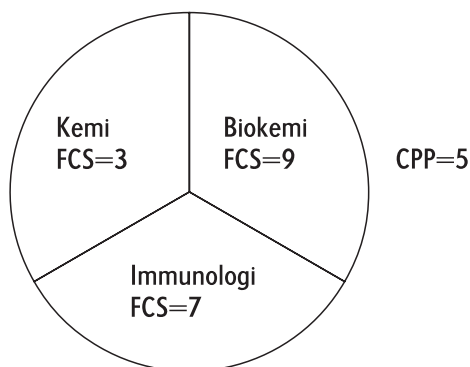
Ett fältmedelvärde (field citation score, FCS) beräknas för varje kombination av fält, publikationsår och typ av publikation ("articles", "reviews" etc.). Eftersom en tidskrift kan tillhöra flera ämnen samtidigt är det nödvändigt att ta hänsyn till detta. För att undvika att räkna artiklar dubbelt delar vi de artiklar som tillhör flera fält i lika många delar som antalet fält, detta medför att vi vet att det vägda medelvärdet för alla 252 FCS blir lika med det totala medelvärdet för citat per publikation i världen. Beräkningen görs på följande sätt för ett fält. Låt N vara antalet publikationer inom ett fält, under ett visst år och av en viss publikationstyp, låt C_i vara antalet citat för den i :e publikationen och F_i vara antalet ämnesområden (fält) som den i :e publikationen klassats till³⁰. Då blir FCS för givet fält, år och typ:

$$FCS = \frac{\sum_{i=1}^N \frac{C_i}{F_i}}{\sum_{i=1}^N \frac{1}{F_i}}$$

Fältnormaliseringen innebär att man beräknar antalet citat av ett papper relativt fältmedelvärdet i det eller de fält den aktuella tidskriften tillhör. Tillhör tidskriften bara ett fält blir alltså den fältnormaliserade citeringen = CPP/FCS, där både CPP och FCS beräknats efter samma tidsperiod (oftast 2 eller 5 år) efter publiceringen.

Vill man jämföra en publikations citering mot ett världsgenomsnitt när tidskriften är klassad till flera ämnesområden måste man anpassa metoden något genom att dela upp papperet i delar. Vi kan exemplifiera detta med nedanstående figur där cirkeln får representera en artikel citerad 5 gånger i en tidskrift klassad till tre olika fält.

³⁰I de flesta fall är F_i lika med ett. Ibland är F_i två eller tre och maximalt är F_i sex.



Vi anser att artikeln bidrar med $\frac{1}{3}$ artikel inom biokemi med $CPP/FCS=5/9$, med $\frac{1}{3}$ artikel inom immunologi med $CPP/FCS=5/7$ och med $\frac{1}{3}$ artikel inom kemi med $CPP/FCS=5/3$. Dessa tre delar bidrar till ett gemensamt medelvärde för hela artikelns CPP/FCS .

Vill man ta fram ett index för enbart medicinska artiklar använder vi bara de "tårtbitar" som tillhör medicin, i det här fallet biokemi och immunologi som bidrar med varsin $\frac{1}{3}$ till indexet.

Om man vill beräkna ett fältnormaliserat medelvärde för publikationerna från en geografisk enhet t.ex. lärosäte eller land, utför vi samma typ av uppdelning genom att inkludera informationen från adressfälten, som säger hur stor del av publikationen som är skriven inom den geografiska enheten. Detta kallas för fraktionalisering och görs genom att vi delar upp publikationen i lika många delar det finns författaradresser angivna på artikeln och sedan ger varje geografisk enhet så många av dessa delar som antalet adresser därifrån. Om t.ex. en artikel har en adress från Sverige, två från Norge och fem ifrån Tyskland så krediteras Sverige $\frac{1}{8}$, Norge $\frac{2}{8}$ och Tyskland $\frac{5}{8}$ av artikeln. Fraktionaliseringen ger en bättre representation av arbetsinsatsen för en artikel än den enklare metoden att tillerkänna varje adress en hel artikel. Dessutom säkerställer fraktionaliseringen att världsmedelvärdet av CPP/FCS blir ett.

En publikation som citeras lika mycket som världsgenomsnittet får då $CPP/FCS = 1$, en kvot på 2 innebär att artikeln citeras dubbelt så mycket som världsgenomsnittet. När vi skriver CPP/FCS för en viss mängd publikationer underförstår vi att ovanstående uppdelning på fält och fraktionalisering på adresser utförs och att kvoten CPP/FCS beräknas som ett vägt medelvärde av de ingående delarna av publikationerna.

Ett index som är mycket snarlikt CPP/FCS har kallats för "the crown indicator" eller kronindikatorn av dess upphovsmän, bl.a. van Raan, eftersom de anser detta vara det bästa sättet att jämföra citeringsstatistik för

grupper eller personer aktiva i olika forskningsfält³¹. Vårt sätt att beräkna FCS skiljer sig emellertid från van Raan's eftersom vi använder ett tidsfönster för FCS som är identiskt med det för CPP. För ett tvåårigt citeringsfönster ser vi efter hur mycket alla publikationer i ett visst fält, publicerade under t.ex. år 2000, citerades t.o.m. år 2002. Enligt originaldefinitionen är FCS för år 2000 hur mycket de artiklar som publicerades i fältet under 1995 till 1998 citerades under år 2000. Vidare jämför vi, genom att använda CPP/FCS, varje artikels citeringsnivå med ett fältmedelvärde medan originaldefinitionen av kronindikatorn beräknas genom att ta medelciteringsnivån för, säg, ett lärosäte och dividera med medelfältnormalen.

En annan metod för normalisering är att använda en grupp med länder som bas. I denna rapport presenterar vi några analyser där medelvärdet för en grupp av 17 europeiska länder använts som bas. Länderna vi valt som referensgrupp är den s.k. EU-15 gruppen (Belgien, Danmark, Finland, Frankrike, Grekland, Nederländerna, Irland, Italien, Luxemborg, Portugal, Spanien, Storbritannien, Sverige, Tyskland och Österrike) plus Schweiz och Norge. Det europnormaliserade medelvärdet för t.ex. Sverige (SV) blir då:

$$\text{medelvärde}_{SV} = \frac{\sum \text{citat}_{SV} / \sum \text{publikationer}_{SV}}{\sum \text{citat}_{EU17} / \sum \text{publikationer}_{EU17}}$$

Både citat och publikationer är fraktioniserade med avseende på adresser på samma sätt som beskrivits ovan.

Andelarna högt och lågt citerade publikationer

Andelen lågt citerade av publikationerna från ett land eller lärosäte definieras som andelen publikationer som inte erhållit några citeringar under de första 2 åren efter publiceringen. Högt citerade publikationer definieras som de 5 % eller 1 % av alla publikationer som citeras mest (dvs citeras lika mycket eller mer än 95 respektive 99 percentilerna), observera att percentilgränserna är inkluderande så att världsgenomsnittet för andelen artiklar bland de 5 % högst citerade blir något högre än 5 %.

³¹ Se t.ex. van Raan (1996, 2004).

Eliminering av självciteringar

Självciteringar tagits bort binärt, dvs en citering räknas bara om inget överlapp finns mellan författarlistorna för det citerade och citerande publikationerna.

Publicerings- och citeringsstatistik används allt oftare vid nationella och internationella jämförelser och utvärderingar av forskningens volym och kvalitet. I denna rapport redovisas vad sådan statistik säger om den svenska forskningens utveckling under de senaste decennierna i ett internationellt perspektiv. Vi presenterar också statistik om variationer mellan svenska lärosäten och ämnesområden när det gäller produktion av vetenskapliga artiklar och hur mycket de citeras. Citeringarna har huvudsakligen mätts i förhållande till världsgenomsnittet för olika forskningsfält, så kallad fältnormaliserad citering, efter det att vi först tagit bort alla självciteringar.

I ett globalt perspektiv har citeringen av svenska vetenskapliga publikationer legat på en relativt konstant nivå, cirka 10% över världsgenomsnittet, sedan början av 1980-talet. Samtidigt har publikationer från flera andra europeiska länder blivit citerade i ökande grad. Jämfört med Europa-genomsnittet har därför svenska forskare förlorat i konkurrensen om att bli citerade. Andelen högt citerade svenska publikationer har också minskat, medan flera europeiska länder har ökat sina andelar av högt citerade verk. De områden där Sverige är framgångsrikt citeringsmässigt är teknik, natur-vetenskap samt skogs- och jordbruksrelaterade ämnen. Svenska publikationer inom medicin och samhällsvetenskap citeras i ungefär samma utsträckning som världsgenomsnittet.



Regeringsgatan 56 103 78 Stockholm Tel 08-546 44 000 Fax 08-546 44 180 vetenskapsradet@vr.se www.vr.se

Vetenskapsrådet är en statlig myndighet som utvecklar och finansierar grundforskning av högsta kvalitet inom alla vetenskapsområden. Vetenskapsrådet arbetar med forskningsfinansiering, strategi och analys samt forskningsinformation. Målet är att Sverige ska vara en ledande forskningsnation.

ISSN 1651-7350
ISBN 91-7307-095-5
