Taking census of physics

Federico Battiston, Federico Musciotto, Dashun Wang, Albert-László Barabási, Michael Szell and Roberta Sinatra

Abstract | Over the past decades, the diversity of areas explored by physicists has exploded, encompassing new topics from biophysics and chemical physics to network science. However, it is unclear how these new subfields emerged from the traditional subject areas and how physicists explore them. To map out the evolution of physics subfields, here, we take an intellectual census of physics by studying physicists' careers. We use a large-scale publication data set, identify the subfields of 135,877 physicists and quantify their heterogeneous birth, growth and migration patterns among research areas. We find that the majority of physicists began their careers in only three subfields, branching out to other areas at later career stages, with different rates and transition times. Furthermore, we analyse the productivity, impact and team sizes across different subfields, finding drastic changes attributable to the recent rise in large-scale collaborations. This detailed, longitudinal census of physics can inform resource allocation policies and provide students, editors and scientists with a broader view of the field's internal dynamics.

There was a time when polymaths such as Galileo Galilei knew all the physics that could be known. Over the centuries, however, the body of knowledge spanned by physics has exploded, encompassing topics as diverse as biophysics, chemical physics and network science. As physics expanded in breadth and depth, physicists were forced to specialize in narrow subfields[1]. The evolving complexity of the field prompts questions such as the following: as a physics student choosing my future specialty, how do I know which subfields are growing; as a funding agency, how do I compare early-career physicists from different subfields; and as a journal editor, how many papers should I expect from each subfield and how do I compare their impact?

So far, there have only been anecdotal or case-by-case answers to such questions, but the recent availability of large data sets of scientific publications and the development of network science techniques offer the opportunity to tackle these questions systematically by exploring the production patterns of scientists[2,3]. Here, we take an intellectual census of physicists and their career trajectories and offer data-driven answers to these questions. We map out the evolution of the subfields of physics and gain quantitative insights relevant to fundamental scientific processes, such as resource allocation to the exchange of knowledge, revealing quantitative footprints not just for physics but also for its intimate relation with the broader scientific community[4,5].

A census of the physics subfields

To take a census of physics, we first identify the relevant physics papers and citations within the Web of Science (WoS). We start by selecting ~3.2 million physics papers, published in 294 physics journals indexed by the WoS. This core represents, however, only a fraction of all physics papers[5,6], missing, for example, those published in interdisciplinary journals such as *Nature* or *Science* or papers published in journals of other disciplines but that are of direct relevance to the physics community. To map out the complete physics literature, we then set out to detect physics papers among the other ~47 million papers in the WoS by virtue of their citation patterns. A paper is a potential physics publication if the number of its references and citations to the core physics literature is significantly higher than the number in a null model in which each paper's citations are assigned randomly, regardless of a paper's journal or research area. We identified ~4.5 million papers whose patterns of citations and references are indistinguishable from papers in physics journals (see Supplementary Section 1), obtaining a data set of ~7.7 million publications of interest to the physics community.

We further characterize this physics corpus by classifying each paper into one of nine major subfields, following the Physics and Astronomy Classification Scheme (PACS) developed by the American Institute of Physics[7] and used by several journals, including the Physical Review series of the American Physical Society (APS) between 1985 and 2015. This time frame reduces the set of considered papers to ~5.6 million. We combine information from the WoS and the APS[8] data sets to reconstruct the publication profile of 135,877 physicists with at least five papers between 1985 and 2015. For more details on the data set curation and validation, refer to BOX 1 and Supplementary Section 3.

The first step in developing a census is to count the number of physicists working in each subfield. Such counting is, however, not straightforward, as physicists may contribute to publications in different subfields. We therefore associate each physicist with a primary subfield if the number of her publications in the subfield is higher, in a statistically significant manner, than expected for a typical physicist (BOX 1 and Supplementary Section 4). The obtained subfield demographics offer us a first overview (FIG. 1a). We find that the largest subfield is CondMat (condensed matter physics), with more than 62,000 physicists, capturing 46% of the entire physicist population. It is followed by General (34,000), HEP (high-energy physics, 33,000), Interdisc (interdisciplinary physics, 32,000), Classical (28,000), Nuclear (24,000), AMO (atomic and molecular physics, 20,000) and Astro (19,000). Plasma is the smallest subfield of physics, with less than 11,000 researchers.

Given the highly specialized nature of the physics subfields, one might suspect that most physicists work in a single subfield. However, we find that specialized physicists are the exception rather than the rule. The majority of physicists (63%) are active in two or more subfields (FIG. 1b). This scenario prompts us to ask which subfields have particularly low or high rates of specialization. The differences between subfields are striking, defining two groups (FIG. 1c): six subfields have less than 10% specialized physicists. Among these subfields, Interdisc has less than 1% of specialized physicists, in line with the expectation that interdisciplinary physicists bridge multiple subfields. By contrast, the percentages of specialized physicists in CondMat, HEP and Nuclear are 42%, 34% and 25%, respectively, at least an order of magnitude larger than the percentage in the other group of subfields, which raises the question of what drives the different levels of specialization in different subfields.

A physicist working in two or more subfields combines the collective know-how of these fields, a process deemed essential for new discoveries in science[9–11]. To understand which of the physics subfields cross-pollinate most significantly, we calculate the co-activities of individual physicists between each pair of subfields. Co-activities are defined by weighted links between subfields, in which the weights measure the observed co-activities compared with those expected in a randomized null model (Supplementary Section 6). Starting with the highest-weighted links, we plot the minimum number of links needed to have a connected network of subfields (FIG. 1d). The network reveals a nontrivial co-activity structure, clustering all physics subfields into three broader areas, all held together by General: Interdisc and CondMat; Classical, AMO and Plasma; and HEP, Astro and Nuclear.

This research space captures the intellectual affinities between subfields, which facilitates movements between close subfields and limits cross-pollination between distant ones such as Interdisc and Nuclear[12–15]. For example, the diversity of topics within CondMat and Classical and their adaptable approaches, such as statistical mechanics applied to multiple systems composed of large numbers of entities, suggests that it would be easier for those working in these subfields to take their tools to different disciplines. By contrast, more specialized subfields such as HEP or Nuclear require their members to acquire familiarity with large-scale, long-term projects. Whereas scientists working in such fields may have deep knowledge and expertise on the subject they specialize in, they may face a greater burden that limits their ability to explore other areas. The observed co-activity network is similar to the citation network[5] between subfields, showing that the flow of knowledge is captured by both metrics[15]. We checked that significant author co-activity

Box 1 | **Methodology**

Identifying subfields

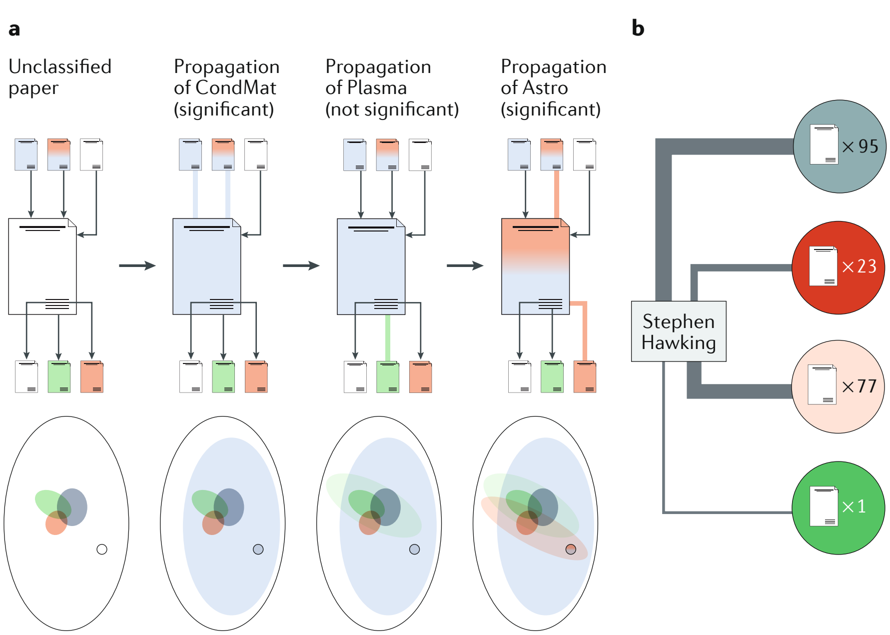
We classify papers into nine subfields on the basis of the 1-digit Physics and Astronomy Classification Scheme (PACS) by the American Institute of Physics[7]:

- General: Mathematical Methods, Quantum Mechanics, Relativity, Nonlinear Dynamics and Metrology
- HEP (high-energy physics): Physics of Elementary Particles and Fields
- Nuclear: Nuclear Structure and Reactions
- AMO: Atomic and Molecular Physics
- Classical: Electromagnetism, Optics, Acoustics, Heat Transfer, Classical Mechanics and Fluid Dynamics
- Plasma: Physics of Gases, Plasmas and Electric Discharges
- CondMat (condensed matter physics): Structural, Mechanical and Thermal Properties, Electronic Structure and Electrical, Magnetic and Optical Properties
- Interdisc: Interdisciplinary Physics and Related Areas of Science and Technology
- Astro: Astrophysics, Astronomy and Geophysics

PACS codes were consistently used in 435,772 papers published in the journals of the American Physical Society (APS) between 1985 and 2015 (Supplementary Section 2). Using an algorithm that evaluates the patterns of citations and references among papers, we propagate subfield labels from APS papers to other papers: if the fraction of references and citations between a given paper and papers in a particular subfield is larger than expected by the null model, the paper is assigned to that subfield. A paper may be assigned to multiple subfields, in line with APS papers reporting multiple PACS codes. As a hypothetical example, we consider a paper that references papers in CondMat, Plasma and Astro and that is cited by CondMat, Astro and another publication still lacking a PACS code (see the figure, part **a**). The publication is first assigned to CondMat and then to Astro but not to Plasma, because it lacks statistically significant links to the subfield. The algorithm is run iteratively until convergence for each subfield, helping us associate at least one subfield to 1,137,670 papers (Supplementary Section 3).

Assigning physicists to subfields

We analyse all careers with at least five labelled papers between 1985 and 2015, capturing the careers of 135,877 physicists. We consider a physicist to be working in a subfield if their share of publications in the subfield is higher than that of the average physicist. The statistical criterion[51] we use guarantees that each scientist is assigned to at least one subfield and takes into account the different sizes of subfields. As an example, we show the result of the criterion applied to the career of Stephen Hawking (see the figure, part **b**). In the physics data set, Hawking has 124 papers associated with one or more subfields. Of these subfields, only General (95 papers) and Astro (77 papers) are assigned to the physicist through the statistical criterion, whereas HEP (23 papers) and Classical (1 paper) are not statistically significant, which is consistent with Hawking being known as a theoretical physicist and cosmologist.



For validation and further methods, see Supplementary Sections 3,4,5.

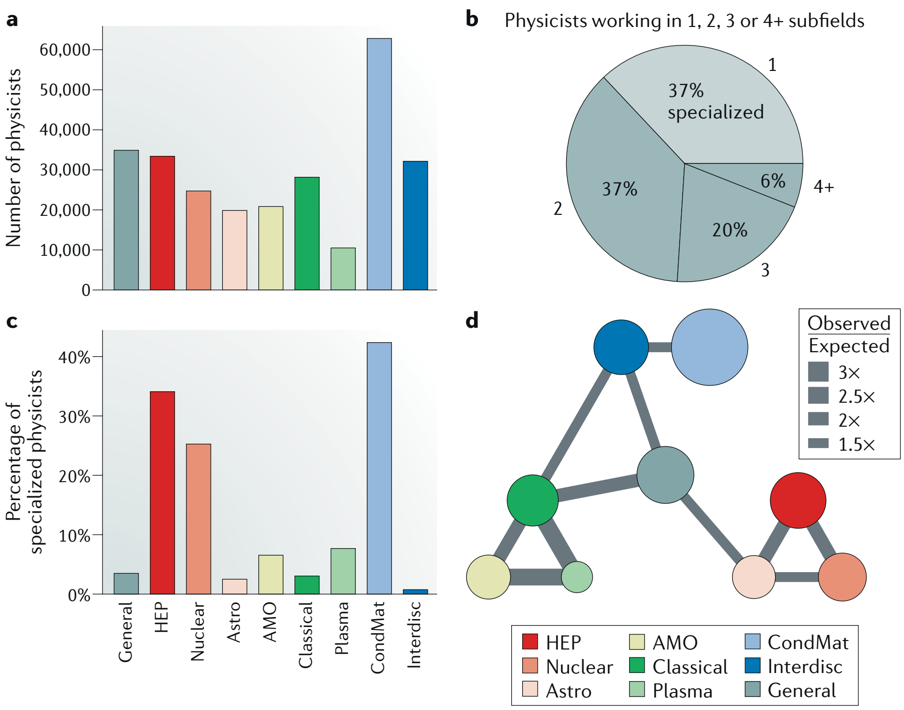


Fig. 1 | **Census of the physics subfields. a** | Number of physicists per subfield. **b** | Percentage of physicists working in 1, 2, 3 or 4+ subfields. We call the 37% of physicists who work in only one subfield specialized. **c** | Percentage of specialized physicists per subfield. Most subfields, except for HEP (high-energy physics), Nuclear and CondMat (condensed matter physics), have a negligible percentage of specialized physicists. **d** | The network of co-activity of individual physicists shows the non-trivial connections between subfields. The node size is proportional to the number of physicists in the subfield, and the link width is proportional to the overlap between subfields, quantified by the ratio between the measured number of physicists working in the two subfields and the expected number based on a randomized null model. AMO, atomic and molecular physics; Interdisc, interdisciplinary physics.

across different subfields is in general a correct observation and not an artefact of our subfield identification method (Supplementary Section 7).

Birth, growth and migration

One may wonder why there are considerable differences in the numbers of specialized physicists between similarly sized subfields, such as Nuclear and Interdisc (FIG. 1a,b). To understand the roots of this heterogeneity, we first assess the relative growth rate of each subfield over time, measuring the fraction of physicists entering a subfield every year (FIG. 2a). We find that the growth rates of Interdisc and Astro increased from a few percent in 1985 to over 20% and 27%, respectively, after 2010. An opposite trend characterizes CondMat: whereas it had the largest share of new physicists in 1985, its share dramatically decreased over time, falling below 5% after 2010. HEP also displayed a receding trend just before 2010.

The reasons behind such substantial changes in the landscape of physics may be traced back to a handful of far-reaching sociopolitical developments in the United States, which had the largest share of physics output in the second half of the 20th century. Most importantly, at the end of the 1980s, CondMat, Nuclear and HEP represented the spearhead of the strong relationship between academia and the Department of Defense in the United States, which was funding up to 80% of the R&D budget during the whole Reagan era[16]. The growth of these subfields during that period was so pronounced that it was sometimes described as a 'speculative bubble' within the academic world[17]. The marked importance of CondMat was the culmination of a conceptual shift that had put condensed matter research in a dominant position with respect to more traditional subfields[18]. At the time, a similar funding focus also prevailed in the Soviet Union, the opposing superpower. The abrupt end of the Cold War era at the beginning of the 1990s quickly reshaped the geopolitical priorities of the major national powers and, subsequently, the relative importance of subfields of physics — the United States Congress vote that cancelled funding for the Superconducting Supercollider in October 1993 provides anecdotal evidence of this trend. The only major exception might be the spur of new research connected to the activity of the Large Hadron Collider in Geneva, which injected new life into HEP. In particular, the sharp peak in 2010 can be attributed to the first ATLAS and CMS publications[19] (Supplementary Section 9).

The number of physicists entering a subfield every year (FIG. 2a) includes both physicists who started their careers in a particular subfield and those who made career transitions from other subfields. There are remarkable examples of physicists who never changed their subfield, such as Klaus von Klitzing, whose first publication was in CondMat and who contributed over 500 papers to the subfield, earning him the Nobel Prize in 1985 for the discovery of the quantized Hall effect. By contrast, Rainer Weiss, best known for inventing the laser interferometric technique at the heart of the Laser Interferometer Gravitational-Wave Observatory (LIGO), which earned him the Nobel Prize in 2017, published his first paper on a rather unrelated topic in AMO. To distinguish such different careers, we systematically explore career transitions within physics[20], asking where physicists are 'born', how they migrate between subfields and when these transitions typically occur.

Remarkably, 64% of the physicists began their careers by publishing in either CondMat (37% of all physicists), HEP or Nuclear (FIG. 2b). These percentages are in agreement with the prominence of these three subfields at the beginning of the investigated period. General, covering topics of interest to a wide set of physicists, accounts for 14% of first publications. By contrast, only 4% of physicists started publishing in Interdisc, and as low as 3% began in Astro. Because Interdisc integrates other disciplines, it might be difficult to start out as an Interdisc physicist. The low percentage of career starts in Astro may be a result of the connections it has recently developed with HEP and Nuclear[21] and to the limited coverage of the Astrophysics and Astronomy literature in our corpus (Supplementary Section 3).

We measure the significant flows between subfields by comparing the subfield in which a physicist published their first paper with the subfields that best characterized their later career (Supplementary Section 6; FIG. 2b). CondMat is the starting point for many physicists who later specialized in Interdisc, Classical and General. HEP and Nuclear tend to swap researchers while feeding talents into Astro, a pattern that may be rooted in the fact that all three subfields study radiation

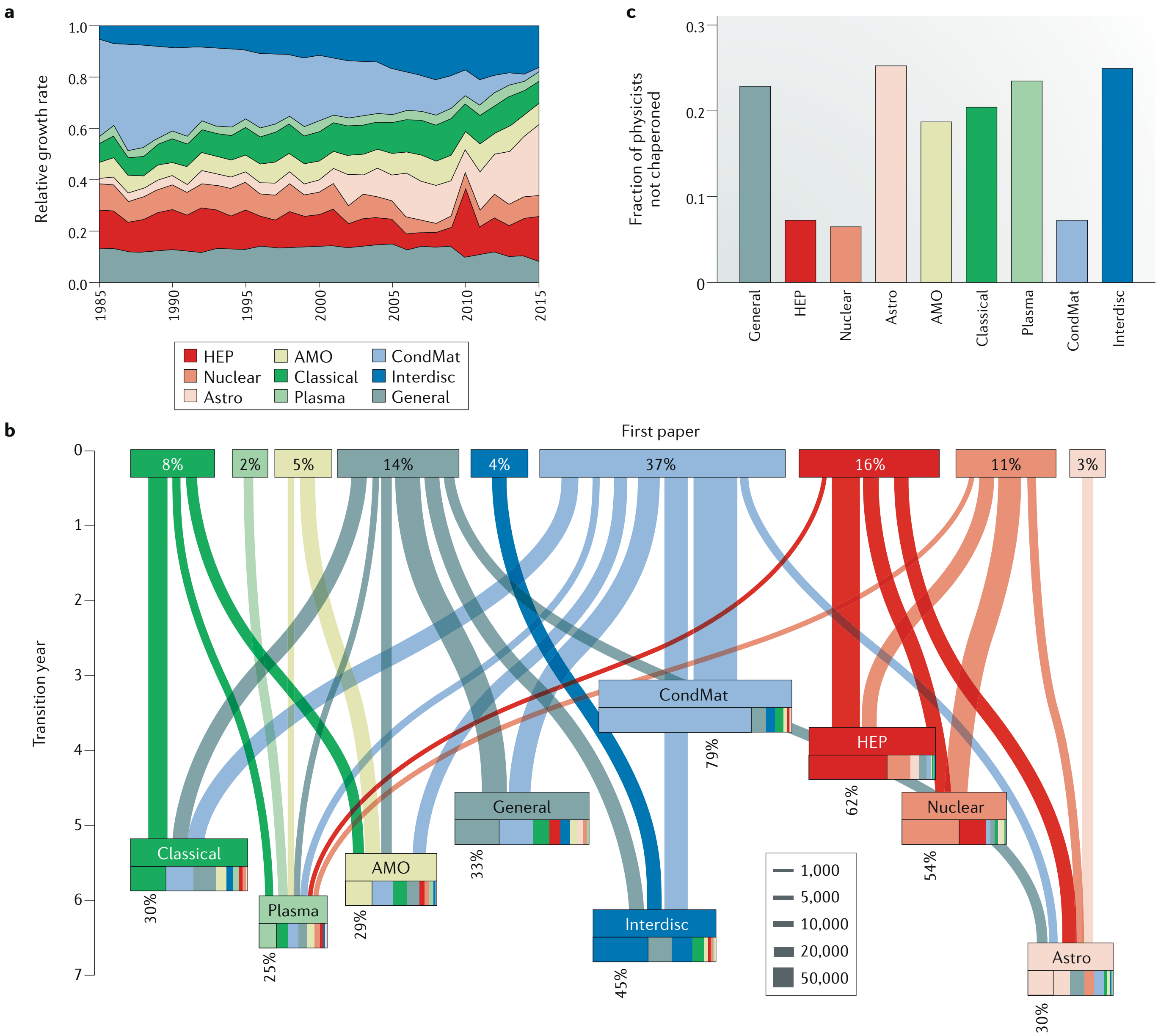


Fig. 2 | **Evolution of physics subfields and careers. a** | Relative growth rate, defined as the yearly fraction of physicists who published their first paper in a subfield new to them. Interdisc (interdisciplinary physics) and Astro grow, while CondMat (condensed matter physics) shrinks considerably. HEP (high-energy physics) displays a spike in 2010 that can be attributed to large-scale collaborations such as ATLAS and CMS (Supplementary Section 9). The relative growth rate is less reliable after 2010 owing to early-career physicists accumulating publications at different rates in each subfield; thus, they reach the five-publication threshold at different times. This outcome distorts the proportion of physicists in favour of more productive and non-specialized subfields. **b** | Fraction of unchaperoned physicists in each subfield. A large majority of physicists starting in HEP, Nuclear or CondMat co-author their first paper with physicists who have already published in the subfield. Other subfields have a higher fraction of physicists who are not chaperoned. **c** | Flow diagram of career transitions. The sizes of the upper rectangles are proportional to the number of career first publications in each subfield. The lower rectangles are proportional to the number of physicists active in each subfield. For example, Astro and AMO (atomic and molecular physics) have roughly the same number of physicists, although Astro starts with 3%, while AMO starts with 5%. The distance from the top reflects the average time at which a career transition occurs. Flows describe the flow of physicists from their first subfield towards the one that characterizes their later body of work. Only significant flows, that is, those larger than expected in the null model, are shown. The percentages adjacent to the lower rectangles report the contribution of the subfield that is contributing most.

or nuclear and subnuclear processes. We find that most Interdisc physicists did not start their career there but migrated from CondMat and General, consistent with the hypothesis that interdisciplinary research builds on disciplinary expertise[10]. Finally, Plasma and Astro welcome physicists with many different backgrounds but rarely feed into other subfields. The diversity of the incoming flows to Plasma and Astro suggests their accessibility to physicists with many different backgrounds.

We also measure the average time it takes to transition to a different subfield (FIG. 2b). Once again, HEP, Nuclear and CondMat top the list: physicists who did not start their career in these subfields tend to transition towards them the earliest, typically by the third or fourth year of their research career. The opposite trend is observed for Interdisc and Astro, which not only have the highest transition rates among subfields but are also characterized by the longest time to transition. Indeed, on average, a physicist publishes their first paper on these two topics 6–7 years into their career, roughly double the transition time towards HEP, Nuclear and CondMat. Interdisc displays a late switch, consistent with the hypothesis that it takes time to gather expertise in multiple fields. Similarly, physicists tend to switch to Astro after a relatively long experience in HEP.

Analysing the flows between subfields (FIG. 2b) helps us better understand the network of co-activity of individual physicists (FIG. 1d). For instance, in the HEP–Nuclear–Astro triple, HEP plays the leading role in producing physicists who transition to its tightly connected subfields, that is, Nuclear and Astro. Likewise, CondMat is the main source feeding Interdisc. The observed widespread career transitions may reflect potential benefits to the whole field, cross-pollinating one physics community with ideas and methods developed by a different subfield[9,10].

The role of chaperones

The future success of young scholars has often been linked to access to valuable mentorship in the early stages of a scientific career[22–24]. For example, a surprising fraction of Nobel laureates had a mentor–mentee or a co-authorship relationship with another Nobel laureate[25,26], and scientists who co-author early with an established scientist are more likely to have higher impact and more chances to publish as a lead author than other scientists[27]. Taken together, a senior scientist who acts as a chaperone during a scientist's early career might foster the acquisition of skills, passing on the experience and knowledge necessary to attain high achievements later in a career.

To quantify the chaperone effect, we measure how many physicists co-author their first paper in a subfield with a physicist who has previously published in that subfield[27]. We find that the chaperone effect is particularly strong for HEP, Nuclear and CondMat, in which over 90% of physicists wrote their first paper with someone who published before in the same subfield (FIG. 2c and Supplementary Section 10). This large share of chaperoned physicists could have several explanations, such as the documented high number of physicists starting their career in these three subfields or the need to access large facilities, which require early-career physicists to collaborate with established scientists. Note that the typical large co-authorship patterns of HEP cannot explain the magnitude of the chaperone effect characterizing this subfield (Supplementary Section 10).

Other subfields have a lower fraction of chaperoned physicists, especially Interdisc and Astro. These subfields are often explored by more senior physicists who received mentorship at a previous stage of their careers in a different subfield and who decided to explore the new area without close supervision (26% of physicists are not chaperoned in both Interdisc and Astro; FIG. 2c). In addition, applications of computational physics, such as computational biophysics or complex systems, classified as Interdisc, require fewer financial resources than experimental research. This lower financial barrier to entry could also play an important role in explaining the low chaperone effect[28]. Overall, the chaperone effect is strong in physics, with an average rate across subfields of 82% of physicists being chaperoned. The effect signals a research culture in which physicists are often introduced to their future research area by senior colleagues in a collaborative setting, in contrast with disciplines such as mathematics, in which the majority of scientists start their career by publishing single-author papers[27].

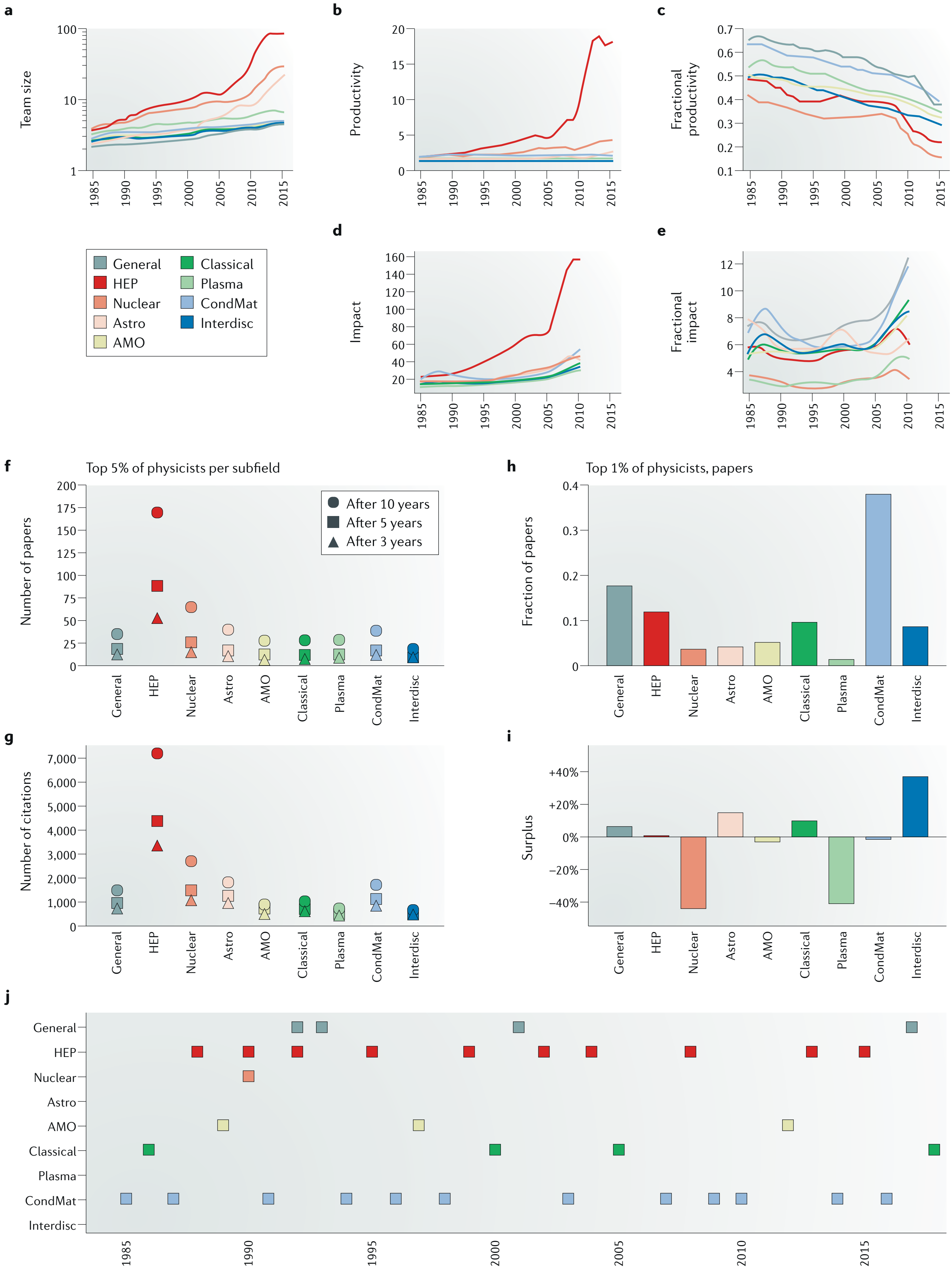
Productivity, impact and team size

Productivity and impact, as measured by the number of papers published and citations received by a physicist, are frequently used metrics in the assessment of scientific careers[29,30]. These quantities have implications for decisions and policies involving predicting, nurturing and funding early-career scientists. However, the proper interpretation of these metrics must account for the highly heterogeneous productivity and citation patterns characterizing different subfields[31] and for different team sizes[32], both of which vary in time.

Team size — the number of co-authors per paper — has been increasing steadily over the past decades in all fields, indicating an increase in collaboration in science[33]. We ask whether there are particular differences in collaborative patterns in the different physics subfields and what their implications are for productivity and impact. To answer this question, we assess the diversity and evolution of collaboration, productivity and citation standards in the different subfields of physics. First, the tendency of scientists to work in increasingly large teams has been particularly pronounced in HEP (especially after 2005), Nuclear (especially after 2010) and Astro (especially after 2000) (FIG. 3a). The observed explosive growth in these three subfields is partly rooted in large-scale projects such as ATLAS (Supplementary Section 9). These large collaborations also result in an increase in productivity: as physicists became involved in more and larger teams, the average number of papers they published each year increased by a factor of ten for HEP and by a factor of two for Nuclear and Astro from 1985 to 2015 (FIG. 3b). However, for the other six subfields, productivity has stayed constant over 30 years, and for all subfields, productivity has increased at a slower rate than team sizes. These different rates of increase explain why fractional productivity, that is, the ratio between the number of papers and the average team size, decreased across all subfields (FIG. 3c). The effect is the strongest in HEP, Nuclear and Astro, in which team size grew disproportionately. In these subfields, authors are usually ordered alphabetically owing to the large average team size, making the assessment of credits for individual authors more problematic[34]. Taken together, we find that the amount of knowledge produced per capita has decreased in all subfields despite the increase in the total number of physicists and physics papers.

Given the explosive increase in both team size and the number of papers per physicist in HEP, we wonder whether HEP physicists today have more or less impact than they had decades earlier. To answer this question, we measured the average impact as the number of citations after 5 years (FIG. 3d) and the fractional impact (ratio between the number of citations and the average team size; FIG. 3e) per physicist per subfield. Interestingly, the average impact of HEP shows a growth of comparable magnitude to the growth in average productivity, leading to an unchanged fractional impact. In other words, large-scale projects such as ATLAS produce papers that generate a large number of citations, compensating for the large numbers of co-authors (hundreds or more).

Our classification of subfields is unable to capture differences between authors at finer scales. For instance, in HEP, two different sub-communities of theorists and experimentalists coexist and have different production patterns. We adopted the following simple a posteriori heuristic to roughly separate these two groups: all papers authored by less than ten scholars were assigned to the theoretical group, whereas all the others were considered as experimental. The two subcategories follow rather different patterns of productivity and impact (Supplementary Fig. 10). On the one hand, the overall productivity and impact


a
Team size
b
Productivity
c
Fractional productivity
General
HEP
Nuclear
Astro
AMO
Classical
Plasma
CondMat
Interdisc
d
Impact
e
Fractional impact
f
Top 5% of physicists per subfield
Number of papers
After 10 years
After 5 years
After 3 years
h
Top 1% of physicists, papers
Fraction of papers
g
Number of citations
i
Surplus
j

◀ Fig. 3 | **Productivity and impact across physics communities. a** | Average team size, defined as the average number of authors per paper, over time. Team sizes grow in all fields, especially in HEP (high-energy physics), Nuclear and Astro owing to large-scale experimental projects. **b** | Average productivity, defined as the number of papers per author, over time. Productivity grows for HEP, Nuclear and Astro but stays roughly constant for other subfields. **c** | Fractional productivity, that is, the number of papers divided by team size, over time. For all subfields, productivity grows slower than team size; hence, fractional productivity decreases. **d** | Average impact, defined as the number of citations per author within a 5-year window. Impact increases in all fields, but only HEP shows exceptional growth. **e** | Fractional impact, the number of paper citations divided by team size, over time. Most subfields show a roughly constant trend until 2005. **f** | Number of papers of the top 5% of physicists for productivity at different career stages. Owing to different collaboration standards, HEP physicists co-author more papers than those in other subfields. Interdisc (interdisciplinary physics) physicists produce an especially low number of papers. **g** | Number of citations of the top 5% of physicists for productivity at different career stages. HEP physicists receive more citations because of their high productivity. **h** | Fraction of top 1% of cited papers per subfield. **i** | Subfield surplus of most-cited papers with respect to the number expected given the subfield size. Interdisc generates the highest number of high-impact papers compared with its size. **j** | Nobel Prizes in physics per year across subfields. Plasma and Interdisc have not received an award. AMO, atomic and molecular physics; CondMat, condensed matter physics.

behaviour of HEP physicists are dominated by the experimental part of that community. On the other hand, fractional productivity and impact show a reversed picture: they are dominated by theorists, indicating the importance of taking into account team size when comparing individual output within the same subfield.

Given some of the large productivity differences between different subfields, we also expect differences in impact[35–38], measured in terms of cumulative citations over a career. For instance, we ask the following: how much impact does it take to be a scientific leader in HEP, and how is that different in CondMat? An answer to this comes from the total number of papers and citations acquired over an average career by the top 5% (in terms of productivity) of physicists in each subfield at different career stages (FIG. 3f,g). For both papers and citations, HEP is by far the most rewarding subfield, the top scientists of which co-author 169 papers and accumulate over 7,000 citations 10 years on from the beginning of their career. By contrast, top Interdisc physicists co-author only 18 papers, with less than 1,000 citations. The large discrepancy is not explained by paper citation rates[39,40], which are roughly constant across subfields (Supplementary Section 11), but by the high or low number of papers per author in the respective subfield (FIG. 3b). As a consequence, when physicists with different specialties compete for positions or grants, caution is needed in comparing their profiles using metrics based on citations or productivity, as subfield-dependent differences appear from the very beginning of a career.

One can also wonder how top papers are distributed across the different subfields. We selected the top 1% of all physics papers (in terms of citations) and assessed into which subfield they fall (FIG. 3h). The majority fall into CondMat, General and HEP; however, this result is trivial, as these fields produce the most papers. To unveil significant effects, we measured the surplus between this top 1% distribution and the distribution of subfields of all physics papers. Interdisc papers are approximately 40% more likely to be in the top 1% than expected, whereas Nuclear and Plasma papers are 40% less likely to be found in the top 1% (FIG. 3i). The high prevalence of Interdisc among the highest-cited papers might be partially explained by the finding that papers that are 15% novel and 85% conventional often have high impact[10]. Interdisc is more likely to achieve this balance, as interdisciplinary research is often novel[10] and, at the same time, must adhere to established principles. Another explanation is that Interdisc physicists are more likely to initiate new topics or emerging subfields. Papers that open such new avenues are known to acquire a high number of citations as they become milestones, cited by subsequent papers once the field is established[41,42].

Recognition of physics subfields

We ask whether differences in impact affect the way in which the overall scientific community perceives the different subfields of physics. As a rough proxy of this recognition, we take the Nobel Prizes awarded from 1985 to the present, highlighting each awarded subfield (FIG. 3j, Supplementary Section 12). Although the Nobel Prize often recognizes research undertaken well before the selection year, the timing of Nobel Prize selections could be affected by the relative importance of different physics communities as perceived by the committee. We find that Nobel Prizes are not related to the number of physicists flocking into specific physics communities, nor do they show significant temporal clusters (compare FIG. 2a and FIG. 3j). However, the general distribution of awarded subfields reveals interesting tendencies: a large fraction of Nobel Prizes have been awarded to 'curricular' topics, such as CondMat, the subfield with the largest number of active researchers, and HEP. Surprisingly, Astro, despite the moderate size of its community, comes in third, with five Nobel Prizes. This success might be linked to the general excitement regarding the grand questions the field addresses, as well as to its strong ties to HEP, a regular recipient of Nobel Prizes. Other well-established areas with a long history, such as AMO and Classical, have also been recognized. By contrast, since 1985, Plasma and Interdisc have not been awarded a Nobel Prize. The omission of Interdisc likely comes from the charter of the Nobel Prize to award clearcut categories (such as physics, chemistry and medicine/physiology), rooted in the 19th century's strict classification of disciplines. This requirement counts against interdisciplinary discoveries[43,44].

Discussion and outlook

Our approach has some limitations. Our study is based on WoS data and lacks the literature that has been exclusively published on preprint servers such as arXiv[45]. This limitation leads to unavoidable differences in subfield representation, owing to diverse publication cultures in different communities. For example, the proportion of HEP and Astro papers on arXiv is higher than that of our data set and the WoS, reflecting the common practice of these communities to communicate findings in preprints rather than journal papers. However, there is high overlap in the coverage of the physics literature between different databases[46] and a high correlation of the representation of physics subfields (Supplementary Section 3), indicating that our findings should not change if our analysis is repeated on a different database.

Furthermore, the analysis has an inherent limitation that stems from the use of the PACS classification by the APS. Indeed, the Physical Review series focuses mostly on pure physics, as opposed to more applied areas such as acoustics, meteorology, crystallography and so on, which are published mainly in specialized applied physics journals. Therefore, stemming from APS journals, our analysis cannot effectively represent the more applied areas of physics, as they might not be sufficiently cited nor referenced in our data set. An extended investigation, for example, through the adoption of a seed data set more comprehensive than the Physical Review series, would be worthwhile. Furthermore, today, many scientists with a physics background contribute to fields outside of physics, from biology to finance, both in academia and the private sectors[47]. For this reason, the analysis of the connections between physics and other scientific disciplines and of the career transitions outside physics remains a fruitful future research direction. Indeed, such an investigation, possibly complemented by data sources that go beyond scientific publications, could shed light on the role of physics and its subfields in the entire ecosystem of science and beyond.

A question that has not been addressed in this Perspective pertains to the role of geographical variations in the evolution of physics subfields. Indeed, the production of science and its impact vary across countries[48,49]. Adding this geographical layer to our analysis through affiliation data[4,50] could enhance our ability to interpret the observed migrations and mobility patterns within physics.

Conclusions

Our survey of the physics landscape shows that subfields rarely live in isolation but rather tend to overlap, with individual scientists working in multiple subfields and transitioning between fields during their career. Mapping these overlaps reveals a highly nontrivial research space, displaying deep intellectual links between some subfields and large gaps between others. Physicists who face the problem of resource allocation to different subfields and departments often use metrics of productivity or impact to seek priority. However, our research suggests that such arguments should be taken with scepticism. Indeed, there are considerable field-specific differences in the patterns of productivity and impact. Publication rates have exploded in recent years in HEP, Nuclear and Astro, whereas fractional productivity is declining. In some subfields, such as HEP, researchers co-author an exceptionally large number of papers, partly rooted in their unique culture of collaboration. By contrast, interdisciplinary physicists produce papers at a much lower rate, but their papers tend to generate a disproportionately higher impact once we factor in the relative size of the subfield. Recognizing these field differences within physics represents the first step towards a deeper understanding of our discipline. As tomorrow's physicists working on different topics compete for the same positions and resources, these insights may prove pertinent for the sustainable vitality of physics as a discipline.

*Federico Battiston[1,9], Federico Musciotto[1,9], Dashun Wang[2,3], Albert-László Barabási[1,4,5], Michael Szell[1,4,6] and Roberta Sinatra[1,4,7,8]**

[1]Department of Network and Data Science, Central European University, Budapest, Hungary.

[2]Kellogg School of Management, Northwestern University, Evanston, IL, USA.

[3]Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL, USA.

[4]Network Science Institute, Northeastern University, Boston, MA, USA.

[5]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, MA, USA.

[6]MTA KRTK Agglomeration and Social Networks Lendulet Research Group, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budapest, Hungary.

[7]Department of Mathematics, Central European University, Budapest, Hungary.

[8]ISI Foundation, Torino, Italy.

[9]These authors contributed equally: Federico Battiston, Federico Musciotto

**e-mail: robertasinatra@gmail.com*

https://doi.org/10.1038/s42254-018-0005-3
Published online 8 January 2019

1. Jones, B. F. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Rev. Econ. Stud.* **76**, 283–317 (2009).
2. Clauset, A., Larremore, D. B. & Sinatra, R. Data-driven predictions in the science of science. *Science* **355**, 477–480 (2017).
3. Fortunato, S. et al. Science of science. *Science* **359**, eaao0185 (2018).
4. Deville, P. et al. Career on the move: geography, stratification, and scientific impact. *Sci. Rep.* **4**, 4770 (2014).
5. Sinatra, R., Deville, P., Szell, M., Wang, D. & Barabási, A.-L. A century of physics. *Nat. Phys.* **11**, 791 (2015).
6. Deville, P. *Understanding social dynamics through big data.* Thesis, Univ. Catholique Louvain (2015).
7. AIP Publishing. PACS 2010 regular edition. *AIP* https://publishing.aip.org/publishing/pacs/pacs-2010-regular-edition (2018).
8. APS Physics. APS data sets for research. *APS* https://journals.aps.org/datasets (2018).
9. Dyson, F. Birds and frogs. *Not. AMS* **56**, 212–223 (2009).
10. Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. Atypical combinations and scientific impact. *Science* **342**, 468–472 (2013).
11. Foster, J. G., Rzhetsky, A. & Evans, J. A. Tradition and innovation in scientists' research strategies. *Am. Sociol. Rev.* **80**, 875–908 (2015).
12. Chen, P. & Redner, S. Community structure of the physical review citation network. *J. Informetr.* **4**, 278–290 (2010).
13. Herrera, M., Roberts, D. C. & Natali, G. Mapping the evolution of scientific fields. *PloS One* **5**, e10355 (2010).
14. Pan, R., Sinha, S., Kaski, K. & Saramäki, J. The evolution of interdisciplinarity in physics research. *Sci. Rep.* **2**, 551 (2012).
15. Guevara, M. R., Hartmann, D., Aristarán, M., Mendoza, M. & Hidalgo, C. A. The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations. *Scientometrics* **109**, 1695–1709 (2016).
16. Leslie, S. W. *The Cold War and American Science*. (Columbia University Press, New York, 1993).
17. Kaiser, D. I. Booms, busts, and the world of ideas: Enrollment pressures and the challenge of specialization. *Osiris* **27**, 276–302 (2012).
18. Martin, J. *Solid State Insurrection: How the Science of Substance made American Physics Matter*. (University of Pittsburgh Press, Pittsburgh, 2018).
19. ATLAS. ATLAS experiment reports. *CERN* https://atlas.cern/updates/atlas-news/atlas-experiment-reports-its-first-physics-results-lhc (2018).
20. Jia, T., Wang, D. & Szymanski, B. K. Quantifying patterns of research-interest evolution. *Nat. Human. Behav.* **1**, 0078 (2017).
21. Kaiser, D. I. Whose mass is it anyway? particle cosmology and the objects of theory. *Social. Stud. Sci.* **36**, 533–564 (2006).
22. Crosta, P. M. & Packman, I. G. Faculty productivity in supervising doctoral students? dissertations at cornell university. *Econ. Educ. Rev.* **24**, 55–65 (2005).
23. Malmgren, R. D., Ottino, J. M. & Amaral, L. A. N. The role of mentorship in protégé performance. *Nature* **465**, 622 (2010).
24. Chariker, J. H., Zhang, Y., Pani, J. R. & Rouchka, E. C. Identification of successful mentoring communities using network-based analysis of mentor–mentee relationships across nobel laureates. *Scientometrics* **111**, 1733–1749 (2017).
25. Zuckerman, H. Patterns of productivity, collaboration, and authorship. *Am. Sociol. Rev.* **32**, 391–403 (1967).
26. Ma, Y. & Uzzi, B. The scientific prize network predicts who pushes the boundaries of science. https://arxiv.org/abs/1808.09412 (2018).
27. Sekara, V. et al. The chaperone effect in science. *PNAS* (*in the press*).
28. Szell, M. & Sinatra, R. Research funding goes to rich clubs. *Proc. Natl. Acad. Sci.* **112**, 14749–14750 (2015).
29. Sinatra, R., Wang, D., Deville, P., Song, C. & Barabási, A.-L. Quantifying the evolution of individual scientific impact. *Science* **354**, aaf5239 (2016).
30. Liu, L. et al. Hot streaks in artistic, cultural, and scientific careers. *Nature* **559**, 396–399 (2018).
31. Radicchi, F., Fortunato, S. & Castellano, C. Universality of citation distributions: Toward an objective measure of scientific impact. *Proc. Natl. Acad. Sci.* **105**, 17268–17272 (2008).
32. Pavlidis, I., Petersen, A. M. & Semendeferi, I. Together we stand. *Nat. Phys.* **10**, 700 (2014).
33. Wuchty, S., Jones, B. & Uzzi, B. The increasing dominance of teams in production of knowledge. *Science* **316**, 1036–1039 (2007).
34. Shen, H.-W. & Barabási, A.-L. Collective credit allocation in science. *Proc. Natl. Acad. Sci.* **111**, 12325–12330 (2014).
35. Lehmann, S., Jackson, A. & Lautrup, B. Measures for measures. *Nature* **444**, 1003–1004 (2006).
36. Lehmann, S., Jackson, A. & Lautrup, B. A quantitative analysis of indicators of scientific performance. *Scientometrics* **76**, 369–390 (2008).
37. Hicks, D., Wouters, P., Waltman, L., Rijcke, S. D. & Rafols, I. Bibliometrics: the Leiden Manifesto for research metrics. *Nature* **520**, 429–431 (2015).
38. Waltman, L. A review of the literature on citation impact indicators. *J. Informetr.* **10**, 365–391 (2016).
39. Lillquist, E. & Green, S. The discipline dependence of citation statistics. *Scientometrics* **84**, 749–762 (2010).
40. Radicchi, F. & Castellano, C. Rescaling citations of publications in physics. *Phys. Rev. E* **83**, 046116 (2011).
41. Newman, M. The first-mover advantage in scientific publication. *EPL (Europhys. Lett.)* **86**, 68001 (2009).
42. Van Noorden, R. Interdisciplinary research by the numbers. *Nat. News* **525**, 306 (2015).
43. Szell, M., Ma, Y. & Sinatra, R. A Nobel Opportunity for Interdisciplinarity. *Nat. Phys.* **14**, 1075–1078 (2018).
44. Bromham, L., Dinnage, R. & Hua, X. Interdisciplinary research has consistently lower funding success. *Nature* **534**, 684–687 (2016).
45. arXiv. The arXiv repository. *Cornell University Library* https://arxiv.org/ (2018).
46. Martín-Martín, A., Orduna-Malea, E. & Delgado López-Cózar, E. Coverage of highly-cited documents in google scholar, web of science, and scopus:

a multidisciplinary comparison. *Scientometrics* **116**, 2175–2188 (2018).
47. Farmer, J. D. Physicists attempt to scale the ivory towers of finance. *Comput. Sci. & Eng.* **1**, 26–39 (1999).
48. May, R. M. The Scientific Wealth of Nations. *Science* **7**, 793–796 (1997).
49. King, D. K. The scientific impact of nations. *Nature* **430**, 311–316 (2004).
50. Zhang, Q., Perra, N., Goncalves, B., Ciulla, F. & Vespignani, A. Characterizing scientific production and consumption in physics. *Sci. Rep.* **3**, 1640 (2013).
51. Balassa, B. Trade liberalization and 'revealed' comparative advantage. *Manchester School* **33**, 99–123 (1965).

Acknowledgements
This work was supported by the John Templeton Foundation Grant #61066 (A.-L.B., F.B., R.S. and M.S.), the Intellectual Themes Initiative (ITI) project 'Just Data', funded by Central European University (F.M. and R.S.), the National Science Foundation grant SBE 1829344 (D.W.) and the Air Force Office of Scientific Research grants FA9550-15-1-0077 (A.-L.B., R.S. and M.S.), FA9550-15-1-0364 (A.-L.B. and R.S.), FA9550-15-1-0162 (D.W.) and FA9550-17-1-0089 (D.W.).

Author contributions
A.-L.B., R.S., M.S. and D.W. conceived the study. All authors designed the research, discussed the results and commented on the manuscript. F.B., F.M. and R.S. developed the methods. F.B. and F.M. analysed the data. M.S. and R.S. directed the research. F.B., F.M., M.S. and R.S. led the writing of the manuscript and A.-L.B. and D.W. edited the manuscript. F.B. and F.M. wrote the supplementary information.

Competing interests
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information
Supplementary information is available for this paper at https://doi.org/10.1038/s42254-018-0005-3.