

Ritchie, Stuart, 2020. *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science*. London: The Bodley Head.

Anmälan av Anders Sundell

Sjuåringen tyckte att det var tråkigt att göra läxor, så jag satte ihop ett litet belöningssystem. Varje dag hon gör en läxa får hon ett klistermärke att sätta upp i en kalender. Det fungerade: nu gör hon dem gärna. Men bara en åt gången, mer vore slöseri. Bättre att sprida ut dem över dagarna för att maximera antalet klistermärken.

Incitament har effekt, på små såväl som stora. Det blir uppenbart när man läser Stuart Richies *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science*, en genomgång av hur de olika problem som räknas upp i titeln fördärvar vetenskapen.

Den bakomliggande orsaken är enligt Ritchie vetenskapssamhällets skeva incitament. Vetenskapen har professionaliserats och expanderat, med en ökande produktion av både forskare och vetenskapliga publikationer. Konkurrensen om tjänster och medel är hård, vilket ökar trycket på att samla karriärpoäng. Målet flyttas därmed, från sökande efter kunskap till sökande efter publikationer. Man kan säga att det viktiga inte är att resultatet är vetenskapligt, utan att det är *vetenskapslikt*.

Konsekvenserna är lika förutsägbara som tragiska. Den vetenskapliga revolutionen och den efterföljande tekniska utvecklingen är orsaken bakom det enorma uppsving i levnadsstandard mänskligheten upplevt de senaste århundradena. Förutsättningslöst kunskapssökande, hypoteser grundade i konsekventa och ofta matematiska teorier, prövade i väldokumenterade experiment, ledde till en ständigt växande kunskapsbank.

Nuförtiden växer dock kunskapen inte i samma takt som antalet publikationer. Studier från en mängd olika vetenskapsfält visar att en stor andel "forskningsresultat" inte visar sig hålla för närmare granskning, och inte kan replikeras av andra. Denna replikationskris har slagit hårdast mot psykologin (Ritchies eget ämne), men naturvetenskapen är inte undantagen – en uppmärksammas genomgång visade att 47 av 53 viktiga resultat inom cancerforskning inte gick att återskapa (Begley och Ellis 2012). Att statsvetenskapen än så länge inte drabbats i samma utsträckning beror antagligen mindre på att våra studier skulle hålla högre kvalitet än det faktum att de är svårare att replikera, eftersom de oftast bygger på observationsdata. Och även om boken i princip bara tar upp kvantitativ forskning tror jag inte heller att kvalitativa forskare skulle vara

mindre benägna att vinkla sina tolkningar för att passa in i ett mer övertygande narrativ. Det är bara svårare att kontrollera.

Ritchie inleder boken med att prata om fusk, och använder pinsamt nog för oss svenskar Macchiariniskandalen som främsta exempel. Och som statsvetare blir det extra genant att ett annat exempel är Lacour (och Green)-skandalen. Efter att en fältstudie misslyckats förfalskade statsvetardoktoranden Michael Lacour data, och fick tillsammans med Donald Green en artikel publicerad i *Science*. En enorm mediauppmärksamhet och ett jobb på Princeton följde. Bedrägeriet avslöjades av två andra doktorander, David Broockman och Joshua Kalla. När de själva senare faktiskt utförde en liknande studie (Broockman och Kalla 2016) visade det sig, ironiskt nog, att resultaten i princip var desamma som rapporterats av Lacour och Green!

En stor förtjänst med boken är emellertid att den inte fastnar i de här exemplen, hur sensationella de än må vara. Medvetet och utstuderat fusk är allvarligt, men inte det största problemet vetenskapen står inför. Istället är det antagligen mer vardagligt fusk och slarv som ligger bakom de flesta studier med opålitliga resultatet. I många fall handlar det troligen om att forskare lurar sig själva.

Drivkraften är även här att få fram publikationer, och eftersom tidskrifter tyvärr är mer benägna att publicera samband än ickesamband leder det till en jakt på statistisk signifikans. Så kallad "p-hacking" innebär att man mixtrar med de oräkneliga parametrar som finns i alla statistiska undersökningar – operationalisering av variabler, val av kontrollvariabler, estimeringsstrategi, hantering av outliers, interaktioner, med mera – för att komma på rätt sida av den magiska gränsen $p=0,05$. En kartläggning av p-värden i prestigetidskrifterna *American Political Science Review* och *American Journal of Political Science* visade till exempel att det är osannolikt få p-värden precis över gränsen, och osannolikt många precis under (Gerber och Malhotra 2008). Ett statistiskt mysterium, men inte alls oväntat för den som vet någonting om forsknings- och publiceringsprocessen. Mönstret är troligen resultatet av flera faktorer: forskarnas val i analysen, huruvida man färdigställer en studie eller inte, och tidskriftens beslut att publicera eller inte när ett manuskript väl skickats in.

Fenomenet p-hacking illustrerar väl målförskjutningen i modern vetenskap. $0,05$ är inte en naturligt magisk gräns. Sannolikheten att ett statistiskt samband är orsakat av slump i urvalet är i princip lika stor när p är $0,049$ som när p är $0,051$. Den stora skillnaden är att man i det förra fallet får sätta ut en stjärna vid koefficienten, och rätten att beteckna sambandet som *signifikant*. Värdet är symboliskt, inte substantiellt.

Till skillnad från andra böcker på liknande tema (till exempel *Return to Meaning – A Social Science with Something to Say* av Mats Alvesson, Yiannis Gabriel och Roland Paulsen 2017) är *Science Fictions* dock aldrig svepande i sin kritik. Ritchie utgår genomgående från empiriska studier, ofta metastudier,

som belägger problemen. Diskussionerna är också nyanserade, och visar på allvaret utan att slå över i sensationalism.

Man kan inte heller misstänka Ritchies för att klanka ner på ett system som behandlat honom illa. Sedan disputationen 2013 har han samlat ihop över 5000 citeringar på Google Scholar. Författaren klarar sig uppenbarligen bra i den dysfunktionella vetenskapsvärld som beskrivs. Boken ger dessutom inte uttryck för någon nostalgi över "den gamla goda tiden". Lösningen är inte att gå tillbaka till ett gentlemannaforskarideal, där kvalitetsgaranten är forskarens fina ideal och genuina kunskapssträvan. Medan Alvesson, Gabriel och Paulsen (2017) vill lösgöra sig från världsligheter som källhänvisningar och uppvärdera intellektualism vill Ritchie snarast gå åt andra hållet, mot större systematik, transparens och kontroll. Ett karriärsystem utan några mätbara kriterier för framgång skulle lämna stort utrymme för godtycke och nepotism.

Ett utmärkt exempel på vad som kan göras för att höja kvaliteten är krav på förregistrering av studier och tillhandahållande av material. Ritchie visar till exempel hur innehållet förändrades i en medicinsk tidskrift efter att förregistrering infördes. Resultaten i publicerade studier blev mindre sensationella, med många fler nollresultat, men ger antagligen en mycket mer representativ bild av forskningsfältet.

Men ökade krav på forskare bör också åtföljas av motsvarande skärpning från tidskriftshåll. En väg framåt kan vara modellen där artiklar antas på grundval av frågeställning, teori och metod. Resultaten avgörs i efterhand, och publiceras oavsett utfall.

Och även om replikationskrisen gett nödvändig krisinsikt ska vi inte tro att allt som publiceras är falskt. En nytt och ambitiöst replikationsförsök visade att resultat i studier med god metod – exempelvis hög statistisk styrka, förregistrering och stor transparens – också kunde återskapas i förväntad utsträckning (Protzko et al 2020).

Lösningen, tror jag, är att ta vetenskapen på större allvar, inte mindre. Vardagsfusk och slarv är mer lockande när man inte tror att någon kommer att använda ens forskningsresultat (om de ens blir lästa). När så är fallet blir vetenskapen till en hobby, eller ett spel, där det viktigaste är att vinna inom spelets ramar – få signifikanta resultat, fina publikationer – snarare än att närma sig en riktig sanning.

Vetenskapen ska i teorin vara mer rigorös och krävande än det allmänna tyckandet. I en krönika kan man skriva vad man tror, medan man i en vetenskaplig text behöver bevis. Men tanken är givetvis att även forskaren ska tro på sina teser – och jag är inte säker på att så alltid är fallet. Risken finns att det formella systemet med peer review slår ut det egna omdömet, att man intalar sig att resultat är rimliga bara de klarar hinderbanan som vetenskapssamhället satt upp. Problemet är att tidskrifternas granskare bara läser den färdiga texten. De ser inte hela processen, och kan därför inte åläggas hela ansvaret.

Om man tänker sig att forskningens slutmål inte är publikation, utan att kunskapen ska komma ut och användas i samhället, borde det bli mindre lockande att skicka in resultat som man själv tvivlar på. Att tvingas registrera sina studier på förhand är ett steg i rätt riktning, eftersom det underlättar kollegial granskning. Men en starkare koppling mellan forskning och undervisning och tydligare krav på att kommunicera resultat till allmänheten skulle också hjälpa. Forskare tvingas då att tydligare stå för sina resultat, inför kollegor, studenter och allmänheten. Gemene man kanske inte kan skilja på en Regression Discontinuity Design och en Time Series Cross Section, men har å andra sidan mindre tolerans för akademiskt navelskåderi. Om det nu är sanningen vi gör anspråk på att avtäcka i forskningen borde den rimligen hålla för kritisk granskning från alla håll.

För den som har följt debatten om replikationskrisen någorlunda erbjuder Stuart Ritchies bok egentligen inget nytt, men är ändå en gedigen och nyanserad översikt från en som vill vetenskapen väl. Den borde vara obligatorisk läsning på alla forskarutbildningar. Vetenskapen är för viktig för att reduceras till en jakt på klistermärken.

REFERENSER

- Alvesson, Mats, Yiannis Gabriel & Roland Paulsen, 2017. *Return to Meaning. A Social Science with Something to Say*. Oxford: Oxford University Press.
- Begley, Glenn & Lee Ellis, 2012. "Raise standards for preclinical cancer research", *Nature*, 483: 531-533.
- Broockman, David & Joshua Kalla, 2016. "Durably reducing transphobia: A field experiment on door-to-door canvassing", *Science*, 352: 220-224.
- Gerber, Alan & Neil Malhotra, 2008. "Do statistical reporting standards affect what is published? Publication bias in two leading political science journals", *Quarterly Journal of Political Science*, 3: 313-326.
- Protzko, John, Jon Krosnick, Leif D. Nelson, Brian A. Nosek, Jordan Axt, Matthew Berent, Nick Buttrick, Matthew DeBell, Charles Ebersole, Sebastian Lundmark, Bo MacInnis, Michael O'Donnell, Hannah Perfecto, James Pustejovsky, Scott Roeder, Jan Walleczek, Jonathan Schooler, 2020. "High Replicability of Newly-Discovered Social-behavioral Findings is Achievable".

Prenumerera på Statsvetenskaplig tidskrift!
www.statsvetenskapligtidskrift.org