

# INFERENCE

International Review of Science

---



**Frank Wilczek**



**Anna Maria Di Sciullo**



**Daniel Kleitman**



**William Press**



**Jeremy Bernstein**



**Neeraja Sankaran**



**Benjamin Davido**



**Tetsuro Matsuzawa**



**Chase Nelson**

## A Christmas Chrestomathy

*Selected Essays, Reviews & Letters*

**INFERENCE-REVIEW.COM**

**@inferencereview**

**(TWITTER, FACEBOOK, INSTAGRAM)**

*Copyright © Inference: International Review of Science 2021*  
*ISSN #2576-4403*

# Contents

## *A Christmas Chrestomathy, December 2021*

### CRITICAL ESSAYS

- 5 Battlefield Medicine Benjamin Davido

### REVIEW ESSAYS

- 11 Quanta of the Third Kind Frank Wilczek  
21 *On Aspects of the Theory of Syntax* Anna Maria Di Sciullo  
29 Primate Memory Tetsuro Matsuzawa  
39 At Lunch with Freeman Dyson William Press

### BOOK REVIEWS

- 45 *Unravelling the Double Helix* by Gareth Williams Neeraja Sankaran  
51 *Restricted Data* by Alex Wellerstein Jeremy Bernstein

### EXPERIMENT REVIEWS

- 57 Reconstructing Ancestral Proteins Chase Nelson

### BIOGRAPHIES

- 63 John Horton Conway: The Game of Life Daniel Kleitman

### LETTERS TO THE EDITORS

- 67 A Brief History of the Muon Sheldon Lee Glashow  
67 More Secrets Kenneth Ford, Jeremy Bernstein  
70 On Algorithmic Amplification Hany Farid  
72 On Modified Gravity Viktor Toth, Jean-Pierre Luminet

## Contributors

**Jeremy Bernstein** is Professor Emeritus of Physics at the Stevens Institute of Technology. He was a staff writer for the *New Yorker* from 1961 to 1995.

**Benjamin Davido** is a physician at the Hôpital Raymond-Poincaré de Garches.

**Anna Maria Di Sciullo** is Professor of Linguistics at the University of Quebec at Montreal.

**Daniel Kleitman** is Emeritus Professor of Applied Mathematics at MIT.

**Tetsuro Matsuzawa** is a Visitor in Psychology at Caltech and former Director of the Primate Research Institute of Kyoto University.

**Chase Nelson** is a Postdoctoral Research Fellow at Academia Sinica in Taipei and Visiting Scientist at the American Museum of Natural History in New York City.

**William Press** is the Leslie Surginer Professor of Computer Science and Integrative Biology at the University of Texas at Austin.

**Neeraja Sankaran** is a historian of science, science writer, editor, and educator who is currently a visiting scholar at Utrecht University.

**Frank Wilczek** is a Nobel Laureate, Herman Feshbach Professor of Physics at MIT, founding director of the T. D. Lee Institute, and chief scientist at the Wilczek Quantum Center in Shanghai.

---

## Masthead

### EDITOR IN CHIEF

Sheldon Lee Glashow

### EXECUTIVE EDITOR

Steven Wheeler

### CONTRIBUTING EDITOR

David Berlinski

### ASSOCIATE EDITOR

Meagan Mason

### MANAGING EDITOR

Theo Radford

### MARKETING DIRECTOR

Sarah Abbas

### CARICATURIST

Jean-Michel Gruet

### BOARD OF EDITORS

Narayana Murthy (Chairman),  
Sheldon Lee Glashow, Barry Barish,  
Noam Chomsky, George Ellis,  
Arthur Jaffe, Richard Roberts,  
Andrew Chi-Chih Yao, James Tour,  
Sergiu Klainerman, Martin Nowak,  
Reidun Twarock, Gregory Chaitin,  
Álvaro de Rújula, Luis Álvarez-Gaumé,  
Robert Berwick, Jay Bhattacharya,  
Jean-Pierre Dupuy, John Iliopoulos,  
Anna Maria Di Sciullo, Tetsuro Matsuzawa,  
Krzysztof Apt, Jeremy Bernstein,  
Richard Lindzen, Jean-Pierre Luminet,  
Pierre Schapira, Linda Bartoshuk,  
David Gelernter, Marc Henry,  
Iris Berent, Jae Chun Choe,  
Andrew Jordan, Juan Uriagereka.

---

### WEBSITE

[inference-review.com](http://inference-review.com)

### SOCIAL MEDIA

[@inferencereview](https://twitter.com/inferencereview)

### CONTACT

[info@inference-review.com](mailto:info@inference-review.com)

CRITICAL ESSAYS

# Battlefield Medicine

*Benjamin Davido*



*Published on July 7, 2021*

**E**VEN THOUGH FRANCE is the country of Louis Pasteur, the French have long been reluctant to get themselves vaccinated. In the 1990s, the vaccine against hepatitis B was suspected of causing multiple sclerosis.<sup>1</sup> If the suspicion proved unfounded, the anxiety it provoked proved unyielding. It is no surprise that when it came to COVID-19, the French remained distrustful. In December 2020, some forty percent of the French public were prepared to accept vaccination; at the end of January 2021, sixty percent.<sup>2</sup> The roll-out of the Pfizer-BioNTech mRNA vaccine was, in any case, glacial in pace: 1.6 million doses were available in France on February 1, 2021. This was a full month *after* the start of the vaccination campaign, which began officially on December 27, 2020. Both general practitioners and pharmacists have traditionally been allowed to administer vaccines, but the mRNA vaccines—Pfizer and Moderna—require deep refrigeration, and neither the general practitioners nor the pharmacists were in a position to provide it.

A more practical vaccine was needed.

Enter AstraZeneca.

Viral vectored, solid, old-fashioned, and, it was hoped, reliable, the AstraZeneca vaccine does not require deep refrigeration. If stored at between two and eight degrees Celsius, it could keep for at least six months.<sup>3</sup> By comparison, the Pfizer vaccine can be kept for up to thirty days, but only if stored at between minus ninety and minus sixty degrees Celsius and replenished with dry ice every five days. Before use, it must be thawed, after which it can be kept for five days, but only if stored at between two and eight degrees Celsius. Each Pfizer vial holds enough for six doses, which must be used within six hours, compared to AstraZeneca's two doses for forty-eight hours.<sup>4</sup>

On January 29, 2021, the European Medicines Agency (EMA), an agency of the European Union that evaluates and supervises medicines, gave the all clear for the AstraZeneca vaccine.<sup>5</sup> On the same day it was approved, French president Emmanuel Macron questioned publicly whether the AstraZeneca vaccine should be restricted to those under sixty-five years of age, describing it as “quasi-ineffective” for anyone above this threshold.<sup>6</sup> “What I can tell you officially today,” he remarked, “is that the early results we have are not encouraging for 60 to 65-year-old people concerning AstraZeneca.”

Four days later, on February 2, 2021, the *Haute Autorité de santé* (HAS)—France's public health authority—recommended the AstraZeneca vaccine to “all citizens” and “all professionals in the health, medico-social and social sectors.”<sup>7</sup> Yet that very same day, the HAS issued a press release in which AstraZeneca was recommended to the young and healthy, but not the elderly, noting that “the current data do *not* allow an assessment of the level of efficacy that this vaccine provides in people over 65 years of age.”<sup>8</sup> The explanation given for these doubts was “the small number of participants aged 65 and over in the trials.”

In the space of four days, two agencies—one European, one French—and the French president had offered differing assessments on the efficacy and potential limitations of the AstraZeneca vaccine. The inevitable result of this mixed messaging from the authorities was confusion among the public concerning the vaccine.

It was against this uncertain background that AstraZeneca vaccinations began. As it turned out, the French public did not have to wait long for further conflicting assessments to emerge. On February 11, 2021, the ANSM (*L'Agence nationale de sécurité du médicament et des produits de santé*) noted that for every 10,000 individuals injected, roughly 150 of them suffered a number of flu-like symptoms.<sup>9</sup> The ANSM report created tremendous anxiety, particularly among nurses: “AstraZeneca ... is a fine vaccine for the general public,” observed Thierry Amoureux, spokesperson for France's national nurse union, “but for a population as exposed as healthcare workers, it is among the three least effective of the authorized vaccines.”<sup>10</sup>

A month later, the news was even worse. On March 16, 2021, the EMA issued a press release concerning a possible link between the AstraZeneca vaccine and episodes of thrombosis.<sup>11</sup> Flu-like symptoms are one thing; thrombosis, another. A blood clot wandering to the brain, the heart, or the lungs may well prove fatal. Nonetheless, the EMA was still persuaded that “the vaccine's proven efficacy in preventing hospitalisation and death from COVID-19 outweighs the extremely small likelihood of developing [blood clots].” To put this into perspective, an article published by the *BMJ* on March 11 noted that just “30 cases of thromboembolic events had been reported among the five million people given the AstraZeneca vaccine in the European Economic Area.”<sup>12</sup>

Given the initial advice from the EMA, France chose not to immediately suspend the AstraZeneca vaccination campaign. In Germany, AstraZeneca vaccinations had already been suspended on March 15 as a precautionary measure, on the advice of the Paul Ehrlich Institute, the country's vaccine authority.<sup>13</sup> The real shock came just two days later on March 17 when the *Élysée* backtracked on their decision to stand by the AstraZeneca jab—an about-face that did nothing to reassure the French public. Speaking off the record, a source in Macron's administration admitted that Berlin's earlier decision had a “psychological impact” on the French government.<sup>14</sup>

The EMA initially appeared confused as to just which principles its member states were appealing to in suspending the vaccine. On March 18, the EMA's executive director, Emer Cooke, fielded a question regarding the possibility of “harmonizing the application of the precautionary principle,” which she diplomatically deflected, responding that, “There's a lot of things that we need to do to make sure that everybody has the same information about the benefits and risks.”<sup>15</sup> The ambassador of the European Union (EU) to the United Kingdom, João Vale de Almeida,

had other ideas. “When there are doubts, the principle of precaution prevails,” he remarked on March 16.<sup>16</sup> And he was not alone. “Acting on the precautionary principle, and pending receipt of further information,” Ireland’s Deputy Chief Medical Officer, Dr. Ronan Glynn, announced that, “the NIAC [National Immunisation Advisory Committee] has recommended the temporary deferral of the COVID-19 Vaccine AstraZeneca vaccination program in Ireland.”<sup>17</sup>

The precautionary principle was formally adopted in the Maastricht Treaty of 1992. In essence, this article allows the EU to take preventative action in the case of risk—specifically, “when a phenomenon, product or process may have a dangerous effect, identified by a scientific and objective evaluation.”<sup>18</sup> Swedish philosopher Per Sandin defined the precautionary principle as “if [emphasis original] there is a threat, which is uncertain, then some kind of action is mandatory.”<sup>19</sup> The article is sufficiently vague as to allow the EU to take action even when scientific evaluation “does not allow the risk to be determined with sufficient certainty,” as the EU themselves have noted.<sup>20</sup> In their short book on the precautionary principle, Gary Marchant and Kenneth Mossman make the point that

[t]he treaty itself does not define or otherwise articulate the requirements of the precautionary principle, nor was any official explanation or definition of the precautionary principle provided during the process of its enactment into the European Treaty. It has therefore been left to the community institutions to define and apply the precautionary principle.<sup>21</sup>

In 2000, the European Commission released a communiqué that sought to “establish a common understanding of the factors leading to recourse to the precautionary principle and its place in decision making, and to establish guidelines for its application based on reasoned and coherent principles.”<sup>22</sup> The document included a series of general principles for the application of the precautionary principle. Among them was a cost-benefit analysis that entailed

comparing the overall cost to the Community of action and lack of action, in both the short and long term. This is not simply an economic cost-benefit analysis: its scope is much broader, and includes non-economic considerations, such as the efficacy of possible options and their acceptability to the public. In the conduct of such an examination, account should be taken of the general principle and the case law of the Court that the protection of health takes precedence over economic considerations.<sup>23</sup>

This dictates that the pros and cons of any decision made on the basis of the precautionary principle must be carefully weighed before any action is taken. In the case of the AstraZeneca vaccine, intervention in France meant halting an already unpopular vaccination drive, and an

extraordinary number of wasted vaccines, whereas inaction would have meant the continued rollout of a vaccine that had already proved effective in helping curb the UK’s spiraling COVID cases.<sup>24</sup> The decision should not have been hard to make. Alas, were the EU forced into a decision, perhaps it would have been this easy. The scenario that unfolded, in which each European country opted to decide on its own, was unforeseeable and inevitably created widespread doubt among the bloc’s populations. This was first the choice of Denmark, then Iceland, then Norway. The crisis has affected all of Europe in much the same way and each member of the EU has applied the precautionary principle in just the same way as well.

Even before the concerns about a possible connection to blood clots began to appear in mid-March 2021, the AstraZeneca vaccine had already been singled out for criticism in France. At the beginning of March, a nurse’s union expressed doubt that AstraZeneca was best adapted to those at high risk of exposure to the virus.<sup>25</sup> At the time, the government declared no reason for doubt and urged that, their concerns notwithstanding, healthcare workers should promptly be vaccinated with the AstraZeneca vaccine.<sup>26</sup>

By March 29, only forty-two percent of nurses in France had received one dose of the vaccine, compared to eighty percent of general practitioners, and twelve percent of the general population.<sup>27</sup> The figures were no less shocking in mid-June, when French daily *Le Monde* gained access to a document from the *Assistance Publique–Hôpitaux de Paris* (Public Hospitals of Paris) suggesting that while ninety-one percent of general practitioners had received one dose, nonmedical staff, including nurses and caregivers, were still well behind at fifty-four percent.<sup>28</sup> This figure was always unlikely to increase rapidly since the European bulk-order strategy led to a shortage of vaccines.<sup>29</sup> The decision to use age as a cutoff for nurses, as well as for the general public, has not helped matters. Using the risk of viral exposure would have made for a more logical strategy with respect to nurses. Anyone working in intensive care should have been first in line for vaccination regardless of age. During the first wave, contamination was directly linked to exposure. Beyond that, the problem is rooted in community transmission, with most hospital personnel becoming infected outside the hospital, at private gatherings.<sup>30</sup>

Accelerating the vaccination campaign means getting vaccinations done quickly in the hope that hospitalizations for COVID-19 decrease. Israel is a model. The symbolic threshold of 4,000 patients in ICU (intensive care units) was passed in France on March 15, exceeding the peak of the second wave. It was not until May 21 that this figure dropped below 4,000 patients once again.<sup>31</sup> The decision to suspend AstraZeneca means that hospitals have been forced to throw away doses of ready-to-use vaccines and to cancel appointments for vaccination.<sup>32</sup> To compensate, physicians have been told to extract a seventh dose of Pfizer from vials initially earmarked for five doses.<sup>33</sup>

As infection rates and ICU admissions remain high, it is clear that France and Europe must continue to accelerate the vaccination drive. The sooner we vaccinate, the sooner any vaccine hesitancy will be eased. To stop vaccinating for forty-eight to seventy-two hours in the name of the precautionary principle is to start work again on vaccine awareness from scratch. Any decision taken must take into account the risk-benefit balance that is largely in favor of *all* of the vaccines available to date. Their effectiveness, as research in both Europe and the US has shown, is, by definition, greater than fifty percent.<sup>34</sup> As a result, the AstraZeneca vaccine has been stigmatized and patients who know nothing of even the simplest risk-benefit calculation are now demanding the right to choose their vaccines.<sup>35</sup>

At the beginning of the pandemic, masks were initially declared to be of no use, and then made mandatory indoors—and then made mandatory outdoors, but only in the summer of 2020. As the first wave swept across Europe in early 2020, France carried out fewer COVID tests than most other countries in Europe,<sup>36</sup> finally screening everyone who wanted to be tested free of charge, including those who were asymptomatic and without a prescription, by September 2020.<sup>37</sup>

It is always easy to see things better backwards, but perhaps a Europe-wide decision should have been applied from the start. The precautionary principle might have suggested that given the real risk of disease and death, all of Europe should have been placed under strict quarantine. If this is well-known as a medieval strategy, it is no less effective for that. No matter the virus, either it dies with its host or its host outlives it; if it cannot spread, it cannot survive. If a complete and effective quarantine is economically unfeasible, the confusion created by the AstraZeneca debacle was unacceptable in a public health crisis—the French government both declaring that the AstraZeneca vaccine is safe in view of the benefit-risk balance of the vaccine, and then, 48 hours later, declaring that, since Germany has applied the precautionary principle, so we will also apply it. The AstraZeneca vaccine is now tainted, and will remain a second-rate vaccine in the eyes of the public, if only because the French, at least, remember that questions have been raised about its efficacy and potential undesirable effects.<sup>38</sup>

It is always safe enough in France to blame any muddle on a failure of communication. The possible effects of the vaccine may not have been sufficiently explained; and no one in government has clearly made the distinction between individual risk and collective welfare. The risk of thrombosis provoked by the AstraZeneca vaccine is *lower* than the risks of thrombosis observed in young smokers taking estrogen-progestogen birth control pills.<sup>39</sup> On the other hand, patients suffering COVID-19 infection run an absolutely clear and significantly increased risk of thromboembolic events.<sup>40</sup> The cost-benefit assessment is entirely in favor of AstraZeneca vaccination. Recent data

analysis from a team at Oxford shows a thirty-nine-in-one-million chance of developing thrombosis in the two weeks following a COVID-19 diagnosis.<sup>41</sup> In the fortnight following a Moderna or Pfizer jab, the same study has shown that this figure drops to four in a million, while the current EMA estimate holds the chances of thrombosis from the AstraZeneca jab at the slightly less favorable five in a million. The scales of profit-risk are leaning heavily to one side. The risk-benefit balance is overwhelmingly in favor of vaccination.

What is new and specific to this epidemic is the association of vaccine and laboratory names. More and more people are calling vaccination centers to ask for the name of their vaccine. The world's geography is now at issue, with vaccines appearing from America, England, Russia, and China. Cost plays a role in all this. It always does. Pfizer and Moderna vaccines are the most expensive at nineteen US dollars and fifty cents, and fifteen dollars per shot, respectively.<sup>42</sup> It is not in the interests of France, nor of Europe, to suspend AstraZeneca vaccinations. AstraZeneca is fourfold less expensive than the mRNA vaccines. Seven million doses of Johnson & Johnson's Janssen vaccine and thirty-five million Pfizer doses are scheduled to arrive in France by the end of the summer. There is no time to wait for one vaccine to be replaced by another of equivalent quality, and that is a real concern. In the end, it is the opposite of what France experienced in 2009 with the H1N1 pandemic, when Roselyne Bachelot, the minister of health at the time, ordered double the amount of required vaccines after failing to follow the up-to-date advice that booster doses were not necessary.<sup>43</sup> Fortunately, the H1N1 pandemic did not hit Europe as expected, and this mistake did not influence public opinion.

What is to be done now, and in the future when the next pandemic comes along? We need something stronger than the current sanitary advertisements in which a grandmother kisses her markedly unenthusiastic grandchildren because she has been vaccinated. To break the chain of transmission, it is not enough to attend to the elderly. Israel began vaccinating sixteen- to-eighteen-year-olds in late January 2021,<sup>44</sup> and, in mid-May, began preparations to vaccinate twelve- to-fifteen-year-olds.<sup>45</sup> There must be, above all, a rediscovery of decision-making principles that have, at least, some clear content. Something like a collective sanitary benefit really exists. A good example of the individual benefits associated with vaccination is the vaccine passport that was implemented in Europe during July 2021. The adoption of the vaccine passport offers a route to return to normal life, free of restrictions. What is crucial in promoting a collective benefit is not exclusively the effectiveness of a vaccine in protecting symptomatic forms of a disease. It is a matter of eliminating the virus in the first place. This cannot be done on the basis of any kind of individual risk assessment.

A final note of caution is, perhaps, warranted. Humanity has been lucky with respect to COVID-19. Some diseases

are far more lethal, but spread slowly. Although COVID-19 spreads rapidly, it is not terribly lethal.

The next pandemic may be both.

*Translated and adapted from the original French by the editors.*



1. “Une augmentation de 65% de SEP après le pic de vaccination contre l’hépatite B en 1994 (A 65% Rise in Multiple Sclerosis after Vaccination against Hepatitis B in 1994),” *Le Quotidien du médecin*, December 1, 2014.
2. “U.S. and U.K. Are Optimistic Indicators for COVID-19 Vaccination Uptake,” *Ipsos*, December 29, 2020.
3. Kyle Blankenship, “AstraZeneca Hopes Warmer Storage Needs for COVID-19 Vaccine Will Be an Advantage against mRNA Competitors,” *Fierce Pharma*, November 23, 2020.
4. Minnesota Department of Health, Interim COVID-19 Vaccine Provider Guide, Appendix A: Pfizer-BioNTech COVID-19 Vaccine (Minnesota: Minnesota Department of Health, 2021).
5. European Medicines Agency, “EMA Recommends COVID-19 Vaccine AstraZeneca for Authorisation in the EU,” January 29, 2021.
6. “Macron: AstraZeneca Vaccine ‘Quasi-Ineffective’ for Over-65s,” *France 24*, January 29, 2021.
7. “Stratégie de vaccination contre la Covid-19 – Place du Covid-19 Vaccine AstraZeneca (Vaccination Strategy against COVID-19 – The Place of the COVID-19 Vaccine AstraZeneca),” *Haute Autorité de santé*, February 2, 2021. Translation by the editors.
8. “AstraZeneca : La HAS recommande son utilisation chez les professionnels de santé et les personnes de 50 à 64 ans (AstraZeneca: The HAS Recommends Its Use among Healthcare Professionals and 50- to 64-Year-Olds),” *Haute Autorité de santé*, February 2, 2021. Translation by the editors.
9. Thomas Dubois, “Effets secondaires du vaccin AstraZeneca : ‘Un moment un peu désagréable’ (Secondary Effects of the Astrazeneca Vaccine: ‘A Bit of an Unpleasant Moment’),” *Paris Normandie*, February 12, 2021.
10. Viviane Le Guen, “Coronavirus : Pourquoi les soignants sont-ils divisés sur la vaccination ? (Coronavirus: Why are Carers Divided over Vaccination?),” *France Bleu*, March 5, 2021. Translation by the editors.
11. “COVID-19 Vaccine AstraZeneca: Benefits Still Outweigh the Risks Despite Possible Link to Rare Blood Clots with Low Blood Platelets,” *European Medicines Agency*, March 18, 2021.
12. Jacqui Wise, “Covid-19: European Countries Suspend Use of Oxford-AstraZeneca Vaccine after Reports of Blood Clots,” *BMJ* 372 (2021): n699, doi:10.1136/bmj.n699. The same figures were quoted in an article published in *The Lancet* at the end of the month: Søren Dinesen Østergaard et al., “Thromboembolism and the Oxford–AstraZeneca COVID-19 Vaccine: Side-Effect or Coincidence?” *The Lancet* 397, no. 10,283 (2021): 1,441–43, doi:10.1016/S0140-6736(21)00762-5.
13. “Germany Suspends Use of AstraZeneca Vaccine, Along with Italy, France, Spain,” *DW*, March 15, 2020.
14. Sarah Wheaton, “Revenge of the Precautionary Principle,” *Politico*, March 18, 2021.
15. “EMA Press Conference,” *European Medicines Agency* (YouTube video), March 18, 2021. The relevant segment begins at 41:57.
16. “Principle of Precaution Must Prevail, Says EU Ambassador,” *BBC*, March 16, 2021; “Revenge of the Precautionary Principle.”
17. “WHO Says Countries Should Keep Using AstraZeneca COVID Vaccine as European Countries Suspend Use,” *EuroNews*, March 16, 2021.
18. “The Precautionary Principle,” European Commission Communication, February 2, 2000.
19. Per Sandin, “Dimensions of the Precautionary Principle,” *Human & Ecological Risk Assessment: An International Journal* 5, no. 5 (1999): 889–907, doi:10.1080/10807039991289185.
20. “The Precautionary Principle.”
21. Gary Marchant and Kenneth Mossman, *Arbitrary and Capricious: The Precautionary Principle in the European Union Courts* (London: International Policy Press, 2005), 25.
22. “Communication from the Commission on the Precautionary Principle,” European Commission Communication, February 2, 2000.
23. “Communication from the Commission.”
24. Merryn Voysey et al., “Single Dose Administration, and the Influence of the Timing of the Booster Dose on Immunogenicity and Efficacy of ChAdOx1 nCoV-19 (AZD1222) Vaccine,” *The Lancet* (pre-print) (2021), doi:10.2139/ssrn.3777268.
25. “Covid-19 : Le vaccin AstraZeneca n’est ‘pas le plus adapté’ pour ‘une population aussi exposée que les soignants’, estime un syndicat d’infirmiers (Covid-19: The AstraZeneca vaccine is ‘not the most adapted’ for ‘a population as exposed as caregivers,’ according to Nurses’ Union),” *Franceinfo*, March 5, 2021.
26. Olivier Véran, Twitter post, March 5, 2021, 5:30pm.
27. “Point Épidémiologique COVID-19 no. 59, 15 avril 2021 (COVID-19 Epidemiological Update no. 59, April 15, 2021),” *Santé Publique France*, 54.
28. Stéphane Mandard and Camille Stromboni, “Covid-19 : Pourquoi la vaccination plafonne chez les infirmiers et les aides-soignants (COVID-19: Why Vaccination Has Plateaued among Nurses and Nursing Aides),” *Le Monde*, June 18, 2021.
29. Ben Hall et al., “Shortage of Coronavirus Shots Heaps Pressure on European Leaders,” *Financial Times*, January 29, 2021.
30. Benjamin Davido et al., “The First Wave of COVID-19 in Hospital Staff Members of a Tertiary Care Hospital in the Greater Paris Area: A Surveillance and Risk Factors Study,” *International Journal of Infectious Diseases* 105 (2021): 172–79, doi:10.1016/j.ijid.2021.02.055.

31. Hannah Ritchie et al., “Coronavirus Pandemic (COVID-19) – the Data,” *Our World in Data* (2021).
32. Académie Nationale de Médecine, “Between the Essential Precautions and the Undesirable Confusion,” March 23, 2021.
33. Klervi Drouglazet, “Covid-19. Est-il vraiment possible d’extraire une septième dose des flacons du vaccin Pfizer ? (COVID-19. Is It Really Possible to Extract a Seventh Dose from the Pfizer Vaccine Vials?),” *Ouest France*, March 13, 2021.
34. Jamie Lopez Bernal et al., “Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca Vaccines on Covid-19 Related Symptoms, Hospital Admissions, and Mortality in Older Adults in England: Test Negative Case-Control Study,” *The British Medical Journal* 373, no. 1,088 (2021), doi:10.1136/bmj.n1088.
35. For example, Huguette Poitras, “Choisir son vaccin (Choosing One’s Vaccine),” *Le Devoir*, March 22, 2021: “Plusieurs voix s’élèvent en ce moment pour revendiquer le droit de choisir son vaccin contre la COVID-19 (Many voices are rising up at the moment to claim the right to choose one’s vaccine against COVID-19).” [Segment translated by the editors of *Inference*.]
36. Louis Tanca, “La France teste beaucoup moins sa population que les autres pays (France Tests Its Population a Lot Less than Other Countries),” *BFMTV*, March 31, 2020.
37. “Tests PCR : Plus besoin d’une ordonnance pour en bénéficier et être remboursé (PCR Tests: No Longer Any Need for a Prescription to Benefit from Them, and Be Reimbursed),” *Direction de l’information légale et administrative* (Premier ministre), September 18, 2020.
38. Anne-Laure Frémont, “71% des Français ne font pas confiance au vaccin AstraZeneca (71% of French People Do Not Trust the AstraZeneca Vaccine),” *Le Figaro*, April 8, 2021.
39. “Birth Control Creates Higher Risk of Blood Clots than AstraZeneca Vaccine,” *Open Access Government*, March 16, 2021.
40. Mahmoud Malas et al., “Thromboembolism Risk of COVID-19 Is High and Associated with a Higher Risk of Mortality: A Systematic Review and Meta-Analysis,” *EClinicalMedicine* 29, no. 100,639 (2020), doi:10.1016/j.eclinm.2020.100639.
41. Maxime Taquet et al., “Cerebral Venous Thrombosis and Portal Vein Thrombosis: A Retrospective Cohort Study of 537,913 COVID-19 Cases,” *medRxiv*, May 11, 2021, doi:10.1101/2021.04.27.21256153. For a summary, see Shawna Williams, “Blood Clot Risk from COVID-19 Higher than after Vaccines: Study,” *The Scientist*, April 16, 2021.
42. Eric Sagonowsky, “Pfizer Eyes Higher Prices for COVID-19 Vaccine after the Pandemic Wanes,” *Fierce Pharma*, February 23, 2021.
43. Sandrine Cabut, “Vaccins H1N1 : Le ministère de la Santé mis en cause (H1N1 Vaccines: Ministry of Health Implicated),” *Le Figaro*, October 14, 2010.
44. “Israel Expands Vaccination Campaign to Teens Aged 16-18,” *Times of Israel*, January 23, 2021.
45. “Israel Set to Vaccinate 12- to 15-Year-Olds after FDA Okays Pfizer Shot,” *Times of Israel*, May 11, 2021.

DOI: 10.37282/991819.21.24

REVIEW ESSAYS

# Quanta of the Third Kind

*Frank Wilczek*



*Published on November 17, 2021*

QUANTUM MECHANICS is nearly one hundred years old, and yet the challenge it presents to the imagination is so great that scientists are still coming to terms with some of its most basic implications. Here I will describe some theoretical insights and recent experimental results that are leading physicists to revise and expand their ideas about what quantum-mechanical particles are and how they behave. These new ideas are centered around a topic traditionally known as quantum statistics. The name is misleading: the basic physical phenomena do not involve statistics in the usual sense. A better title might have been the quantum mechanics of identity, but the new developments make that name obsolete too. A more accurate description would be the quantum mechanics of world-line topology. Since that is quite a mouthful, most researchers now simply refer to anyon physics.

QUANTUM MECHANICS achieves a strange and wonderful unification between forms of matter that appear to be vastly different. Prior to the advent of quantum theory, electrons and atomic nuclei were regarded as particles, conforming to Isaac Newton's classic definition: "hard, massy, impenetrable." During the nineteenth century, light came to be understood in terms of waves or, ultimately, space-filling electromagnetic fields. That description supplies a rich and accurate account of interference, diffraction, and many aspects of the interaction between light and matter.

In quantum theory, electrons, light, and all other forms of matter are described using the same mathematics. A more general concept, sometimes expressed as the *wavicle*, governs everything. The *wavicle* is a space-filling function—the wave function—that describes the probability of finding a particle at different places. This common framework accommodates both the wave behavior of electrons, manifested in electron diffraction, and the particulate nature of light, manifested in the all-or-none response of the photoreceptor cells responsible for color vision.

Most popular and even introductory textbook accounts of quantum theory stop there. But the great *wavicle* unification has an important qualification: it applies only to single particles. When we compare the quantum description of two or more electrons with that of two or more photons, we find fundamental differences.

ELECTRONS, NEUTRONS, and protons are examples of fermions, named in honor of Enrico Fermi, who pioneered their study.<sup>1</sup> Fermions are antisocial by nature. More precisely, they obey the Pauli exclusion principle,<sup>2</sup> which states that no two fermions of the same kind can be in the same quantum state. The exclusion principle plays a central role in our understanding of atoms, atomic nuclei, white dwarfs, neutron stars, and matter in general:

- In many-electron atoms, the exclusion principle forces the electrons to occupy different orbitals.

This behavior is essential for building up the shell structure of atoms, which underlies the periodic table of elements and chemistry.

- Similarly, in atomic nuclei, the exclusion principle governs the behavior of the protons and neutrons, building up the shell structure that controls the nuclear chemistry of fission and fusion.
- Freeman Dyson and Andrew Lenard demonstrated mathematically that if the equations of quantum theory are applied to the ingredients of ordinary matter without taking into account that electrons are fermions, the mixture implodes.<sup>3</sup> Once the electrons are treated as fermions, all is well.
- White dwarfs are the evolutionary final state of moderate-sized stars, such as our sun. After they have exhausted their nuclear fuel, these stars collapse to much smaller sizes. The sun, for example, will eventually become Earth sized. Beyond a certain point, the electrons within these objects prevent further compression because too small an object would contain too few orbitals to accommodate them.
- Neutron stars are the final stage in the evolution of somewhat larger stars. Subrahmanyan Chandrasekhar demonstrated that there exists a limit to how much pressure electrons can withstand.<sup>4</sup> Past the so-called Chandrasekhar limit, stellar remnants up to about twice the sun's mass collapse further, down to a few kilometers in radius, whereupon the exclusion principle for neutrons halts the process.

Due to their Fermi statistics, identical fermions exhibit a repulsive force of quantum-mechanical origin that is over and above the four conventional forces of the Standard Model—strong, weak, electromagnetic, and gravitational. That effective force is not merely an esoteric addition to the basic forces, but a central pillar in our understanding of nature.

Photons, together with gravitons, the Higgs particle, and many other particles, are examples of bosons,<sup>5</sup> named in honor of Satyendra Bose, who was the first to study them.<sup>6</sup> In contrast to fermions, bosons are natural conformists and prefer to be in the same state. The probability for multiple occupancy, which vanishes for fermions, is enhanced for bosons, which are said to obey Bose statistics.

Laser beams epitomize the bosonic behavior of photons. Within a laser beam, many photons have succeeded in occupying the same state with the same color, same direction, and same spatial profile. More complex material manifestations of Bose statistics are superfluidity and superconductivity. In those low-temperature states of matter, large numbers of  $^4\text{He}$  atoms or Cooper pairs of electrons, respectively, occupy the same quantum state and thus flow coherently. At low temperatures, when they are less distracted by the noise of the external world, one might say that bosonic particles get to do what they want to do—which is to do the same thing.

QUANTUM MECHANICS coalesced during the 1920s. Decades of adventurous discovery followed, in which many new particles were identified. The behavior of these particles involved many qualitatively new phenomena, including antimatter, strangeness, oscillatory changes in identity, and violation of spatial parity and time-reversal symmetry. Particles that could not be observed in isolation—quarks and gluons, which are fermions and bosons,<sup>7</sup> respectively—became fundamental ingredients in our best description of nature. During all these upheavals, the division of the world of particles into just two kingdoms, those of fermions and bosons, remained intact. By the 1970s these notions had become conventional wisdom, bordering on dogma.

In 1977, two Norwegian physicists, Jon Leinaas and Jan Myrheim, challenged that consensus.<sup>8</sup> Subsequent investigation has clarified the profound roots of quantum statistics, why bosons and fermions are so pervasive, and the possibility of alternatives. Before proceeding further, it is worth taking a moment to review the ultimate source of boson and fermion behavior, as presently understood.

The fundamental task of quantum mechanics is to calculate the probability for a specified event to occur. This is done by calculating an auxiliary quantity, the *amplitude* of the event, and then squaring the amplitude to obtain the probability.<sup>9</sup>

There are several ways to calculate quantum mechanical amplitudes. The most transparent method was discovered by Richard Feynman.<sup>10</sup> The “sum over histories” approach involves a consideration of all possible ways in which a process might have happened. The dynamical description of the system provides a definite mathematical rule, or algorithm, that assigns a numerical base-amplitude to each possible history. The total amplitude is obtained by adding all the base-amplitudes. In this framework, the central task of fundamental quantum theory is to discover the rules for calculating base-amplitudes in different physical situations. Physicists often look to classical physical for guidance, because for large objects the quantum rules must reproduce observed classical behavior.

With that framework in mind, consider a process in which two indistinguishable particles—two electrons, say, or two photons—start at two positions (A, B) and end up at two other positions (C, D). The possible histories underlying this process fall into two distinct classes. In one class, the particle originating at A travels to C, while the particle originating at B travels to D. In the other class, the particle at A travels to D while the particle at B travels to C. Since the particles are indistinguishable, one cannot tell, by looking only at the outcome, which class of historical process led to it. Guided by classical physics, physicists can develop rules for how to assign base-amplitudes within each of the two classes. Adding the base-amplitudes within each class yields two partial amplitudes.

The remaining issue is to determine a rule for combining the two partial amplitudes into the total amplitude.

Classical physics offers no guidance here. Indeed, classical physical theory assures us that in principle we can keep accurate tabs on particles. But if that is the case, the two topologically distinct classes of histories correspond to physically distinct processes, each of which is characterized by a separate probability. In quantum theory, on the contrary, one cannot keep tabs. That is an aspect of Heisenberg’s uncertainty principle, which limits how well a particle’s position can be resolved. When the uncertainties in the positions of our two particles overlap, it becomes impossible to keep track of who is who.<sup>11</sup>

The rule for combining the two partial amplitudes must then involve some essentially new consideration that goes beyond classical physics. Quantum statistics, with all its weighty implications for physics, ultimately comes down to this rule. The traditional rules are as follows: for bosons, add them, and for fermions, subtract them. These are the only two available choices. These are the only two available choices, it seems, because quantum theory imposes an important general consistency requirement.<sup>12</sup> If we apply our rule twice to the process  $(A, B) \rightarrow [(A, B) \text{ or } (B, A)] \rightarrow (A, B)$ —we must obtain the same result as we get by applying it directly—e.g., to  $(A, B) \rightarrow (A, B)$ . Thus, since a double exchange is equivalent to no exchange at all, the factor  $x$  that we can associate with an exchange must satisfy  $x^2 = 1$ . This implies that either  $x = 1$ , as for bosons, or  $x = -1$ , as for fermions.

THIS ELEGANT and superficially profound understanding of why there can be bosons and fermions, and nothing else, relies on an important implicit assumption that escaped attention of physicists for more than fifty years. Consider two particles whose motion is confined to two dimensions—specifically, a plane. To carry out the sum over histories, one must consider how the particles move in time as they progress from (A, B) to (C, D). In visualizing this problem, it is convenient to regard time as a third dimension, perpendicular to the two spatial dimensions. The motion of each particle then defines a path in a three-dimensional space-time known as a *world-line*.

The world-lines of two particles can wind around one another; and if they do, there is a discrete topological distinction among the histories from (A, B) to (C, D)—namely, the number of times the world-line of the first particle winds around the world-line of the second particle. Mathematicians refer to this as the winding number.<sup>13</sup> In cases of more than two particles, world-lines can become interwoven in elaborate patterns termed *braids*.

For particles that move in three dimensions, the need to consider winding and braiding processes no longer arises. The contrast between the rich topology of multiparticle histories in two dimensional spaces—i.e., three-dimensional space-times—and the paltry topology of multiparticle histories in three-dimensional spaces—i.e., four-dimensional space-times—is closely related to the fact that in four dimensions, though not of course in three, knots are

always easy to untangle.<sup>14</sup> In the topology of braids, less is more, since a smaller ambient space means less room for untangling maneuvers.

In three dimensions, the only two topologically distinct classes of histories are the ones involving exchanges of position. Given that mathematical fact, the potentially profound argument given in the previous section then becomes definitive. The meagre topology of multiparticle paths in three-dimensional space offers only the choice between bosons and fermions, with no other options. This is a very satisfying result because it justifies the classification into bosons and fermions that physicists had discovered empirically.

By contrast, the richer topology of multiparticle paths in two-dimensional space supports a much bigger menu of consistent quantum-mechanical rules. One can add partial amplitudes that arise from infinitely many topologically distinct classes, and the consistency conditions are less constraining. In a 1982 paper, I introduced the term *anyon* to describe this situation, with the connotation that anything goes.<sup>15</sup> While it is not literally true that anything goes, theoretically speaking, the move to flatland opens up many new possibilities for the quantum statistics of particles. Indeed, the kingdom of anyons has many mansions.

The consistent rules for quantum statistics in two space dimensions can become quite complicated—at least as complicated as braids. I will describe here only the very simplest anyons precisely, and briefly acknowledge some others. Even in the simplest case, the anyon rule for combining different partial amplitudes uses basic concepts about complex numbers.<sup>16</sup> There are different topological classes of braids, distinguished among other things by the number of times different strands wind around one another. A rule is needed for combining the partial amplitudes from those sectors. The simplest anyon rule is as follows: multiply the partial amplitude for each class by the complex number  $e^{i\alpha W}$ , where  $W$  is equal to the total number of windings, counting all pairs of particles. Different values of  $\alpha$  define different species of anyons.

More structured, so-called non-abelian anyon rules are sensitive to other details associated with braids. These rules usher in more complicated wave functions that are not simply complex numbers, but arrays of numbers. The different components in the array represent possible values of an emergent degree of freedom, roughly analogous to the possible colors of quarks. When non-abelian anyons wind around one another, their joint wave function is transformed by more complicated operations than multiplication by a number—that is, multiplication by a unitary matrix. In this way, non-abelian anyons acquire a strange, capacious storage capacity. Their quantum-mechanical wave functions carry a more detailed record of the braids their world-lines build up, which tracks more information than total winding.

The possible rules defining different species of non-abelian anyons are intricate and diverse. Here, a few names

and references will have to suffice: Ising anyons, Fibonacci anyons, parafermions of several kinds, and Majorinos.<sup>17</sup> It is fascinating to observe that world-lines can wind even among particles that are not indistinguishable. That possibility, which I named *mutual statistics*, highlights the novelty of the new, more general perspective on quantum statistics.

These extraordinary new possibilities for physical behavior are fun to think about, but they might also seem somewhat academic or fantastical. After all, we do not live in flatland. But we can still visit. In fact, the physical world abounds in flatlands, and they play starring roles in modern technology. Planar circuits photolithographically etched onto layered surfaces, otherwise known as chips, are the workhorses of modern microelectronics. In electronic chips, the motion of electrons is essentially confined to two dimensions. If the electrons that live on a chip were zapped with enough energy they could be removed. But so long as their energies don't get too big, the electrons are confined to two dimensions, and the quantum mechanics of flatland applies.

PHYSICISTS HAVE learned to organize their fundamental descriptions of the quantum world by focusing on the behavior of energy concentrations that are reasonably stable and exhibit reproducible properties. Such entities are termed elementary particles, and they are used as the building-blocks in our best model of the physical world.

In thinking about exotic materials and states of matter, it has been fruitful to consider them as worlds in themselves: *quasi-worlds*, inhabited by *quasiparticles*.

Suppose that a crystalline solid is zapped with a well-focused laser pulse and an electron is ejected. The remaining material will then contain a localized unit of positive charge where the electron used to be. After radiating some excess energy, this excitation may settle into a stable, reproducible form—a quasiparticle. This kind of quasiparticle, which reflects the absence of an electron, is usually referred to as a hole. In semiconductors, holes are units of positive charge that are cheap to produce and much easier to move than protons. Understanding the properties of holes was a key step in the invention of solid-state transistors and the emergence of modern microelectronics.

That success story, and others like it, has inspired some physicists to cultivate an art that might be described as designing quasi-worlds.<sup>18</sup> To begin with, one must imagine quasi-worlds with interesting properties, and then seek out or manufacture materials and states of matter that embody them.<sup>19</sup> Of course, that strategy can only work if the quasi-worlds are not too outlandish. Success requires discipline and good taste, as well as inspiration.

I first began to consider the ideas and new possibilities for quantum statistics in 1982. At the start, I was unaware of the work of Leinaas and Myrheim, which had attracted

little attention. I was simply imagining quasi-worlds. In my conceptions, three lines of thought came together:

- Fractionalization: properties of particles that ordinarily appear only as whole-number multiples of a fundamental unit might occur in smaller multiples within quasi-worlds. Roman Jackiw and Claudio Rebbi abstractly, and Wu-Pei Su, Robert Schrieffer, and Alan Heeger concretely, demonstrated that quasiparticles could carry half a unit of electric charge, i.e., half the charge of an electron.<sup>20</sup> Jeffrey Goldstone and I had shown that in other quasi-worlds different fractions of charge could occur.<sup>21</sup> I wanted to see if a similar fractionalization could happen for angular momentum, that is, spin.<sup>22</sup>
- Flux tubes: I soon realized that fractional angular momentum was indeed possible by means of a very specific physical mechanism: particles orbiting around tubes of magnetic flux. That was an encouraging result, because the theory of flux tubes was already a well-developed, respectable subject. Flux tubes occur in a large class of superconductors, so-called type II superconductors, and in promising, though speculative, unified field theories.
- Dimensional reduction: from a slice of a narrow tube, one can obtain a small, essentially point-like structure that can be considered a particle. Thus, the calculated behavior of narrow tubes in three-dimensional space could be used to construct new kinds of particles in two-dimensional space.

When I gave a seminar about these ideas at Caltech, my friend and colleague John Preskill reminded me that there is a deep connection between the spin of a particle and its quantum statistics.<sup>23</sup> If there is fractional spin, he asked, shouldn't there also be fractional statistics? This was a question I had not considered. In fact, I didn't see right away what the term fractional statistics could even mean. On the drive back to Santa Barbara, I realized that the right thing to think about was braiding, and that braiding flux tubes would yield behavior that could be interpreted as fractional quantum statistics. Within a few days I pulled my thoughts together in two short papers.<sup>24</sup> Anyons had now acquired a name and a more-or-less plausible, though not yet concrete, path to physical embodiment.

**S**TRANGE THINGS can happen when one exposes two-dimensional droplets of mobile electrons—in a narrow range of densities and held at ultralow temperatures—to strong magnetic fields. Under these conditions, as one varies the strength of the magnetic field, a large family of new states of matter is produced. These states, known as fractional quantum Hall liquids (FQHL) are interrelated, but still distinctly different. Though the required conditions are very special and hard to achieve,

the properties of FQHL states are so new and interesting that they have become a preoccupation for many researchers since FQHLs were first discovered in 1982.

The most direct physical manifestation of FQHL states is the phenomenon referred to as quantized Hall conductance. This can be observed if one puts four leads—A, B, C, and D, in that order—at different points on the boundary of the sample. A voltage is applied across A and C by connecting them through a battery, for example. The current flowing between B and D is then measured. As one varies the magnetic field, the conductance—that is, the ratio of voltage to current—does not always vary continuously. Instead, it holds constant around a series of so-called plateau values.

To the uninitiated this behavior might seem an esoteric curiosity, but to physicists it came as a shocking departure from prior experience and expectations. Each plateau represents a new state of matter—a distinctive quasi-world, with its own quasiparticles. Robert Laughlin shared the 1998 Nobel Prize in Physics for his theoretical elucidation of the phenomenon, together with Horst Störmer and Daniel Tsui, who discovered the effect experimentally.<sup>25</sup> Laughlin's theory remains the foundation of our present understanding.

For the purposes of the current discussion, the most relevant part of Laughlin's theory, originally published in 1983, is his picture of quasiparticles in the FQHL states. In short, one produces a quasiparticle by subjecting the material to the influence of a notional flux tube.<sup>26</sup> The resulting quasiparticles have remarkable properties, that differ from one plateau to another. Notably, their electric charge is a fraction of the electric charge of an electron. In the FQHL state that is easiest to produce, the one-third state, the electric charge of a quasiparticle is one-third of the charge of an electron, or  $e/3$ . There is also a one-fifth state on a different plateau, where the quasiparticles have charge  $e/5$ . There are many other states in which the quasiparticle charges are other fractions.

My colleagues Daniel Arovas and Schrieffer were aware of my interest in the theory of fractional quanta and provided tutorials on the nascent theory of the FQHL. When we came to the part about flux tubes, I was able to teach them something in return: the ideas about anyons, and their realization through flux tubes, that I sketched above. Within a few days, we figured out how to bring those general ideas to bear on the FQHL. In a paper published in 1984, we demonstrated mathematically that the FQHL quasiparticles *are* anyons, in the precise sense that when you move them around each other—that is, when you braid their world-lines—their wave function does exactly what anyon wave functions are supposed to do.<sup>27</sup> At the time, I thought it would be easy to test our prediction experimentally. In the years that followed, many people tried to do just that. But nobody succeeded, due to a variety of technical challenges. It was not until 2020 that two groups were able to obtain decisive results.

### **Experiment One: Levels of Conformity**

Hugo Bartolomei et al. set up a kind of quasi-accelerator within the quasi-world of the  $1/3$  FQHL state.<sup>28</sup> As part of their experiment, they produced channels in the shape of an X through which quasiparticles could flow and injected quasiparticle beams flowing upward at the bottom. At the crossing, quasiparticles could meet and scatter from one another. By studying the output at the top, one obtains information about how the quasiparticles interact, through a process physicists term *scattering*.

The quantum statistics of the quasiparticles affects how they scatter in ways that can be calculated confidently. Since two bosons like to do the same thing, they will have a much-enhanced probability to scatter in the same direction, i.e., to enter the same upper arm. Fermions, on the contrary, will strongly prefer to enter different arms. The observed results, falling in between, fit neither of those expectations. Instead, they align well with predictions derived from the kind of anyon quantum statistics predicted for the  $1/3$  FQHL state.<sup>29</sup>

### **Experiment Two: The Beauty of Braiding**

James Nakamura et al. set up an ingenious arrangement modeled on interferometers, a workhorse tool in optics.<sup>30</sup> The central idea in interferometry is to offer a light beam—or quasiparticle—two alternative paths from source to detector. Influences that alter the balance of the paths then show up as changes in the output.

In their experiment, which was likewise performed in the  $1/3$  FQHL state, a flow of quasiparticles can follow one of two paths that together enclose an island within the sample. From the perspective of a quasiparticle within that island, those two paths differ by a winding.<sup>31</sup> If the quasiparticles are anyons, each island anyon will alter the way the sub-amplitudes for the flowing anyons should be added together, in a predictable way. Thus, anytime an additional anyon appears in the island, the output will suddenly jump, also in a predictable way. This is what they observed.

The beauty of this experiment is how clearly it maps onto the most basic defining characteristic of anyons, namely their response to braiding. The jumps can also be measured accurately, which allows for a clean quantitative comparison with the theoretical predictions. Thankfully, they are in agreement.

Why were Nakamura et al. able to succeed where many others fell short? Over the years there have been steady improvements in the purity of the materials and in the techniques available for setting up tiny currents and measuring them accurately. The crucial innovation added by Nakamura et al. was to surround the sample with a bath of electrons that can move to compensate for inhomogeneities of charge within the sample. The compensation process damps out other forces, while leaving the effect of quantum statistics intact. As a result, the behavior of the quasiparticles

becomes more reproducible, and more clearly dominated by their quantum statistics. The strategy employed in this experiment should be adaptable to other FQHL states, which are predicted to support other kinds of anyons.

### **Experiment Three: Switching from Afar**

Using related ideas and as part of work that has extended over several years, Robert Willett et al. applied interferometry to several FQHL states, including some that are suspected, theoretically, to contain non-abelian anyons.<sup>32</sup> In this scenario, the addition of anyons to the central island, one at a time, changes the behavior of the output in different ways at each step. The anyons create something akin to a sophisticated toggle switch.

As part of their work, Willett et al. collected signals that were consistent with theoretical predictions for anyons. At present, their experiments seem to call for more cautious interpretation than those of Nakamura et al., mainly due to questions about the integrity of the island. That said, there is every reason to think that with further work these delicate experiments will become more clear-cut, while their central conclusions will remain valid.

### **Experiment Four: Engineering Anyons**

To conclude this brief survey of recent experimental results, it is appropriate to mention a rather different kind of endeavor that involves realizing anyons within designed systems, as opposed to in natural states of matter.

Collaborators from Austria, China, and Germany set up ingenious circuits involving a mix of conventional and superconducting electronics that support two different kinds of localized excitations.<sup>33</sup> The circuits were designed so that those two kinds of quasiparticles would exhibit nontrivial mutual statistics—and indeed they do.

This construction is meant to be more than a one-off demonstration. It is part of a program to produce fault-tolerant elements for use in quantum computers, based on the ideas of Alexei Kitaev.<sup>34</sup>

**E**XISTING REALIZATIONS of anyons in fractional quantum Hall states are not an ideal vehicle for detailed studies or possible applications. This is because the realizations require ultrapure materials, ultralow temperatures, and high magnetic fields.

It seems possible that different realizations might be free of these drawbacks. Anyons have been predicted to occur in many other quasi-worlds of two-dimensional matter. Numerical simulations have offered support for these ideas. In some cases, there is also suggestive evidence for the predicted behaviors. Crucial experimental tests have been proposed, but they are technically challenging, and they have not yet been carried out.<sup>35</sup>

Another topic for future research is what happens when many anyons are close together, in the same mate-

rial. Fermion and boson behaviors, as noted previously, have dramatic consequences for many-particle systems. Even simple anyons are predicted to support a rich variety of collective behaviors, including a new mechanism of superconductivity.<sup>36</sup> The behavior of an ensemble of more complex anyons and mixtures containing several kinds of anyons with mutual statistics remains largely unexplored.

SEVERAL ANCIENT Andean civilizations, including the Inca, developed a versatile and nonverbal method to record and process information that served them well for many centuries. *Quipus* are formed from a sequence of colored strings containing knots.<sup>37</sup> Each of the strings is tied at one end to a common cord, so that when the cord is suspended the strings hang down and can be scanned easily. The colors of the strings and the placement of the knots might convey an accounting ledger, a historical chronicle, or a military roster. In the quantum world, one might imagine using braids to represent information in a similar manner.<sup>38</sup> Anyons empower such an approach, because the wave functions of multi-anyon systems store memories of the braids formed by their world-lines.

The memory capacity of the simplest anyons, which only keep track of the total winding, is very limited. But well-designed systems that use more complex anyons, whose wave-function rules bring in non-abelian and mutual statistics, can capture much more detailed information. Such systems are the basis for topological quantum computing.<sup>39</sup>

This way of representing information could have important advantages:

- Capacity: the storage capacity of multi-anyon braids grows exponentially fast as the number of anyons or the length of the braids increases, quickly outstripping the capacity of more conventional memories.
- Parallelism: in weaving a single strand through the others, one is confronted with many choices that can produce many different new braids. From the opposite perspective, a single anyon operates in parallel on the information that the preexisting braids encode.
- Noise immunity: the two preceding advantages are characteristic of quantum computers in general. The most distinctive advantage of anyons is their potential reliability. Anyons store and manipulate information about braids, and braids retain their overall form—their topology—even if they are jostled a bit.<sup>40</sup> Since the main technical challenge in quantum computing is avoiding errors, ensuring reliability is a big deal.

Topological quantum computing is currently an active area of research. Microsoft has made substantial investments in the area, and has put forward a concrete, long-range plan for making it into a practical, largescale technology. The process will be far from easy, but the chal-

lenges appear more technical than fundamental. In this, as on several other fronts mentioned above, anyons will keep physicists fruitfully engaged for years to come.



1. Enrico Fermi, “Sulla quantizzazione del gas perfetto monoatomico,” *Rendiconti Lincei* 3 (1926): 145–49; see also, Paul Dirac, “On the Theory of Quantum Mechanics,” *Proceedings of the Royal Society of London, Series A* 112, no. 762 (1926): 661–77, doi:10.1098/rspa.1926.0133.
2. Wolfgang Pauli, “Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren,” *Zeitschrift für Physik* 31, no. 1 (1925): 765–83, doi:10.1007/bf02980631.
3. Freeman Dyson and Andrew Lenard, “Stability of Matter. I,” *Journal of Mathematical Physics* 8, no. 3 (1967): 423–34, doi:10.1063/1.1705209.
4. Subrahmanyan Chandrasekhar, “XLVIII. The Density of White Dwarf Stars,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 11, no. 70 (1931): 592–96, doi:10.1080/14786443109461710.
5. In general, assemblies containing an even number of fermions behave as bosons, while assemblies containing an odd number of fermions behave as fermions. For example, the most common kind of helium atom, <sup>4</sup>He, contains a nucleus of two protons and two neutrons, surround by two electrons. Thus, it has six fermions in all, and so it is a boson. <sup>3</sup>He atoms contain one neutron fewer, so they are fermions. Reflecting this difference, purified liquids based on <sup>4</sup>He and <sup>3</sup>He behave drastically differently at low temperatures. Their contrasting behavior is widely exploited in modern cryogenics.
6. Satyendra Bose, “Plancks Gesetz und Lichtquantenhypothese,” *Zeitschrift für Physik* 26, no. 1 (1924): 178–81, doi:10.1007/bf01327326; see also Albert Einstein, “Quantentheorie des einatomigen idealen Gases,” *Sitzungsberichte der Preußischen Akademie der Wissenschaften* 1 (1925): 3–14.
7. The so-called color of quarks was originally introduced as a distinction without a difference, to allow quarks to be described as fermions. Years later, within quantum chromodynamics, the color degree of freedom got promoted into a new kind of charge, generalizing electric charge. Indeed, this is the crux of quantum chromodynamics.
8. Jon Leinaas and Jan Myrheim, “On the Theory of Identical Particles,” *Il Nuovo Cimento B* 37, no. 1 (1977): 1–23, doi:10.1007/bf02727953; see also Frank Wilczek, “Magnetic Flux, Angular Momentum, and Statistics,” *Physical Review Letters* 48, no. 17 (1982): 1,144–46, doi:10.1103/physrevlett.48.1144; Gerald Goldin, Ralph Menikoff, and David Sharp, “Representations of a Local Current Algebra in Non-simply Connected Space and the Aharonov-Bohm Effect,” *Journal of Mathematical Physics* 22, no. 8 (1981): 1,664–68, doi:10.1063/1.525110; and Yong-Shi Wu, “General Theory for Quantum Statistics in Two Dimensions,” *Physical Review*

- Letters* 52, no. 24 (1984): 2,103–106, doi:10.1103/physrevlett.52.2103.
9. In general, the amplitude is a complex number, and the probability is its absolute square.
  10. Richard Feynman and Albert Hibbs, *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill, 1965).
  11. Less intuitively but more precisely, according to quantum theory, if we do not measure the position of a particle then it simply does not have a definite position. In this situation of ignorance, we must allow for all virtual possibilities and add up their base amplitudes.
  12. But see below!
  13. To do the counting properly we should count the total winding in, say, the counterclockwise direction. Clockwise windings are then assigned negative numbers. With this understanding, the total winding for two successive histories is the numerical sum of the windings for each history separately.
  14. We can move different parts of the candidate knot through one other, while avoiding collisions, by maneuvering into the fourth dimension. Specifically, if two strands threaten to collide, we can avoid the collision by nudging the strand we want to move a little in the direction that lies outside the three-dimensional space defined by its desired but obstructed direction of motion and the two tangents to the strands.
  15. Wilczek, “Magnetic Flux.”
  16. I will not make a long detour to explain that bit of mathematics. Nothing else in this article makes use of it, so if it’s unfamiliar, you can just skip to the next paragraph without losing the thread.
  17. Chetan Nayak et al., “Non-Abelian Anyons and Topological Quantum Computation,” *Reviews of Modern Physics* 80, no. 3 (2008): 1,083–59, doi:10.1103/revmodphys.80.1083.
  18. In the physics literature, it would be called something like, “constructing effective Lagrangians to predict prospective universality classes.”
  19. The theorists who do the imagining are usually different people from the experimenters who do the making.
  20. Roman Jackiw and Claudio Rebbi, “Solitons with Fermion Number  $\frac{1}{2}$ ,” *Physical Review D* 13, no. 12 (1976): 3,398–409, doi:10.1103/physrevd.13.3398; and Wu-Pei Su, J. Robert Schrieffer, and Alan Heeger, “Solitons in Polyacetylene,” *Physical Review Letters* 42, no. 25 (1979): 1,698–701, doi:10.1103/physrevlett.42.1698.
  21. Jeffrey Goldstone and Frank Wilczek, “Fractional Quantum Numbers on Solitons,” *Physical Review Letters* 47, no. 14 (1981): 986–89, doi:10.1103/physrevlett.47.986.
  22. The angular momentum of particles ordinarily is a whole-number multiple of a minimal unit that is equal to one half of Planck’s constant.
  23. According to the spin-statistics theorem, which is one of the profound general consequences of relativistic quantum field theory, particles whose spin is an odd multiple of the basic unit are fermions, while those whose spin is an even multiple of the basic unit are bosons. Of course, that theorem, as stated, only applies to particles that live in three-dimensional space. In two dimensions there still is a tight connection between quantum statistics and spin, but the spin is less constrained, and so is the statistics.
  24. Frank Wilczek, “Remarks on Dyons,” *Physical Review Letters* 48, no. 17 (1982): 1,146–49, doi:10.1103/physrevlett.48.1146; and Frank Wilczek, “Magnetic Flux.”
  25. Robert Laughlin, “Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations,” *Physical Review Letters* 50, no. 18 (1983): 1,395–98, doi:10.1103/physrevlett.50.1395; and Daniel Tsui, Horst Störmer, and Arthur Gossard, “Two-Dimensional Magnetotransport in the Extreme Quantum Limit,” *Physical Review Letters* 48, no. 22 (1982): 1,559–62, doi:10.1103/physrevlett.48.1559.
  26. In this process, the final magnitude of the flux is a critical value called a flux quantum. At this value, the flux tube becomes invisible to the electrons. Technically, it can be compensated by a so-called large gauge transformation. Nevertheless, a quasiparticle is left behind, as a scar from the process of turning the flux tube on.
  27. Daniel Arovas, J. Robert Schrieffer, and Frank Wilczek, “Fractional Statistics and the Quantum Hall Effect,” *Physical Review Letters* 53, no. 7 (1984): 722–23, doi:10.1103/physrevlett.53.722. At about the same time, Bert Halperin independently inferred that FQHL quasiparticles are anyons using less direct arguments. Bert Halperin, “Statistics of Quasiparticles and the Hierarchy of Fractional Quantized Hall States,” *Physical Review Letters* 52, no. 18 (1984): 1,583–86, doi:10.1103/physrevlett.52.1583.
  28. Hugo Bartolomei et al., “Fractional Statistics in Anyon Collisions,” *Science* 368, no. 6,487 (2020): 173–77, doi:10.1126/science.aaz5601.
  29. Bernd Rosenow, Ivan Levkivskiy, and Bertrand Halperin, “Current Correlations from a Mesoscopic Anyon Collider,” *Physical Review Letters* 116 (2016): 156801, doi:10.1103/physrevlett.116.156802.
  30. James Nakamura et al., “Direct Observation of Anyonic Braiding Statistics,” *Nature Physics* 16 (2020): 931–36, doi:10.1038/s41567-020-1019-1.
  31. More precisely, if a quasiparticle takes one path from source to detector, and then returns by the other path, it will have wound once around a quasiparticle within the island.
  32. Robert Willett et al., “Magnetic-Field-Tuned Aharonov–Bohm Oscillations and Evidence for Non-Abelian Anyons at  $\nu = 5/2$ ,” *Physical Review Letters* 111, no. 18 (2013), doi:10.1103/physrevlett.111.186401.
  33. Roughly speaking, they involve concentrations of charge and of circulating current. Chao-Yang Lu et al., “Demonstrating Anyonic Fractional Statistics with a Six-Qubit Quantum Simulator,” *Physical Review Letters* 102 (2009): 030502, doi:10.1103/physrevlett.102.030502.
  34. Han-Ning Dai et al., “Four-Body Ring-Exchange Interactions and Anyonic Statistics within a Minimal Toric-Code Hamiltonian,” *Nature Physics* 13, no. 12 (2017): 1,195–1,200, doi:10.1038/nphys4243; and Alexei Kitaev, “Fault-Tolerant Quantum Computation by Anyons,” *Annals of Physics* 303, no. 1 (2003): 2–30, doi:10.1016/S0003-4916(02)00018-0. For

- more about this, see the discussion of “Noise Immunity” in the next section below.
35. Siddhardh Morampudi et al., “Statistics of Fractionalized Excitations through Threshold Spectroscopy,” *Physical Review Letters* 118, no. 22 (2017), doi:10.1103/physrevlett.118.227201.
  36. Martin Greiter and Frank Wilczek, “Heuristic Principle for Quantized Hall States,” *Modern Physics Letters B* 4, no. 16 (1990): 1,063–69, doi:10.1142/s0217984990001331.
  37. See the excellent article “Quipu” in *Wikipedia*.
  38. It is entertaining and enlightening to do a Google search on “braids” and then choose the Images results. You’ll quickly appreciate that braids can get very complicated, very fast.
  39. Torsten Karzig et al., “Scalable Designs for Quasiparticle-Poisoning-Protected Topological Quantum Computation with Majorana Zero Modes,” *Physical Review B* 95, no. 23 (2017), doi:10.1103/physrevb.95.235305.
  40. Elaborate braided hairstyles would be hopeless to maintain if they could be undone by passing breezes or nodding of the head.  
DOI: 10.37282/991819.21.64

*From the editors of Inference*

# THE RAMBLER

---

A CURATED FEED, UPDATED DAILY



*Articles of interest from across  
the scientific community*

---

INFERENCE-REVIEW.COM/THE-RAMBLER

FOLLOW @RAMBLER\_FEED

REVIEW ESSAYS

# *On Aspects of the Theory of Syntax*

*Anna Maria Di Sciullo*



*Published on August 11, 2021*

*This essay is the first part of a series on classic texts that have come to be seen as landmark achievements in their fields.*

NOAM CHOMSKY PUBLISHED *Aspects of the Theory of Syntax* in 1965.<sup>1</sup> The publication of *Syntactic Structures* in 1957 had already sounded like the roll of distant thunder. A natural language could be studied at the level of explicitness and rigor common in mathematical logic.<sup>2</sup> A revolution was in prospect.<sup>3</sup> Having heard thunder, linguists were eager to see lightning. They were not disappointed. *Aspects* consolidated the revolution. Old-fashioned linguists and behavioral psychologists were scattered into exile.

In undertaking a revolution, Chomsky did what revolutionaries often do. He created his own predecessors, Plato and René Descartes among them. Reviving the notion of Universal Grammar from the seventeenth-century Port-Royal grammarians, Chomsky argued that since every human child could learn any human language, a single abstract grammatical system must be the common property of the human race. *Syntactic Structures* had offered linguists a theory in the sense understood by the serious sciences. In *Aspects*, the offer was carried forward and justified. Writing almost thirty years later, David Pesetsky struck just the right note:

The linguistic capacity of every human being is *an intricate system* [emphasis added], full of surprises but *clearly law-governed* [emphasis added], in ways that we can discern by *scientific investigation* [emphasis added]. Though we still have much to learn about this system, a great deal has been discovered already.<sup>4</sup>

These are ideas that, in *Aspects*, Chomsky compelled some linguists to accept: that many have accepted them is a measure of the book's importance.

A DISCUSSION OF HUMAN creativity typically proceeds from a handful of examples: Aristotle, William Shakespeare, Isaac Newton, Wolfgang Amadeus Mozart, Albert Einstein. Whatever the list, and no matter its length, it embodies the assumption that human creativity is in short supply. All honor to the geniuses, if only because they are rare. Noam Chomsky's very greatest contribution to thought has involved turning this assumption on its head. Human creativity is an acquisition of the species, the common property of the human race. By virtue of having mastered a natural language—Pesetsky's intricate system—every human being is in possession of a rich, complex, and creative system of thought.

In *Syntactic Structures*, Chomsky identified creativity with the recursive structure of a natural language. The human faculty of language is unbounded in precisely the way that the natural number system is unbounded. It is always possible to extend a sentence, as when *the cat is on the mat* is enlarged to encompass *John believes that the cat*

*is on the mat*, and it is possible to do this without obvious limit. In making this possibility the gravamen of his concerns, Chomsky revived Wilhelm von Humboldt's view that language “must make infinite use of finite means.”<sup>5</sup> If this is what language does, until the development of the theory of recursive functions in the first four decades of the twentieth century, no one knew how it was done. Chomsky had read and studied the masters: Kurt Gödel, Alan Turing, Alonzo Church, and, above all, Emil Post.<sup>6</sup> They gave him a theory, and in *Syntactic Structures* he made use of it.

He was the first linguist to do so.

Following the publication of *Syntactic Structures*, Chomsky enlarged this idea of linguistic creativity by appealing to his Cartesian camouflage: “[O]ne fundamental contribution of what we have been calling ‘Cartesian linguistics,’” he wrote,

is the observation that human language, in its normal use, is free from the control of independently identifiable external stimuli or internal states and is not restricted to any practical communicative function, in contrast, for example, to the pseudo language of animals.<sup>7</sup>

This is a large and dramatic claim because it assigns to the ordinary use of language an aspect of human freedom. Thoughts and their expression in language are inclined by circumstances but they are not impelled by them: they are free *both* from “the control of independently identifiable external stimuli” and “internal states.” If this is a claim with overwhelming intuitive plausibility, there is no underestimating its radical nature. It exalts human creativity, but in doing so, places it beyond the scope of the physical sciences as they are now understood. About this kingdom, as Chomsky recognized, modern science has virtually nothing to say.

THE TRUE AND PROPER object of linguistic theory, Chomsky argued in *Aspects*, is the competence of a native speaker—what he knows and not what he says.<sup>8</sup>

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.<sup>9</sup>

A speaker's performance is compromised by limitations of memory, hesitations, repetitions, and any number of throat clearings or verbal tics. The object of linguistic theory is the generative system that accounts for a native speaker's competence; and not the use of this scheme by systems of parsing and production. This, at once, raised a profound and difficult question: if the performance of

a native speaker—what he says—is compromised in various ways, how might he have acquired the underlying system of rules that makes his performance possible? Linguists find the task very difficult, and it is, even today, by no means complete for any natural language. It is hardly possible that children perform a remarkable inductive feat on being presented with data that are compromised and thus degenerate, and under circumstances that are characterized by what Chomsky, with his gift for memorable formulations, called the poverty of the stimulus. Having posed the problem, Chomsky also proposed its solution:

The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance ... The grammar of a particular language, then, is to be supplemented by a universal grammar that accommodates the creative aspect of language use and expresses the deep-seated regularities which, being universal, are omitted from the grammar itself.<sup>10</sup>

The goal of linguistic theory is to provide a theory to describe any human language by principles general enough to apply to every one of them. Unless such a theory exists, there could be no accounting for the fact that human languages are all learnable.

A generative grammar is a system of rules that assigns structural descriptions to sentences.<sup>11</sup> There is no end to sentences and no end to their structural descriptions. The generative grammar represents the linguist's theory, but it also represents the adult speaker's tacit linguistic knowledge.

It represents both.

ASPECTS PRESENTED LINGUISTS with what, at once, became the Standard Theory. *Syntactic Structures* had already offered the essentials. A grammar of a natural language comprises phrase structure and transformational rules. Phrase structure rules break sentences into constituents, the process ultimately yielding a terminal string in which constituents no longer contain constituents. These rules generate hierarchical structures or phrase markers—tree diagrams, in fact. Transformational rules, on the other hand, map phrase markers onto phrase markers. Transformational rules had, in fact, been introduced by Chomsky's mentor, Zellig Harris, but in *Syntactic Structures* they were, for the first time, embedded in a purely formal context.

In *Aspects*, the ideas found first in *Syntactic Structures* found themselves amplified. The Standard Theory is a computational system. Rules are formal because they are explicitly specified: there is no appeal to meaning. The grammar consists of syntactic, semantic, and phonological components; and in addition it contains, or makes use of, a lexicon, something like a formal dictionary.<sup>12</sup> Syntax is

under the control of phrase structure and transformational rules. Phrase-structure rules are formulated as context-free rewriting rules.<sup>13</sup> A category symbol  $A$ , where  $A$  might designate  $S$  (for sentence), is dissected into a string  $Z$  of one or more symbols:  $A \rightarrow Z / X\_Y$ , where the context afforded by  $X$  and  $Y$  is null. The symbols themselves may represent lexical categories, such as noun ( $N$ ) or verb ( $V$ ); syntactic categories such as sentence ( $S$ ); and syntactic constituents such as noun and verb phrases ( $NP$  and  $VP$ ).

The grammar also contains context-sensitive rules:  $A \rightarrow Z / X\_Y$ , where  $X$  or  $Y$  are not null. These rules serve to insert lexical items into phrase markers.<sup>14</sup> It matters a great deal where they are inserted. *Morris plays lapta* is fine: not so *Lapta plays Morris*. The appeal to context is ineliminable. Context-free and context-sensitive rules generate the phrase marker underlying sentences:  $[_S [_{NP} [N]] [_{VP} V [_{NP} [N]]]]$  is an example drawn down to the level of syntactic categories; and on lexical insertion, there is  $[_S [_{NP} Morris] [_{VP} plays [_{NP} lapta]]]$ . From these phrase markers, it is possible to recover old-fashioned grammatical functions—the fact that *Morris* is the subject of the sentence in which he is playing *lapta*. Functions are treated as two-place relations:  $x$  is the subject of  $y$ . These functional relationships may be seen in plain sight on the phrase marker itself, with one node marking the subject of a sentence, and another, its object. The result is what *Aspects*, in a phrase now famous, called deep structure. Transformational rules then map deep structures onto surface structures—those structures ready to enter the gabble of communication.

Chomsky electrified the community of linguists by persuasively arguing that the surface structures of a natural language are no good guide to its deep structures, and, indeed, the distinction between deep and surface structures was widely appreciated as one of the theory's greatest insights. In insisting on the distinction, and its importance, Chomsky appealed to brilliantly chosen examples. In *Syntactic Structures*, he had introduced the now famous sentence *Colorless green ideas sleep furiously* in order to demonstrate that there exist perfectly grammatical English sentences that don't mean a thing. It followed that syntax and semantics were independent; a large conclusion derived from a small example. In *Aspects*, examples multiplied. The sentences *John is easy to please* and *John is eager to please* are on their surface very similar, differing as they do in only one word and otherwise conforming to the same grammatical pattern: NP Cop Adj to VP. Appearances are misleading. These sentences are not at all similar. From *John is easy to please* it follows that it is easy to please John, but nothing like this follows from *John is eager to please*. On the other hand, John's eagerness to please follows from the fact that John is eager to please, but there is nothing like John's easiness to please, even though it *is* easy to please John. These two sentences are radically different. It is on the level of deep structure that these differences are evident. In arguing in this way with

respect to a great many examples, Chomsky was making specific points, but he was also doing more. He was introducing linguists to a new *style* of argument.

**R**ECURSION FIGURED PROMINENTLY in *Syntactic Structures*. Syntactic rules can refer back to themselves and thus may apply to their own outputs. In *Aspects*, sentences themselves became objects of recursive looping and replaced certain transformational rules. This was a major technical development. A sentence (S) may be dissected into a noun phrase and a verb phrase

$$S \rightarrow NP VP.$$

Well and good. A noun phrase may now be dissection into a noun phrase *and* a sentence

$$NP \rightarrow NP S.$$

A verb phrase may then be dissected into a verb and a noun phrase

$$VP \rightarrow V NP.$$

And in view of  $NP \rightarrow NP S$ , into a verb, a noun phrase, *and* a sentence in virtue of  $NP \rightarrow NP S$ . This makes possible the generation of structures such as

$[_s \text{ John } [_s \text{ who met Mary}] \text{ knows Sue}]$ ,

as well as

$[_s \text{ the linguist } [_s \text{ that met the mathematician } [_s \text{ that knows the student } [_s \text{ that... } ]]$

The introduction of sentential recursion, with S hanging on for dear life from both sides of a phrase marker, introduced a notable economy into the Standard Theory. *Syntactic Structures* had handled the matter by hand, inserting sentential phrasemakers in other sentential phrase markers. Fewer symbols were now required, the derivation of complex clauses simplified, the theory streamlined.

With recursion, there is in *Aspects*, a return to the creativity of language:

The infinite generative capacity of the grammar arises from a particular formal property of these categorical rules, namely that they may introduce the initial symbol S into a line of a derivation. In this way, the rewriting rules can, in effect, insert base Phrase-markers in other base Phrase-markers, this process being iterable without limit.<sup>15</sup>

**T**HE STANDARD THEORY offered linguists a formal structure with two different kinds of formal rules—phrase-structure *and* transformational. Recursion got rid of some transformations, but not all. The result-

ing structure is, if not inelegant, then, at least, somewhat clumsy. Why two? Empirical justifications for transformational rules arose from the mismatches between deep and surface structures. The passive voice is an example. In a passive sentence, the logical object of a verbal predicate occurs in the subject position. *John was convinced by Bill to leave* consists of two sentences

(S):  $[_s \text{ John was convinced by Bill } [_s \text{ _ to leave}]]$ .

*John* is the grammatical subject of the main sentence, but not its logical subject, which is *Bill*. On the other hand, *Bill* is not the logical subject of the embedded sentence, which is *John*.

Transformational rules apply from the embedded constituent of a sentence to its outermost constituent. They can insert, erase, substitute, and reorder linguistic constituents. The passive transformation is again an example:

$$NP_1 V NP_2 \Rightarrow NP_2 \text{ be} + V\text{-ed by} + NP_1.$$

This transformation applies to a phrase maker consisting of a nominal constituent  $NP_1$  followed by a verb (V), itself followed by second distinct nominal constituent  $NP_2$ . The transformation specifies the result of this operation:  $NP_1$  and  $NP_2$  are reordered, the auxiliary *be* is added to V as well as the passive morphology *-ed*, and the preposition *by* is added to the postposed  $NP_1$ .

None of this can be handled by phrase structure rules, unless the phrase-structure rules are themselves allowed to increase without limit. If transformational rules are ineliminable within the context of phrase-structure grammars, they seemed, nevertheless, to carry just something of the arbitrary. It is therefore one of the ironies of intellectual history that, far from being purged in theoretical syntax, it has been the other way around, with phrase structure rules themselves dwindling in favor of transformational rules in the minimalist program.

**B**EYOND ITS OBVIOUS contribution to syntactical theory, *Aspects* offered linguists a rich and subtle analysis of old-fashioned grammatical categories—*noun*, *verb*, *adjective*, *adverb*, and the like. Although obviously answering to something, these categories were never clearly defined. A noun was traditionally defined as an expression designating a person, place, or thing. The definition is obviously inadequate. In the sentence *Luck is a great virtue*, “luck” is a noun but not one designating a person, place, or thing. There are many other examples. Making use of a technique first introduced by Roman Jakobson, Chomsky purged these didactic definitions in favor of a scheme in which each syntactic category was flagged by a finite set of binary-valued features. The word *dog* thus enters the lexicon marked as [+N] for noun; the word *barks* by [+V] for verb. Neither [+N] nor [+V] receives any further definition, but they do determine how lexical categories

behave.<sup>16</sup> Their meaning is in their use, as Ludwig Wittgenstein remarked, and their use is governed by their rules, the rules in turn governed by their features. These features serve to discriminate transitive verbs such as *frighten* from intransitive verbs such as *sleep*. Both *frighten* and *sleep* are specified with an inherent [+V] feature: they are both verbs; but *frighten*, contrary to *sleep*, is specified by a trailing [+N]. It takes an object. The introduction of categorical selection rules—what goes where—ensures that verbs like *frighten* are inserted in a phrase marker in the context of a nominal constituent ([+N]), while verbs like *think* are not. *The professor frightens the boy* is grammatical. *The professor thinks the boy* is not.

Chomsky also proposed to distinguish between categorical and semantic selectional features. A verb like *frighten* requires a [+animate] object; not so, a verb such as *praise*. The sentence *The professor frightens sincerity* is grammatical, even though it is semantically deviant, whereas *The professor praises sincerity* is grammatical and otherwise just fine.<sup>17</sup> The introduction of contextual selection rules ensures that *frighten* is inserted in the phrase marker in the structural context of a [+N] [+animate] object.

In developing his theory of syntactic features, Chomsky was heeding methodological constraints: he was responding to the imperative to keep his theory simple. Context-sensitive rules could well be used to settle the distinctions between *frighten* and *sleep*, but only by adding complexity to the grammar. The introduction of syntactic features is one of the most important contributions of *Aspects*.<sup>18</sup> It leads to one of Chomsky's boldest and most dramatic conclusions. The lexicon of a natural language, with its constituents flagged by various syntactic, semantic, and phonological features, is the very place where one language is unlike another. Beyond the lexicon, every human language is governed by the same structures of universal grammar, and in this sense, Chomsky argued, there is only one human language.

One human language! This is surely among the most provocative and dramatic claims of the last half century.

LINGUISTIC THEORY AIMS to derive linguistic facts from first principles, an ultimate goal linguistics shares with science. What would these principles be for language? We point to one universal principle stemming from the Standard Theory: the structure dependency of syntactic rules. Thus S goes over to NP and VP. NP and VP are sister nodes, both structural dependents of S. Ditto for NP → Det N and VP → V NP. The top-down application of the rewriting rules generates structural dependencies between syntactic constituents. The rule governing relative clauses rewrites an NP into an [<sub>NP</sub> NP S]. This rule ignores the linear position of the NP. Relative clauses can be generated both in subject position [<sub>S</sub> *The student of physics* [<sub>S</sub> *who met your advisor*] *is in my class*], and object position, e.g. [*I know the student of physics* [<sub>S</sub> *who met your advisor*]]. A relative clause modifies an

NP and *not* the embedded nominal constituent within that NP. The relative clause [<sub>S</sub> *who met your advisor*] does not modify the nominal constituent [*physics*], even though this nominal constituent immediately precedes it. Structural dependency is a first principle of the language faculty. Linear order is not.

Transformational rules, as defined in the Standard Theory, are structure dependent and they apply to the structural description of a sentence, specify the structural changes, and derive the resulting transformed structure. Transformations may also be associated with conditions on their application. For example, certain transformations apply to main clauses but not to embedded clauses. This is the case for closed yes or no questions. This transformation applies to the underlying structure of sentences such as [<sub>S</sub> *John is here*] and yields the underlying structure [<sub>S</sub> *Is John here*]. Even though these examples seem to indicate that this transformation relies on surface linearity, inverting the auxiliary and the immediately preceding nominal constituent, the following example includes a more complex subject: [<sub>S</sub> [<sub>NP</sub> *The professor of John*] *is here*], and illustrates that this transformation is in fact structure dependent. If it were not the case, this transformation could apply to the auxiliary and the immediately preceding nominal constituent *John*, yielding [*The professor of is John here*]. Instead, this transformation applies to the *full* NP structure and yields [<sub>S</sub> [<sub>NP</sub> *the professor of John*] *here*]. It might very well be the case that structure dependency of syntax is rooted in language design and so a first principle of the language faculty.

Why? No one knows.

ASPECTS LEFT OPEN several questions for further inquiry. Alternative hypotheses are considered in *Aspects*, including with respect to the relevant levels of representation, the properties of the syntactic rules, and the principles of Universal Grammar.<sup>19</sup> These questions have been investigated in the course of the development of generative grammar. The discovery that syntactic rules apply across categories led to the elimination of the multiple rewriting rules postulated in *Aspects*, in favor of a general rule schemata in Government and Binding theory. Transformational rules were reduced to two general operations: move NP (displacing nominal constituents), and move *wh*- (displacing operators such as *who*, *what*, *where*, *when*, in open question formation). In the minimalist program,<sup>20</sup> syntactic operations are reduced to Merge (*x*, *y*), where *x* and *y* are two syntactic objects. Current work investigates the consequences of distinguishing Set Merge, a symmetrical operation deriving unordered sets of constituents, from Pair Merge, an asymmetrical operation deriving ordered sets of constituents.

Another interesting question left open in *Aspects* is whether syntactic rules yield the linear order of syntactic constituents, as in the Standard Theory, where *John eats flies*, or whether they leave the constituents they combine unordered, as in the set {John, eats, flies}, which is on

set-theoretical grounds identical to {eats, John, flies}. The minimalist program investigates, and, indeed, champions the second hypothesis. The linearization of syntactic constituents is handled by the phonological component of the grammar. The very deepest operations of the human mind are indifferent to what might appear to be the most fundamental fact about human language—that words follow one another in a particular order. In all of these arguments, a greater, grander argument is always at work. Universal Grammar must account for the rapid emergence of language in the species, and it must account for its rapid acquisition in the individual. Nothing less than radical simplicity can serve either goal.

**B**Y DEFINING THE OBJECT of inquiry of linguistic theory as internal to the mind, linguistic theory led to the creation of a new interdisciplinary field of inquiry devoted to the study of the biological basis of language, the so-called Biolinguistic Program.<sup>21</sup> Recent research confirms the importance of generative grammar for an understanding of the language faculty as a specifically human trait.<sup>22</sup> The language faculty, like other biological systems, is genetically rooted. Under normal conditions, it develops very early in the child without conscious efforts or extensive training. Animals cannot learn a human language, much to their regret and ours. Monkeys can spontaneously master the weakest of finite-state grammars, but they cannot reach the context-free grammars, which are characteristic of human language, and hierarchical structures are, for this reason, beyond them.<sup>23</sup>

Nothing in the neurosciences is yet as subtle and detailed as the Standard Theory, but it has been established that Broca's area supports the processing of syntax. Human beings are programmed to compute linguistic recursion. A part of Broca's area would appear dedicated to complex syntactic structures: Brodmann area 44 is activated for center embeddings, and Brodmann area 45, adapted to movement.<sup>24</sup> Other studies in cognitive neurosciences indicate that the human brain is sensitive to structure-dependent computation when processing language. This is the case for sentence processing as well as for the processing of phrasal constituents.<sup>25</sup> Yet other studies indicate that the brain processes deep structures, largely ignoring their surface form.<sup>26</sup> "Linguistic theory is mentalistic," Chomsky wrote somewhat defiantly, "since it is concerned with discovering a mental reality underlying actual behavior."<sup>27</sup> Linguistic theory is still mentalistic, but step-by-step, research is uncovering its physical roots in the neurophysiology of the human brain.

**A**SPECTS INTRODUCED A revolution within linguistics. The subject has never been the same again. It promoted linguistics into a science, one that accepted the methods and the standards of the serious sciences themselves. It did more. It championed an integrated study of organic systems, an interdisciplin-

ary field of inquiry bridging results from linguistics and other sciences. And it did still more. It achieved what only the most profound of scientific revolutions achieves and that is transformation of what initially seemed outrageous to what currently seems commonplace. Children *do* learn their native language without effort or instruction; a human language *is* a system of dazzling and poorly understood complexity; some things *must* be innate if anything is ever to be acquired; there *is* a distinction between competence and performance; the most robust system of assessment in studying grammar *is* a native speaker's intuitions; and the ability of every human being to use his language for creative means *is* a mystery that we have not penetrated and may never understand.<sup>28</sup>



1. Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge, MA: MIT Press, 1965).
2. The mathematical basis of generative grammar was published in different articles including Noam Chomsky, "Three Models for the Description of Languages," *IEEE Transactions on Information Theory* 2, no. 3 (1956): 113–24, doi:10.1109/tit.1956.1056813; and George Miller and Noam Chomsky, "Finitary Models of Language Users," in *Handbook of Mathematical Psychology*, vol. 2, ed. Duncan Luce, Robert Bush, and Eugene Galanter (New York: Wiley, 1963), 419–91. See also Thomas Bever, "The Cognitive Basis for Linguistic Structures," in *Cognition and the Development of Language*, ed. John Hayes (New York: Wiley & Sons, 1970), 277–360.
3. Transformational generative grammar stands in contradiction with the structuralist-behaviorist paradigm, prevalent in the first half of the twentieth century. By targeting surface phenomena, structuralist grammars were inevitably drawn to listing exceptions and irregularities instead of capturing language regularities and generalizations. Structuralist grammars were not concerned with Universal Grammar, the human-specific trait enabling the child's ability to develop language naturally, and they endorsed the behaviorist view of language according to which language is acquired by general mechanisms such as induction, analogy, training, and reinforcement. See Chomsky's famous review: Noam Chomsky, "Review of Skinner's Verbal Behavior," *Language* 35 (1959): 26–58, doi:10.2307/411334. See also Eric Lenneberg, "On Explaining Language," *Science* 164, no. 3,880 (1969): 635–43, doi:10.1126/science.164.3880.635.
4. David Pesetsky, "Forecast: Sunny with Scattered Annoyances, but with a Chance of Storms (Recommended Action: Very Basic Linguistics Education)," paper presented at the conference *Generative Syntax in the Twenty-First Century: The Road Ahead*, Athens, Greece, May 28–30, 2015, 1.
5. Wilhelm von Humboldt, "*Ueber das vergleichende Sprachstudium in Beziehung auf die verschiedenen Epochen der Sprachentwicklung* (On the Comparative Study of Language

- and its Relation to the Different Periods of Language Development),” in *Gesammelte Schriften*, vol. 7 (Berlin: Behr, 1907), 98–99. Translation by the editors.
6. For a detailed examination of Post’s work and significance, see Allyn Jackson, “Emil Post: Psychological Fidelity,” *Inference: International Review of Science* 4, no. 2 (2018), doi:10.37282/991819.18.48.
  7. Noam Chomsky, *Cartesian Linguistics* (Cambridge, UK: Cambridge University Press, 2009), 76.
  8. In B. F. Skinner’s *Verbal Behavior*, the distinction is missing, one reason that Chomsky dismissed his theories with disdain. To accept the distinction is almost at once to reject behaviorism in psychology. B. F. Skinner, *Verbal Behavior* (Hoboken, NJ: Prentice-Hall, Inc., 1957).
  9. Chomsky, *Aspects of the Theory of Syntax*, 3.
  10. Chomsky, *Aspects of the Theory of Syntax*, 4, 6.
  11. Chomsky writes: “The term ‘generate’ is familiar in the sense intended here in logic, particularly in Post’s theory of combinatorial systems.” Chomsky, *Aspects of the Theory of Syntax*, 9.
  12. Chomsky, *Aspects of the Theory of Syntax*, 84.
  13. A generative grammar is a computational system analogue to an automaton, with the capacity to read, write, and display information. Chomsky defines a hierarchy of formal grammars, ranked according to their increasing generative complexity (Chomsky, “Three Models”). Each grammar is associated with an automaton of equivalent capacity. It has been established that the generative capacity of a grammar to describe human language must at least have the capacity of phrase structure grammars. See also Noam Chomsky, “On Certain Formal Properties of Grammars,” *Information and Control* 2, no. 2 (1959): 137–67, doi:10.1016/s0019-9958(59)90362-6; Noam Chomsky, “Formal Properties of Grammars,” in *The Handbook of Mathematical Psychology*, vol. 1, 323–418.
  14. For example, verbs such as *see* and *meet* select an NP object and an animate NP subject, which is not the case for verbs such as *grow*.
  15. Chomsky, *Aspects of the Theory of Syntax*, 142.
  16. In Chomsky’s essay “Remarks on Nominalization,” the binary syntactic features  $[\pm N]$  and  $[\pm V]$  are used to define the major syntactic categories, N:  $[+N, -V]$ , V:  $[-N, +V]$ , ADJ:  $[+N, +V]$ , P:  $[-N, -V]$ . This allows for the identification of natural classes of categories, that is, categories undergoing the same syntactic operations, by using a formal property of binary feature systems. See Noam Chomsky, “Remarks on Nominalization,” in *Readings in English Transformational Grammar*, ed. Roderick Jacobs and Peter Rosenbaum (Waltham, MA: Ginn, 1970): 184–221.
  17. The autonomy of syntax with respect to semantics is established in *Syntactic Structures* on the basis of Chomsky’s example *colorless green ideas sleep furiously*. While being semantically deviant, such sentences are generated by the grammar of English. This is not the case for ungrammatical sentences such as *\*colorless sleep furiously ideas green*.
  18. See, among other works, David Pesetsky and Esther Torrego, “The Syntax of Valuation and the Interpretability of Features,” in *Phrasal and Clausal Architecture*, ed. Simin Karimi, Vida Samiian, and Wendy Wilkins (Amsterdam: John Benjamins, 2007), 262–94.
  19. Another question left open in *Aspects* is the distribution of labor between components of the grammar, including the lexicon, the syntax, and the semantic components. As mentioned previously, alternative hypotheses are considered in *Aspects*, along with their consequences. These alternatives have been investigated in the course of the development of generative grammar, giving rise to further questions and problems to solve. One question is whether derived nominals, such as *destruction* and *refusal* are derived by syntactic rules, on a par with gerundive nominals, such as *destroying* and *refusing*, or subject to combinatorial rules in the lexicon or in a distinct morphological component of the grammar. See, for example, Chomsky, “Remarks on Nominalization”; Mark Aronoff, *Morphology by Itself* (Cambridge, MA: MIT Press, 1996); Anna Maria Di Sciullo and Edwin Williams, *On the Definition of Word* (Cambridge, MA: MIT Press, 1987); Morris Halle and Alec Marantz, “Distributed Morphology and the Pieces of Inflection,” in *The View from Building 20*, ed. Ken Hale and Samuel Keyser (Cambridge, MA: MIT Press, 1993), 111–76.
  20. Noam Chomsky, *The Minimalist Program* (Cambridge, MA: MIT Press, 1995). Noam Chomsky, Angel Gallego, and Dennis Ott, “Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges,” *Catalan Journal of Linguistics* (2019), doi:10.5565/rev/catjl.288.
  21. Noam Chomsky, “On Minds and Language,” *Biolinguistics* 1 (2007): 9–27; Massimo Piattelli-Palmarini, Juan Uriagereka, and Pello Salaburu, *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country* (Oxford: Oxford University Press, 2009), 469–72; Anna Maria Di Sciullo et al., “The Biological Nature of Human Language,” *Biolinguistics* 4, no. 1 (2010): 4–34; Robert Berwick and Noam Chomsky, “The Biolinguistic Program: The Current State of Its Development,” in *The Biolinguistic Enterprise*, ed. Anna Maria Di Sciullo and Cedric Boeckx (Oxford: Oxford University Press, 2011), 19–41; Anna Maria Di Sciullo and Lyle Jenkins, “Biolinguistics and the Human Language Faculty,” *Language* 92, no. 3 (2017): e205–36, doi:10.1353/lan.2016.0056; Noam Chomsky, “The Language Capacity: Architecture and Evolution,” *Psychonomic Bulletin and Review* 24, no. 1 (2017): 200–203, doi:10.3758/s13423-016-1078-6; Anna Maria Di Sciullo, ed., *Biolinguistics: Critical Concepts in Linguistics*, vols. 1–4 (New York: Routledge, Taylor and Francis, 2017).
  22. Karin Stromswold, “Genetics and the Structure, Acquisition and Evolution of Language,” paper presented at *Biolinguistic Investigations*, Santa Domingo, Dominican Republic, February 24, 2007; Karin Stromswold, “The Genetics of Speech and Language Impairments,” *New England Journal of Medicine* 359, no. 22 (2008): 2,381–83; Karin Stromswold, “Genetics and the Evolution of Language: What Genetic Studies Reveal about the Evolution of Language,” in *The Evolution of Human Language: Biolinguistic Perspectives*, ed. Richard Larson, Viviane Déprez, and Hiroko Yamakido

- (Cambridge, MA: Cambridge University Press, 2010), 176–90, doi:10.1017/CBO9780511817755.013.
23. See, among other works, W. Tecumseh Fitch and Marc Hauser, “Computational Constraints on Syntactic Processing in a Nonhuman Primate,” *Science* 303, no. 5,656 (2003): 377–80, doi:10.1126/science.1089401.
  24. Michiru Makuuchi et al., “Segregating the Core Computational Faculty of Human Language from Working Memory,” *Proceedings of the National Academy of Sciences* 106, no. 20 (2009): 8,362–67, doi:10.1073/pnas.0810928106; Yosef Grodzinsky and Andrea Santi, “Working Memory and Syntax Interact in Broca’s Area,” *NeuroImage* 37, no. 1 (2007): 8–17, doi:10.1016/j.neuroimage.2007.04.047; Yosef Grodzinsky and Andrea Santi, “fMRI Adaptation Dissociates Syntactic Complexity Dimensions,” *NeuroImage* 51, no. 4 (2010): 1,285–93, doi:10.1016/j.neuroimage.2010.03.034.
  25. Andrea Moro et al., “Syntax and the Brain: Disentangling Grammar by Selective Anomalies,” *NeuroImage* 13, no. 1 (2001): 110–18, doi:10.1006/nimg.2000.0668; Esti Blanco-Elorrieta and Liina Pykkänen, “Composition of Complex Numbers: Delineating the Computational Role of the Left Anterior Temporal Lobe,” *NeuroImage* 124 (2016): 194–203, doi:10.1016/j.neuroimage.2015.08.049.
  26. Christos Pliatsikas et al., “Processing of Zero-Derived Words in English: an fMRI Investigation,” *Neuropsychologia* 53 (2014): 47–53, doi:10.1016/j.neuropsychologia.2013.11.003. See also Angela Friederici, *Language in Our Brain: The Origins of a Uniquely Human Capacity* (Cambridge, MA: MIT Press, 2020), doi:10.7551/mitpress/9780262036924.001.0001, for a summary of the results on the neurobiological foundations of language indicating that species-specific brain differences may be at the root of the human capacity for language.
  27. Chomsky, *Aspects of the Theory of Syntax*, 4.
  28. I gratefully acknowledge lively feedback during the preparation of this essay from David Berlinski.

DOI: 10.37282/991819.21.35

REVIEW ESSAYS

# Primate Memory

*Tetsuro Matsuzawa*



*Published on October 11, 2021*

*My life as a primatologist and researcher has been spent studying chimpanzees, both in the wild and in the laboratory. Our closest living relatives in the animal world are fascinating creatures. In this essay, I recount some of the most striking aspects of my work to illustrate the development of my views about primates, memory, and the evolution of the human mind.*

**T**HE MOST RECENT common ancestor of chimpanzees and humans lived between five and seven million years ago. This shared heritage became evident when sequencing revealed a 1.2% DNA difference between species.<sup>1</sup> Chimpanzees have a living sister species, bonobos, that is equally closely related to humans.<sup>2</sup> Both chimpanzees and bonobos are found only in Africa; this is also true of gorillas. Chimpanzees and humans shared a common ancestor with gorillas between eight and nine million years ago. Another species of great ape closely related to humans are orangutans, which live in Southeast Asia, mostly on the islands of Borneo and Sumatra.

The family Hominidae consists of four genera: humans, chimpanzees (and bonobos), gorillas, and orangutans. The most recent common ancestors of hominids is thought to have lived around fourteen million years ago. In 2019, the International Union for Conservation of Nature listed no less than 512 living species of primates belonging to seventy-nine genera. These species were previously classified as hominids, gibbons, old-world (Afro-Eurasian) monkeys, new-world (American) monkeys, lemurs, and prosimians. With the exception of humans, all the living species of primates reside in Africa, Asia, and Central and South America.

Japan has its own indigenous monkey species in the Japanese macaque, which are sometimes referred to as snow monkeys. The monkeys found in the Shiga Kogen area of central Japan have become famous as the only species known to bathe in hot springs during winter. Another monkey community living on the tiny island of Kōjima near the southern tip of Japan is known for washing sweet potatoes. To wash away the sand from the potatoes they pick up from the beach, the monkeys carry the potatoes to the shallows and dip them in the sea, where they acquire a salty taste. This behavior was first observed in 1953 and is passed from one generation of monkeys to the next. The sweet-potato washing observed on Kōjima was the first example of cultural behavior found in nonhuman animals.<sup>3</sup>

**O**VER THE COURSE of my career, I have undertaken numerous studies investigating the use of tools by chimpanzees in the wild. In recent years, this research has centered on a community of chimpanzees living in the woodlands around the town of Bossou in the West African country of Guinea. These chimpanzees have developed a unique method of gathering drinking water from tree hollows. The technique involves using their mouths to fold leaves and create a receptacle for the water.<sup>4</sup> The appearance of the folded leaves has

been compared to origami. There are numerous other examples of tool use within the group. The same community use a pair of stones as a hammer and an anvil to crack open oil palm nuts and gain access to the edible kernel.<sup>5</sup> I tried the kernel for myself and found it to be quite tasty. Another example involves the chimpanzees using a twig as a fishing stick to gather safari ants from both the ground and their nests. I also tried eating safari ants and found that they are crunchy but may bite your tongue.

The Bossou chimpanzees have also found a novel approach for scooping up algae floating on the surface of a pond. Tatyana Humle and I discovered this behavior in 1995.<sup>6</sup> After following a group of chimpanzees to a glade of tall grass, we observed a four-year-old female that we had named Fotaiu stepping into the thick vegetation surrounding a pond. She then began manufacturing a scooping tool. The first step was choosing a fern (*Cyclosorus afer*, Polypodiaceae) from among the plants growing in the glade and finding a stalk on the ground. Fotaiu then bit off the distal end of the stalk, using her teeth to shorten it to around fifty centimeters. While holding the stalk in her mouth, she stripped off the leaves along its stem using one hand in a swift, downward motion.<sup>7</sup> Holding the stick between her index and middle fingers, she dipped it into the pond and made a swiveling action with her wrist, moving the stick back and forth to collect the floating algae. She then lifted the stick to her mouth and pulled it sideways through her lips, gathering the algae in her mouth.<sup>8</sup> I have attempted to mimic this approach with mixed results. I also tasted the algae and found it muddy and rather unpalatable.

In 1960, Jane Goodall was the first to observe wild chimpanzees engaged in termite fishing.<sup>9</sup> During the six decades since that first sighting in Gombe, Tanzania, many further examples of tool manufacture and use among chimpanzees have been discovered. It has become clear that each chimpanzee community has its own repertoire of tools and associated behaviors.<sup>10</sup> The Bossou chimpanzees pick up and eat termites as they emerge from their mounds, but have not been observed using tools to extract termites hidden inside these structures. Even though stones, nuts, and algae are all readily available for the chimpanzees in Gombe, there is no evidence that they use stone tools to crack open oil-palm nuts, nor have they been observed engaging in algae scooping.

**I**N MY OWN WORK, I have often described the social learning techniques of chimpanzees as education by master-apprenticeship.<sup>11</sup> Mothers and other adults take on the role of the master. The young chimpanzees in the community learn by carefully observing the behavior of the masters. Observational learning has three important aspects: the master models behavior but does not actively teach it; the apprentice has a strong and intrinsic motivation to copy the behavior; and, importantly, the masters are tolerant toward their apprentices while they learn.

Consider the learning process involved in stone tool use. Despite their best efforts, the first attempts made by infant chimpanzees to crack nuts end in failure. It is not until they are around four or five years of age that young chimpanzees are able to start cracking nuts for themselves. To compensate for this shortcoming, very young chimpanzees are allowed to steal kernels from their mothers. I have observed young infants stealing up to seven kernels from the mother in quick succession. As the infants grow, they steal fewer and fewer kernels. At the same time, their attempts at using stones to crack nuts become more frequent and varied. This demonstrates that the motivation for young chimpanzees to adopt the same behaviors as their mothers and other adults is not solely derived from the need for sustenance.<sup>12</sup>

The Bossou forests are home to around 600 plant species. About a third of these are used by chimpanzees, who eat their fruits and leaves, as well as their flowers, stems, roots, and bark. Confronted with so many different plant species, it may take a long time for young chimpanzees to acquire knowledge about plants and the skills that they need to survive in the forest.<sup>13</sup> As part of my fieldwork in Bossou, I tracked chimpanzees between 6:30 in the morning and 6:30 in the evening. With the help of local assistants, I was able to learn about their daily activities and social life.<sup>14</sup> I was often amazed to discover how well the chimpanzees knew when particular fruits were available and where to find them.

**F**IELD STUDIES ARE helpful to learn more about the cognition and behavior of chimpanzees. But they are less well suited for more detailed investigations of their cognitive functions. How do chimpanzees distinguish red fruits from the surrounding green leaves? How do they memorize their experiences in the forest? Are their cognitive processes similar to those of humans, or completely different? These questions are all best examined in a laboratory setting.

The nineteenth-century physicist and philosopher Gustav Fechner is considered one of the founders of experimental psychology. Fechner was the leading figure in the development of psychophysics, a discipline focused on the relationship between perceptual and physical worlds. The Weber–Fechner laws are named after the physician Ernst Weber and his student Fechner. The perceived heaviness of an object, for example, is not simply proportional to its weight, nor is the relationship linear. Instead, Fechner found that the relationship between the perceived weight and the actual weight is a logarithmic function.

The nature of memory also emerged as one of the central issues in early psychological studies. Hermann Ebbinghaus was the first to apply psychophysics to its experimental study. He invented a memory test using three-letter nonsense syllables such as SIF, PLJ, RIT, and TAS. To find out how much time was needed to memorize these syllables, Ebbinghaus became his own test subject. By analyzing the

results of these tests, he was able to discern the forgetting curve plotting memory decay over time, and identify factors that helped account for shrinking retention.

The knowledge and skills acquired by chimpanzees in their natural habitats are retained in long-term memory. In humans, a distinction is made between two categories of long-term memory: declarative (explicit) memory and procedural (implicit) memory. Declarative memory stores factual and episodic information. Procedural memory makes it possible to recall how to do things, such as swimming.<sup>15</sup> From my own observations in the field, I am persuaded that chimpanzees in the wild possess both categories of long-term memory. A more rigorous experimental approach is needed to determine precisely what kinds of memories are retained by chimpanzees, and how similar or different those might be to the memories of humans.<sup>16</sup>

In parallel with my fieldwork and research, I have long been involved in laboratory testing using involving captive chimpanzees.<sup>17</sup> The Primate Research Institute of Kyoto University (KUPRI) has hosted a long-running study aimed at elucidating the cognitive capacities and intelligence of chimpanzees. The main test subject is a female chimpanzee, Ai.

Ai was born in Africa in 1976. When she arrived at KUPRI, she was approximately one year old.<sup>18</sup> I first met the tiny chimpanzee in the basement of the institute in November 1977. The Ai project began under the leadership of my mentor, Kiyoko Murofushi, in collaboration with Toshio Asano and myself. As the youngest professor on the team, I assumed responsibility for the daily care and testing of chimpanzees. A key feature of the Ai project that has remained constant is a simplified experimental setting involving a chimpanzee seated in front of an automated device, with little or no direct human involvement.<sup>19</sup>

Although the Ai project was originally conceived as an ape-language study, I hoped it might be possible to use the same language system to study the perceptual world of chimpanzees. My goal was to compare the perceptual worlds of chimpanzees and humans using the same methods and following the same procedures. By establishing chimpanzee psychophysics as a discipline in its own right, I sought to create the basis for a broader study in comparative cognitive science examining the minds of humans and nonhuman animals.<sup>20</sup>

**W**E BEGAN THE project with three young chimpanzees—Ai, Akira, and Mari—when they were between 1.5 and 2.5 years of age.<sup>21</sup> During the experiments, the chimpanzee sat in a test booth with a backlit keyboard configured in a 5 × 7 matrix. Each key was about the size of a fingertip. The keys were illuminated when they were available and kept dark when they were inactive. On April 15, 1978, Ai pressed a key for the first time.

To begin with, the chimpanzees needed to adapt to both the test booth and the keyboard. The first task was to press an illuminated key. The chimpanzees were then prompted

to perform an identity matching-to-sample task. When a sample color—say, red—appeared on the small screen of the projector, the chimpanzees had to press the corresponding red-colored key on the keyboard. The chimpanzees subsequently learned to discriminate between three colors: red, green, and blue. They also learned to discriminate between nine elemental figures, or graphemes: (H)orizontal line, (O)blique line, (W)ave, (S)nake, (B)lock, (D)ot, (L)ozege, (C)ircle, and (R)ectangle.

As part of the project, I came up with the Kyoto University Lexigram system (KUL). By combining two or three graphemes, it becomes possible to create a unique visual pattern that can be used as a symbol. The idea behind KUL comes from the Japanese–Chinese letters used in kanji. Each kanji letter can be broken down into elements known as radicals. An object such as “pencil” is represented in KUL by combining the graphemes for B and R. The color blue, for example, is represented by DC, while black is OLR.

The chimpanzees also learned a symbolic matching-to-sample task. The sample was the real object, shown to them by the experimenter. The objective was to press the key with the corresponding lexigram. Ai required fifty-seven daily sessions to learn the symbols for eight objects: ball, brick, bowl, glass, glove, paper, pencil, and toothbrush. Akira took eighty-three days, and Mari took 104 days. They also learned the reciprocal use of the lexigrams, decoding the lexigrams to match to the real objects. Both the productive use and receptive use of the lexigrams may provide the basis for symbolic thinking.

Each color name was represented by a lexigram. Later in the project, when touch screens became available, the lexigrams were replaced with the kanji characters representing the eleven colors.<sup>23</sup> Each color name was taught using a specific sample color. I tested the categorization of the various colors using 224 pieces of Munsell color chips that differed in hue, saturation, and brightness. Working in the 1960s, Brent Berlin and Paul Kay found that the universality of color terms outweighed cultural differences.<sup>24</sup> Ai’s results showed that the universality of color categorization is a trait shared by both species. Chimpanzees memorize color terms just like humans.

At the age of five, Ai learned that Arabic numerals can represent numbers.<sup>25</sup> She was the first chimpanzee to successfully label numbers. Combining her acquired skills of object and color naming, Ai can assign the label “Red/Pencil/5” when five red pencils are shown to her. Her spontaneous word order preference was either color-object-number or object-color-number; the number was always placed at the last position in the three-word naming schema.

At the age of six and a half, Ai started learning the letters of the alphabet. After some time, she was able to discriminate between all twenty-six uppercase letters.<sup>26</sup> Each of the letters was available on a keyboard, meaning that the chimpanzee keyboard was essentially the same as a regular computer keyboard. In the initial phase of learning, she confused letters such as E and F, D and O, O and Q, V and Y, and M and W. It seems that these shape pairs have a similar appearance for chimpanzees. Once learned, the letters were also used to measure visual acuity. Placed at a distance from the chimpanzee, the letters were gradually reduced in size to determine the threshold of detection. The visual acuity of Ai was found to be 1.5, which is comparable to that of humans with normal eyesight. In a subsequent test, the letters represented the names of individual humans and chimpanzees. The letter A, for example, was allocated to chimpanzee Akira, M was for Mari, and Z was for Matsuzawa. For Ai, recognizing the faces of chimpanzees was easier than those of humans, while the human subjects showed greater difficulty in identifying chimpanzee faces.<sup>27</sup>

At the age of seven and a half, Ai learned the lexigrams for apple (consisting of the graphemes DCR), banana (SLC), carrot (HBC), cabbage (BDL), potato (OSB), and chow (HWS) as part of a symbolic matching-to-sample task. She also learned symbols for six food items. She then completed a constructive matching-to-sample task in which the sample was a lexigram and Ai had to match the sample using the nine graphemes on the keyboard. Finally, she was tested using another task known as symbolic constructive matching-to-sample. When a real apple was shown to Ai as a sample, she had to construct the corresponding lexigram by choosing the graphemes D, C, and R on the keyboard. On the very first trial of the first day of testing, she constructed the lexigram for carrot from the

**Figure 1.**



A sample of the color red is displayed for Ai on the touch-screen. She is about to touch the corresponding kanji letter for “red.” Photo by the author.

At the age of four, Ai expanded her repertoire of learned symbols to include eleven color names: red, orange, yellow, brown, green, blue, purple, pink, white, grey, and black.<sup>22</sup>

nine graphemes. On average, she needed 2.8 attempts to construct the corresponding lexigrams for six food items upon first encounter.<sup>28</sup> The Ai project proved that the chimpanzees have sufficient long-term memory to learn double-segmented symbol systems—making a symbol from scratch, such as food names from graphemes, and combining the symbols into a noun phrase, such as red/pencils/5.

The Ai project developed into an integrated research effort that continues to the present day. Current research efforts are focused on a multigenerational group of twelve chimpanzees. The forty-five-year history of the Ai project has now included several generations of scientists, chimpanzee participants, research topics, and technological methods.<sup>29</sup> It has also emphasized the importance of maintaining a naturalistic physical and social environment for the chimpanzee subjects, including the development of a seminatural living space with fifteen-meter-high climbing towers in an outdoor compound, connections between multiple habitats, and a multigenerational social group of conspecifics.<sup>30</sup>

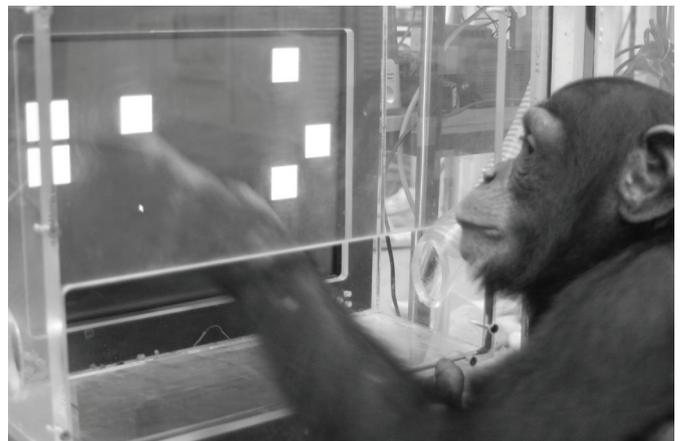
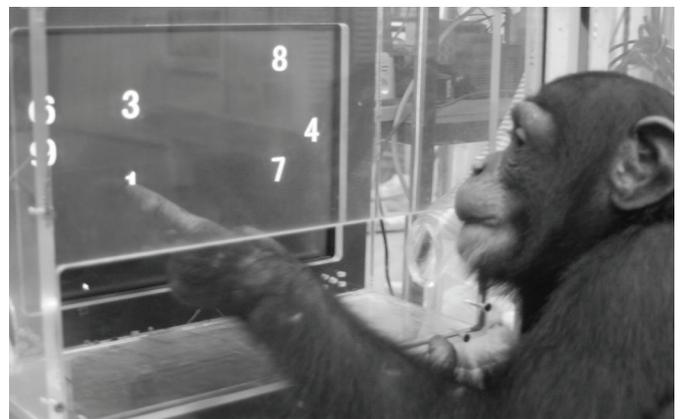
**I**N 2000, AT THE age of twenty-three, Ai gave birth to a son, Ayumu. Ayumu's father was Akira, a long-term friend of Ai. Two other females in the group also gave birth that year. As a result, we were able to study three mother-infant pairs at the same time. Since then our research has been mainly concerned with cognitive development.<sup>31</sup> Studies investigating memory have been the most important aspect of our work involving young chimpanzees. It has been short-term memory, or working memory, rather than long-term memory that has been our main focus.

In tests where three numerals—say, 3, 5, and 8—were presented on a touch screen, all the chimpanzees tested were able to touch the numerals in ascending order. We were amazed by the speed at which they completed the task. A postdoctoral student, Dora Biro, now an Oxford University professor, came up with an interesting short-term memory test that involved switching the second and third numerals immediately after the first was touched.<sup>32</sup> When 3 was touched, for example, the positions of 5 and 8 were swapped. In these tests, the chimpanzees often made the error of touching 3 and then touching 8 instead of 5. The results demonstrated that the chimpanzees had memorized the three numerals at a glance and had planned the order to touch them in advance.

The ability to memorize three numerals at a glance led to another question: How many numerals can chimpanzees memorize in this manner? Human cognition includes a process known as subitizing. For up to five items, we do not need to count them one by one; we can immediately recognize the number of items at a glance. George Miller, one of the founders of cognitive science, was the first to point out human short-term memory was, on average, limited to seven items.<sup>33</sup> A postdoctoral student working on the proj-

ect, Nobuyuki Kawai, now a professor at Nagoya University, proposed a new test that we named the masking task.<sup>34</sup> Kawai's test was a modified version of the swapping trial: when the first numeral was touched, the second and third numerals were immediately masked by white rectangles. When 3 was touched, for example, 5 and 8 were replaced with white rectangles. The goal was for the chimpanzees to touch the white rectangle where 5 had been. The masking task with three numerals can be easily extended to a masking task with four numerals—using a sequence such as 3, 5, 8, 9—and a masking task with five numerals, and so on. Ai's performance on masking tasks involving four and five numerals was comparable to that of human adults.

**Figure 2.**



Ayumu is shown working on the masking task involving seven numerals. As soon as Ayumu touches the first numeral, all the rest are replaced with white squares. Photos by the author.

While conducting these tests, we were amazed by the performance of the three young chimpanzees: Ayumu, Cleo, and Pal, all of whom were four years of age. Their performance on the masking task was much better than that of their mothers. Ayumu was even able to complete a masking

task involving nine numerals.<sup>35</sup> Humans are typically able to complete this task if they spend ten to twenty seconds memorizing the positions of each numeral from 1 to 9.<sup>36</sup>

We then decided to make a direct comparison between the working memory of humans, adult chimpanzees, and young chimpanzees. The masking task was modified slightly for the purposes of this comparison. Sana Inoue, now an associate professor of Ritsumeikan University, and a postdoctoral student at that time, introduced the limited-hold task, in which the numerals were presented to the chimpanzees for only a brief duration.<sup>37</sup> Suppose that there are five numerals—2, 3, 5, 8, and 9—displayed on the touch screen. After 650 milliseconds (ms), 430 ms, or 210 ms, all the numerals are automatically replaced by white squares. The goal is to touch the white squares in the ascending order of the now-masked numerals.

Ai's performance in the limited-hold task was comparable to that of university students facing the same test for the first time. The performances of three young chimpanzees were much better than those of humans. We also tested the impact of overtraining among human subjects, allowing them to repeat the memory test many times over. Although their performances improved with practice, no human has ever been able to match Ayumu's speed and accuracy in touching the nine numerals in the masking task.<sup>38</sup>

One day, a chance event occurred that illustrated the retention of working memory in chimpanzees. While Ayumu was undertaking the limited-hold task for five numerals, a sudden noise occurred outside. Ayumu's attention switched to the distraction and he lost concentration. After ten seconds, he turned his attention back to the touch screen, by which time the five numerals had already been replaced with white squares. The lapse in concentration made no difference. Ayumu was still able to touch the squares in the right order. This incident clearly shows that the chimpanzee can memorize the numerals at a glance, and that their working memory persists for at least ten seconds.

**C**HIMPANZEES ARE ABLE to remember the meaning of visual symbols, such as lexigrams, kanji, letters of the alphabet, and Arabic numerals. Despite this capability, the proto-linguistic capability of chimpanzees, although it clearly exists, remains limited. There have been a number of attempts to teach chimpanzees human methods of communication, such as vocal communication, gestural sign languages, plastic sign languages, lexigrams, and so on.<sup>39</sup> In these studies, chimpanzees have been unable to reach the same level of proficiency as humans. The Ai project proved that although chimpanzees were not good at memorizing arbitrary associations of symbols and their corresponding meanings, they were very good at memorizing items at a glance.

Recent research using eye-tracking technology has revealed some interesting differences between humans and chimpanzees. As part of these studies, no training

or verbal instruction was provided other than requesting that the participants remain seated in front of a monitor. A small visual stimulus, such as a face, is displayed in a random position on the screen and is quickly changed in succession to other images in different positions. The participants observe and spontaneously track the ever-changing stimuli. The studies have shown that chimpanzees possess a remarkable aptitude for switching focus from one stimulus to another. This is not the case for humans. We often stick to one stimulus and are unable to follow the quick change to the next. It seems that our way of perceiving the world is highly skewed toward focusing on a stimulus to determine its meaning. By contrast, chimpanzees are able to shift their attention to capture the whole scene as quickly as possible.<sup>40</sup>

In 2013, I proposed the cognitive tradeoff theory of language and memory.<sup>41</sup> Our most recent common ancestor with chimpanzees may have possessed an extraordinary chimpanzee-like working memory, but over the course of human evolution, I suggested, we have lost this capability and acquired language in return.<sup>42</sup> Suppose that a creature passes in front of you in the forest. It has a brown back, black legs, and a white spot on its forehead. Chimpanzees are highly adept at quickly detecting and memorizing these features. Humans lack this capability, but we have evolved other ways to label what we have witnessed, such as mimicking the body posture and shape of the creature, mimicking the sounds it made, or vocally labeling it as, say, an antelope.

Our human ancestors are thought to have moved out of the forests and onto the savannah millions of years ago.<sup>43</sup> In doing so, they developed a collaborative way of living characterized by group hunting and male–female division of labor for raising offspring. These arrangements are quite different for chimpanzees, among whom only the mothers raise the offspring. Very little in the way of parental care is provided by the father, the grandparents, or other potential helpers. Female chimpanzees start their pregnancies at around twelve years, with an inter-birth interval of about five years. The lifespan of female chimpanzees can be fifty years or more, but there is no social role for grandmothers in their communities.<sup>44</sup> Humans need to share information for survival; transmission of knowledge, skills, and experience from one generation to the next was essential. Although humans may have lost the power of immediate memory we see in chimpanzees, we have acquired greater powers of imagination.<sup>45</sup> These capabilities provide the fundamental basis for language and the ability to discern, to varying degrees, the thoughts of others. In my view, it is the intrinsic motivation for sharing based on the power of imagination that is the evolutionary driving force that has made us human.



1. The Chimpanzee Sequencing and Analysis Consortium, "Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome," *Nature* 437 (2005): 69–87, doi:10.1038/nature04072; and Kay Prüfer et al., "The Bonobo Genome Compared with the Chimpanzee and Human Genomes," *Nature* 486 (2012), doi:10.1038/nature11128.
2. Takayoshi Kano, *The Last Ape* (Redwood City: Stanford University Press, 1992); Christophe Boesch, Gottfried Hohmann, and Linda Marchant, *Behavioural Diversity in Chimpanzees and Bonobos* (Cambridge: Cambridge University Press, 2002); Brian Hare and Shinya Yamamoto, *Bonobos: Unique in Mind, Brain, and Behavior* (Oxford: Oxford University Press, 2017); and Takeshi Furuichi, *Bonobo and Chimpanzee* (Tokyo: Springer, 2019).
3. For a history of primatology in Japan, see Tetsuro Matsuzawa and William McGrew, "Kinji Imanishi and 60 Years of Japanese Primatology," *Current Biology* 18, no. 14 (2008): R587–91, doi:10.1016/j.cub.2008.05.040; Tetsuro Matsuzawa and Juichi Yamagiwa, "Primatology: The Beginning," *Primates* 59 (2018), doi:10.1007/s10329-018-0672-9. On the topic of hot-spring bathing, see Tetsuro Matsuzawa, "Hot-Spring Bathing of Wild Monkeys in Shiga-Heights: Origin and Propagation of a Cultural Behavior," *Primates* 59 (2018), doi:10.1007/s10329-018-0661-z. On sweet-potato washing, see Satoshi Hirata, Kunio Watanabe, and Masao Kawai, "'Sweet-Potato Washing' Revisited," in *Primate Origins of Human Cognition and Behavior*, ed. Tetsuro Matsuzawa (Tokyo: Springer, 2001), doi:10.1007/978-4-431-09423-4\_24.
4. Cláudia Sousa, Dora Biro, and Tetsuro Matsuzawa, "Leaf-Tool Use for Drinking Water by Wild Chimpanzees (Pan troglodytes): Acquisition Patterns and Handedness," *Animal Cognition* 12 (2009): 115–25, doi:10.1007/s10071-009-0278-0.
5. Tetsuro Matsuzawa, "Field Experiments on Use of Stone Tools by Chimpanzees in the Wild," in *Chimpanzee Cultures*, ed. Richard Wrangham et al. (Cambridge: Harvard University Press, 1994), 351–70; Noriko Inoue-Nakamura and Tetsuro Matsuzawa, "Development of Stone Tool Use by Wild Chimpanzees (Pan troglodytes)," *Journal of Comparative Psychology* 111, no. 2 (1997): 159–73, doi:10.1037/0735-7036.111.2.159; Dora Biro et al., "Cultural Innovation and Transmission of Tool Use in Wild Chimpanzees: Evidence from Field Experiments," *Animal Cognition* 6 (2003): 213–23, doi:10.1007/s10071-003-0183-x; and Susana Carvalho et al., "Chaînes Opératoires and Resource-Exploitation Strategies in Chimpanzee (Pan troglodytes) Nut Cracking," *Journal of Human Evolution* 55 (2008): 148–63, doi:10.1016/j.jhevol.2008.02.005.
6. Humle was an undergraduate student at Edinburgh University at that time. She is now a Reader in Conservation and Primate Behaviour at the University of Kent and the International Primatological Society's Vice President for Conservation.
7. Tetsuro Matsuzawa, "Chimpanzees Foraging on Aquatic Foods: Algae Scooping in Bossou," *Primates* 60 (2019): 317–19, doi:10.1007/s10329-019-00733-0.
8. For footage of this algae-scooping behavior, see "Algae Scooping," *YouTube* video, May 31, 2012.
9. Jane Van Lawick-Goodall, "The Behaviour of Free-Living Chimpanzees in the Gombe Stream Reserve," *Animal Behaviour Monographs* 1, no. 3 (1968): 161–311, doi:10.1016/S0066-1856(68)80003-2; and Jane Goodall, *The Chimpanzees of Gombe: Patterns of Behavior* (Cambridge: Belknap Press, 1986).
10. William McGrew, *Chimpanzee Material Culture: Implications for Human Evolution* (Cambridge: Cambridge University Press, 1992); and Andrew Whiten et al., "Cultures in Chimpanzees," *Nature* 399 (1999): 682–85, doi:10.1038/21415.
11. Tetsuro Matsuzawa et al., "Emergence of Culture in Wild Chimpanzees: Education by Master-Apprenticeship," in *Primate Origins of Human Cognition and Behavior*, ed. Tetsuro Matsuzawa (Tokyo: Springer, 2001), doi:10.1007/978-4-431-09423-4\_28. The idea of education by master-apprenticeship was also introduced by Frans de Waal in his *The Ape and the Sushi Master: Cultural Reflections by a Primatologist* (New York: Basic Books, 2001).
12. For more about the learning processes involved in stone tool use, see "Education by Master-Apprenticeship," *YouTube* video, March 23, 2012.
13. Yukimaru Sugiyama and Jeremy Koman, "A Preliminary List of Chimpanzees' Alimentation at Bossou, Guinea," *Primates* 28 (1987): 133–47, doi:10.1007/BF02382192; and Yukimaru Sugiyama and Jeremy Koman, "The Flora of Bossou: Its Utilization by Chimpanzees and Humans," *African Study Monograph* 13, no. 3 (1992): 127–69, doi:10.14989/68093.
14. Tetsuro Matsuzawa, Tatyana Humle, and Yukimaru Sugiyama, *The Chimpanzees of Bossou and Nimba* (Tokyo: Springer, 2011).
15. Endel Tulving, "Episodic and Semantic Memory," in *Organization of Memory*, ed. Endel Tulving and Wayne Donaldson (New York: Academic Press, 1972), 381–403.
16. Kazuo Fujita and Tetsuro Matsuzawa, "Delayed Figure Reconstruction by a Chimpanzee (Pan troglodytes) and Humans (Homo sapiens)," *Journal of Comparative Psychology* 104, no. 4 (1990): 345–51, doi:10.1037/0735-7036.104.4.345.
17. Tetsuro Matsuzawa, "The Ai Project: Historical and Ecological Contexts," *Animal Cognition* 6 (2003): 199–211, doi:1007/s10071-003-0199-2; and Tetsuro Matsuzawa, "The Chimpanzee Mind: In Search of the Evolutionary Roots of the Human Mind," *Animal Cognition* 12 (2009): 1–9, doi:10.1007/s10071-009-0277-1.
18. These events occurred prior to 1980, when Japan ratified the Convention on International Trade in Endangered Species of Wild Fauna and Flora.
19. Prior to the Ai project, we used a DEC-PDP-8 minicomputer to test learning capabilities among monkeys. Asano created the operating system, and I learned the machine code to operate it. The commands were similar to the nonsense syllables created by Ebbinghaus, such as TAD, DCA, ISZ, and JMS. To run the cognitive tests, we needed to switch on lights for different keys and record the key presses by the chimpanzee. The interface itself was built from transistors,

- resistors, condensers, and electromagnetic relays—integrated circuits were still gaining popularity at the time. The Ai project started out using a PDP-11/V03 minicomputer connected to a keyboard. At the time, floppy disks were replacing paper tape readers for data input.
20. Tetsuro Matsuzawa, *The Perceptual World of a Chimpanzee* (Tokyo: University of Tokyo Press, 1991) (in Japanese).
  21. Toshio Asano et al., “Object and Color Naming in Chimpanzees,” *Proceedings of the Japan Academy, Series B* 58, no. 5 (1982), doi:10.2183/pjab.58.118; and Tetsuya Kojima, “Generalization Between Productive Use and Receptive Discrimination of Names in an Artificial Visual Language by a Chimpanzee,” *International Journal of Primatology* 5 (1984): 161–82.
  22. Tetsuro Matsuzawa, “Colour Naming and Classification in a Chimpanzee (Pan troglodytes),” *Journal of Human Evolution* 14, no. 3 (1985): 283–91, doi:10.1016/S0047-2484(85)80069-5.
  23. Toyomi Matsuno, Nobuyuki Kawai, and Tetsuro Matsuzawa, “Color Classification by Chimpanzees (Pan troglodytes) in a Matching-to-Sample Task,” *Behavioural Brain Research* 148, no. 1–2 (2004): 157–65, doi:10.1016/S0166-4328(03)00185-2; and Camille Pene, Akiho Muramatsu, and Tetsuro Matsuzawa, “Color Discrimination and Color Preferences in Chimpanzees (Pan troglodytes),” *Primates* 61 (2020): 403–13, doi:10.1007/s10329-020-00790-w.
  24. In their book *Basic Color Terms*, Berlin and Kay proposed an evolutionary account for color terms. Brent Berlin and Paul Kay, *Basic Color Terms: Their Universality and Evolution* (Berkeley: University of California Press, 1969).
  25. Tetsuro Matsuzawa, “Use of Numbers by a Chimpanzee,” *Nature* 315 (1985): 57–59, doi:10.1038/315057a0.
  26. Tetsuro Matsuzawa, “Form Perception and Visual Acuity in a Chimpanzee,” *Folia Primatologica* 55 (1990): 24–32, doi:10.1159/000156494.
  27. This is referred to as the other-race effect in human subjects. There is another phenomenon termed the face-inversion effect. Faces presented upside down are harder to recognize than those presented right-side up. Ai showed a reduction or loss of such difficulty in naming inverted faces. Chimpanzees use arboreal habitats, in which they often hang upside down. In such an environment, the recognition of inverted faces may serve an important biological function. See Masaki Tomonaga, Shoji Itakura, and Tetsuro Matsuzawa, “Superiority of Conspecific Faces and Reduced Inversion Effect in Face Perception by a Chimpanzee,” *Folia Primatologica* 61 (1993): 110–14, doi:10.1159/000156737. The face-inversion effect in humans and nonhuman primates remains controversial. See Christoph Dahl et al., “The Face Inversion Effect in Non-Human Primates Revisited – An Investigation in Chimpanzees (Pan troglodytes),” *Scientific Reports* 3 (2013): 2,504, doi:10.1038/srep02504.
  28. Tetsuro Matsuzawa, “Spontaneous Pattern Construction in a Chimpanzee,” in *Understanding Chimpanzees*, ed. Paul Heltne and Linda Marquardt (Cambridge: Harvard University Press, 1989), 252–65.
  29. Chris Martin and Ikuma Adachi, “Automated Methods and the Technological Context of Chimpanzee Research,” in *Chimpanzee in Contexts: A Comparative Perspective on Chimpanzee Behavior, Cognition, Conservation, and Welfare*, ed. Lydia Hopper and Stephen Ross (Chicago: The University of Chicago Press, 2020), 182–207.
  30. Tetsuro Matsuzawa, “WISH Cages: Constructing Multiple Habitats for Captive Chimpanzees,” *Primates* 61 (2020): 139–48, doi:10.1007/s10329-020-00806-5.
  31. Tetsuro Matsuzawa, Masaki Tomonaga, and Masayuki Tanaka, *Cognitive Development in Chimpanzees* (Tokyo: Springer, 2006), doi:10.1007/4-431-30248-4; Masaki Tomonaga et al., “Development of Social Cognition in Infant Chimpanzees (Pan troglodytes): Face Recognition, Smiling, Gaze, and the Lack of Triadic Interactions,” *Japanese Psychological Research* 46, no. 3 (2004): 227–35, doi:10.1111/j.1468-5584.2004.00254.x; Satoshi Hirata and Maura Celli, “Role of Mothers in the Acquisition of Tool-Use Behaviours by Captive Infant Chimpanzees,” *Animal Cognition* 6 (2003): 235–44, doi:10.1007/s10071-003-0187-6; and Misato Hayashi and Tetsuro Matsuzawa, “Cognitive Development in Object Manipulation by Infant Chimpanzees,” *Animal Cognition* 6 (2003): 225–33, doi:10.1007/s10071-003-0185-8.
  32. Dora Biro and Tetsuro Matsuzawa, “Numerical Ordering in a Chimpanzee (Pan troglodytes): Planning, Executing, and Monitoring,” *Journal of Comparative Psychology* 113, no. 2 (1999): 178–85, doi:10.1037/0735-7036.113.2.178.
  33. George Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information,” *Psychological Review* 63 (1956): 81–97, doi:10.1037/h0043158.
  34. Nobuyuki Kawai and Tetsuro Matsuzawa, “Numerical Memory Span in a Chimpanzee,” *Nature* 403 (2000): 39–40, doi:10.1038/47405.
  35. Sana Inoue and Tetsuro Matsuzawa, “Acquisition and Memory of Sequence Order in Young and Adult Chimpanzees (Pan troglodytes),” *Animal Cognition* 12 (2009): 59–69, doi:10.1007/s10071-009-0274-4.
  36. The masking task using Arabic numerals can be a good index of working memory regarding number of numerals and time for remembering. Verbal instruction is minimal: “Please touch the numerals in ascending order.” Arabic numerals are universal, which means that we can test memory capacity in terms of developmental changes, aging, and recovery from disease.
  37. Sana Inoue and Tetsuro Matsuzawa, “Working Memory of Numerals in Chimpanzees,” *Current Biology* 17, no. 23 (2007): R1004–05, doi:10.1016/j.cub.2007.10.027.
  38. Footage of the performance can be viewed online: “Working Memory of Numerals in Chimpanzees,” *YouTube* video, January 21, 2019.
  39. There have been numerous ape-language projects in chimpanzees: chimpanzee Gua by Winthrop Kellogg, “Communication and Language in the Home-Raised Chimpanzee,” *Science* 162, no. 3,852 (1968), doi:10.1126/science.162.3852.423; chimpanzee Washoe by Allen Gardner and Beatrice Gardner, “Teaching Sign Language to a Chimpanzee: A Standardized System of Gestures Provides a Means of Two-Way Commu-

- nication with a Chimpanzee,” *Science* 165, no. 3,894 (1969): 664–72, doi:10.1126/science.165.3894.664; the Washoe project continued by Roger Fouts, “Acquisition and Testing of Gestural Signs in Four Young Chimpanzees,” *Science* 180, no. 4,089 (1973): 978–80, doi:10.1126/science.180.4089.978; chimpanzee Sarah by David Premack, *Intelligence in Ape and Man* (Hillsdale: Lawrence Erlbaum Associates, 1976); chimpanzee Lana by Duane Rumbaugh, *Language Learning by a Chimpanzee: The LANA Project* (New York: Academic Press, 2014); and chimpanzee Nim by Herbert Terrace et al., “Can An Ape Create a Sentence?” *Science* 206, no. 4,421 (1979): 891–902, doi: 10.1126/science.504995, and Sue Savage-Rumbaugh, *Ape Language: From Response to Symbol* (New York: Columbia University Press, 1986). Ape language studies have revealed their cognitive ability to understand other’s mind; see David Premack and Guy Woodruff, “Does the Chimpanzee Have a Theory of Mind?” *Behavioral and Brain Sciences* 1, no. 4 (1978): 515–26, doi:10.1017/S0140525X00076512. Studies reveal physical-causal understanding; see Daniel Povinelli, *Folk Physics for Apes: The Chimpanzee’s Theory of How the World Works* (Oxford: Oxford University Press, 2000). On apes’ gestural communication in their natural habitats, see Catherine Hobaiter and Richard Byrne, “The Meanings of Chimpanzee Gestures,” *Current Biology* 24, no. 14 (2014): 1,596–600, doi:10.1016/j.cub.2014.05.066.
40. Fumihiro Kano and Masaki Tomonaga, “Face Scanning in Chimpanzees and Humans: Continuity and Discontinuity,” *Animal Behaviour* 79, no. 1 (2010): 227–35, doi:10.1016/j.anbehav.2009.11.003; and Fumihiro Kano and Masaki Tomonaga, “Species Difference in the Timing of Gaze Movement between Chimpanzees and Humans,” *Animal Cognition* 14 (2011): 879–92, doi:10.1007/s10071-011-0422-5.
  41. Tetsuro Matsuzawa, “Symbolic Representation of Number in Chimpanzees,” *Current Opinion in Neurobiology* 19, no 1 (2009): 92–98, doi:10.1016/j.conb.2009.04.007.
  42. See “The Cognitive Tradeoff Hypothesis,” *YouTube* video, December 5, 2018.
  43. Tim White et al., “Ardipithecus ramidus and the Paleobiology of Early Hominids,” *Science* 326, no. 5,949 (2009): 64–86, doi:10.1126/science.1175802.
  44. Melissa Emery Thompson et al., “Aging and Fertility Patterns in Wild Chimpanzees Provide Insights into the Evolution of Menopause,” *Current Biology* 17, no. 24 (2007): 2,150–56, doi:10.1016/j.cub.2007.11.033; Kristin Havercamp et al., “Longevity and Mortality of Captive Chimpanzees in Japan from 1921 to 2018,” *Primates* 60 (2019): 525–35, doi:10.1007/s10329-019-00755-8. Long postmenopausal lifespans distinguish humans from all other primates. This pattern may have evolved with mother–child food sharing, a practice that allowed aging females to enhance the fertility of their daughters, thereby increasing selection against senescence. This is the grandmother hypothesis raised by Kristen Hawkes et al., “Grandmothering, Menopause, and the Evolution of Human Life Histories,” *Proceedings of the National Academy of Sciences* 95, no. 3 (1998): 1,336–69, doi:10.1073/pnas.95.3.1336.
  45. Aya Saito et al., “The Origin of Representational Drawing: A Comparison of Human Children and Chimpanzees,” *Child Development* 85, no. 6 (2014): 2,232–46, doi:10.1111/cdev.12319; see also Tetsuro Matsuzawa, “Pretense in Chimpanzees,” *Primates* 61 (2020): 543–55, doi:10.1007/s10329-020-00836-z.

DOI: 10.37282/991819.21.51



REVIEW ESSAYS

# At Lunch with Freeman Dyson

*William Press*



*Published on May 12, 2021*

IN THIS ESSAY, I would like to tell the story of a minor discovery in mathematical game theory that Freeman Dyson and I made in 2011. Dyson was a personal friend and one of the great mathematical physicists of the twentieth century. He died in 2020, at the age of ninety-six. He was famously self-effacing, which is not to say that he lacked an accurate opinion of his own abilities. Freeman would deny that he had done anything at all and then allow friends—or even strangers—to vehemently contradict him. Our discovery was not of that character. It really *was* very minor. The reasons for telling the story now are less about the discovery itself and more about the tendency of scientists to seek lessons in moral philosophy in the least likely of places—high-school algebra, for example.

Imagine that a group of scientists gather to play a kind of terror game. They must propose scenarios that, should they eventuate, would shake their belief in the foundations of their fields. The mathematician’s proposed terror is that a long message, in English, is found to be encoded—in excess of any plausible random probability—somewhere in the first billion digits of pi.<sup>1</sup> The physicist’s terror is that the interaction cross-section of a fundamental particle will have significantly different values when measured in different places on earth, or in the same place at different times.<sup>2</sup> The biologist’s terror is that some feature of the living world will be unexplainable by the principle of natural selection. Within biology’s subspecialty of evolution theory, there is a small area of study known as *evolution of cooperation*. That study, some would say, lies closest to the biologist’s terror. That makes it worth poking at.

IN BIOLOGY, a cooperator is an individual who pays a cost for another individual to receive a benefit. When cooperation is mutually beneficial to two individuals of the same or different species—a condition termed *direct reciprocity*—then it is favored by natural selection. There are other possibilities. In so-called kin selection, an individual’s self-sacrifice may be favored if, on average, it helps another individual in the same gene pool to survive.<sup>3</sup> The unit of survival is understood in this case to be not the individual, but the gene that two individuals share.<sup>4</sup> It is harder to understand why individuals cooperate when defection would be more favorable or when the reciprocity is only indirect.

Suppose that two microbe species, *A* and *B*, both need processed nutrients *a* and *b*. The cooperative state might be that *A* produces *a*, *B* produces *b*, and each secretes a portion of its nutrient for the benefit of the other. But this equilibrium is not evolutionarily stable: a defecting *A* with a mutation that halts its sharing of *a* becomes a free rider, benefitting from *B* without paying the fare. Free riders, avoiding a cost, will tend to take over the population. The evolutionarily stable endpoint is noncooperation, even though cooperation would be better for both species.

Cooperation among humans seems hardest of all to understand. “Humans are the champions of cooperation,”

Martin Nowak has remarked. “From hunter-gatherer societies to nation-states, cooperation is the decisive organizing principle of human society.”<sup>5</sup> In much, if not most, of our cooperation, reciprocity is indirect. To be sure, some people give money to universities in the hope of getting their own children admitted—kin selection—but many more give to charities that are of no direct benefit to themselves or their kin. Many billionaires become philanthropists, but from the standpoint of evolution theory, why is this? A quirk of our culture, maybe? But cultures, too, compete for dominance with other contemporaneous cultures, and by a process akin to natural selection. Are we to understand that generosity is selectively favored? Or are the generous billionaires only transient?

Charles Darwin recognized that cooperation posed a challenge to his theory of natural selection. He described an elegant experiment to ferret out whether the aphid yields its excretion to the ant voluntarily, or involuntarily with the ant as a parasite.<sup>6</sup> He provided a convincing argument that it was the former. Darwin, the consummate naturalist, hated overgeneralized theory. Yet the significant literature on the evolution of cooperation that has flourished in the last fifty years is almost entirely theoretical. Much of it is cast in the formalism of mathematical game theory, a subject that came into existence more than half a century after Darwin’s death in the work of John von Neumann and Oskar Morgenstern. Game theory describes how competing, sentient players, in a well-defined universe of choices and payoffs, may knowingly seek to optimize their own outcomes. Evolution is the blind watchmaker,<sup>7</sup> optimizing only by trial and error. Exactly how the achievable outcomes of evolution correspond to the mathematical optima of game theory is not a settled question.

GO BACK TO MICROBES *A* and *B*, but now promote them to sentience. They become Alice and Bob, who are arrested on suspicion of committing, together, a serious crime. Each has sworn not to betray the other. They are questioned in separate rooms.

“We already have enough evidence to convict you both of a misdemeanor,” the detective says to each, “that will put you away for one year.” Each, separately, says nothing. “But if you defect, rat out your partner and turn state’s evidence,” the detective continues, “we’ll let you go, scot-free. Your partner will get a felony conviction, six years in the state penitentiary.”

“What if we *both* turn state’s evidence?” Alice and Bob each ask.

“Well, I can’t let you both go free,” the detective says. “You’ll each get three years.”

Alice reasons as follows: there are only two possibilities. Either Bob will rat me out, or else he won’t. If he rats, then I’ll get six years—unless I rat also, in which case I’ll get just three years. So, if he rats, I should too. But what if he doesn’t rat? What a chump! I can rat on him and be out today. So, *either way*, I should defect. Bob employs the

same reasoning and defects on Alice. They each get three years. The pair spend the time wishing that they had both kept their promises not to betray each other and escaped with misdemeanor convictions.

The prisoner's dilemma (PD) game, played once, has no direct bearing on evolution. But consider the *iterated* prisoner's dilemma (IPD) game, first posited at the RAND Corporation in the 1950s: Alice and Bob play many rounds of the same game with each other. After following the same logic for a few games, Alice tentatively tries a round of cooperation. In that round, Bob still defects and Alice gets six years. But Bob has now seen Alice's signal. He tries cooperation himself. Alice reciprocates. And, for a string of games, they are both cooperating, receiving only misdemeanor convictions. In the IPD, there is information in the previous plays, and each player can try to use that information to devise a superior strategy that remains self-interested.

Is the best strategy to cooperate always? Certainly not. If Alice adopts that strategy, Bob will always defect, getting off scot-free, while Alice will always get six years. A good strategy would seem to be something like, "Cooperate most of the time, but don't be a chump if the other player doesn't follow suit." Can this be formalized, or made crisp, in some way? By one estimate, more than 200 experiments on IPD, with human or computer players, were conducted between 1965 and 1971. Robert Axelrod called IPD "the *E. coli* of social psychology."<sup>8</sup> Axelrod's own 1980 experiments are the most famous.<sup>9</sup> Human contestants submitted algorithmic strategies that, programmed on a computer, were all played against each other in a tournament. A strategy could be very complex. Alice could, in principle, look at Bob's previous thousand plays and try to crack the code of his decision process. It was a big surprise when a very simple strategy, known as tit for tat (TFT), won the Axelrod tournaments. TFT starts by cooperating. Then, if the other player cooperates, it cooperates on the next round. If the other player defects, then, on the next round, it defects.

In later experiments, a related strategy, generous tit for tat (GTFT), was found to do even better. GTFT is the same as TFT, except that, when a tit-for-tat defection would be prompted, GTFT *sometimes*, and with a fixed probability, cooperates rather than defects—it is generous in that way. There seemed to be moral lessons in these results, pseudo-mathematical justifications of high human values. TFT embodied the Golden Rule. GTFT went further: turn the other cheek.

Science and ethics were in harmony.

I had long ago read Axelrod's book and William Poundstone's popular exposition, and I knew about TFT and GTFT.<sup>10</sup> That was about all I knew. During the slow period between Christmas and New Year's in 2011, I set out to understand why TFT and GTFT did so well. I was struck by the fact that they both were memory-one strategies. That is, a player's move—cooperate or defect—depended

only on the immediately previous round of play. Had anyone ever proved that GTFT, or any other strategy for that matter, was optimal among memory-one strategies? If so, I couldn't find it in Google Scholar.

Without getting into too much detail, every memory-one strategy is defined by four numbers, probabilities in the range between zero and one. A match between two players, each playing a fixed memory-one strategy, is thus described by eight parameters—equivalent to a point inside, or on the surface of, an eight-dimensional hypercube. It is not hard to derive formulas for the statistically expected gain or loss of each player as a function of the match position in the hypercube. With these formulas, it is not necessary to play out the games at every point, a huge saving in computer time.

Game theory has the important concept of Nash equilibrium. As the players each try to improve their strategies, one imagines the match-strategy point moving around in the hypercube. A Nash equilibrium is a point where no small adjustment of Bob's four parameters by Bob improves his score, and the same is true for Alice's four parameters and her score. If the players reach a Nash equilibrium, then neither will see any benefit in changing strategy further. Each sees an optimum that is at least local. Over the holiday, I wrote what I thought was an elegant computer program to explore the game hypercube, always seeking better strategies. I expected it to find that the Nash equilibrium was something like GTFT for both players.

On the computer screen, I watched the progress of the optimization. The match seemed to be getting closer and closer to a Nash equilibrium—until the program crashed. I tried again, starting from a different point in the hypercube. Again, the program crashed, but at a different place. I automated the procedure and ran the program a thousand times—and got a thousand crashes. But they weren't at random points in the eight-dimensional hypercube. All of the crashes occurred on a particular four-dimensional hyperplane. I traced their cause to a faulty assumption in my program. It assumed that when Bob changed his strategy, it would, generically, have some effect on his own score, and similarly for Alice. How could this not be true? Apparently it wasn't. The computer could find instances, but not explain them.

That was when I knew that I needed Freeman Dyson. The exact complement to computer intelligence, as yin to yang, is Freeman Dyson intelligence. This problem needed both kinds. I emailed a description of my puzzling results to Freeman. A day later, he sent back a note with the general result all worked out, which became equations 1 through 7 in our paper.<sup>11</sup> When the confusion was hacked away, it came down to a simple equation in high-school algebra, as Freeman noted. Hidden within the IPD was a matrix whose determinant could be forced, by either Alice or Bob acting alone, to be zero—with quite unexpected implications. We called the resulting strategies zero determinant (ZD).

Overlooked in fifty years of research on IPD was the simple fact that Alice had the ability, by choosing a certain ZD strategy, to set Bob's score. She could pick any value between that of full cooperation (the one-year misdemeanor) and full defection (the three-year felony). No matter how Bob played, that value, on average, would be his outcome. And Bob, correspondingly, could do the same thing to Alice. Game theorists already had a name for this situation: an ultimatum game. They had no idea that there was an ultimatum game hidden inside the IPD.

In the classic ultimatum game, a hundred dollars appears out of nowhere on the table between Alice and Bob. By a coin flip, one player, say Alice, goes first. "I'll take \$60, and you can have \$40," she tells Bob. He can either accept and take the \$40, or else the whole hundred disappears and neither gets anything. What makes it a game is that Alice can pick any number between \$0 and \$100 for her initial offer. What is her optimal strategy? Mathematical game theory doesn't provide an answer; it turns out that all values are Nash equilibria, but not in a useful way. Psychologists and economists have tried the ultimatum game experimentally across a wide range of locations and ethnic groups.<sup>12</sup> The first player never offers more than \$50—why should she?—and the second player rarely accepts less than \$20 or \$30, so a 60:40 split is quite typical. The game has been played between humans, and chimpanzees using raisins, with ambiguous results.<sup>13</sup>

The ultimatum variant revealed by the ZD strategies could be this: "Set my score to the lightest-sentence misdemeanor," Alice tells Bob, "and I'll do the same for you. But if for any reason you cross me, I'll change my strategy to punish you severely." This is reminiscent of TFT, but it is played at the meta-level of altering strategies, not at the game level of individual moves. To see the difference, imagine a scenario in which Alice is facing Darwin rather than Bob. Darwin is the blind watchmaker who can only try, by small mutations in strategy, to maximize his score. He doesn't know that he is playing against a sentient being. Evolutionary biologists speak of the fitness landscape of hills and valleys in which evolution by natural selection can be viewed as taking place. But here, Alice completely controls Darwin's fitness score. She can simulate any fictitious evolutionary landscape she wants. Darwin cannot distinguish it from a natural one. Thanks to ZD, she completely controls the apparent rules of the game. Biologists like to say, "You can't fool evolution," but this example shows that, within the constructed space of IPD, the ZD strategies actually can. That surprised a lot of people.<sup>14</sup>

The ZD strategies had yet more surprises in store. Instead of setting Bob's score to a value, there is a ZD strategy by which Alice can set Bob's score to be related to her own, extortionately. Alice picks an extortion factor: five, for example. She implements the corresponding ZD strategy for her own game, then lets Bob play any strategy he wants. If his strategy reduces his felony sentence by an

amount  $x$ , which we will call his bonus, then Alice's bonus will be  $5x$ . If he optimizes his strategy by blind evolution, then Alice will score better than even if both cooperated—not even a misdemeanor sentence, but only a traffic violation. Bob ends up worse off. There is nothing he can do about it. Alice sets the strategy.

Really? Nothing? Can't he do back to Alice exactly the same thing as she is doing to him? Yes, he can. He can adopt a ZD strategy that sets his bonus to be five times hers. How can each of them have five times the bonus of the other? Easy: both get zero. In effect, a return to noncooperating three-year felony sentences. So, they are back to bargaining, as in the ultimatum game: "If I do this for you, will you do this for me?"

The most interesting result Dyson and I reached was this: that the outcome of the game depended on whether each player had a theory of mind about the other. Psychologists use the term *theory of mind* to mean the ability to attribute mental states—such as belief and intention—to others. When Alice plays against Darwin, Darwin has no theory of mind. Alice can extort him, or lead him around a fictitious fitness landscape, at will. But when Alice plays against Bob, he may have a theory of mind. Noticing that his opponent's bonuses are always five times his own, he thinks, "I am being extorted! I will make us both get zero. She will notice this. Then, we can do a deal." But Bob may be wrong about this. Alice may have set her ZD strategy once and for all, put it on autopilot, and then vanished. Bob, with or without a theory of mind, can try anything he wants. He can only hurt himself. His best option is to accede to the extortion. "Press and I have solved the Prisoner's Dilemma game," Dyson told people jocularly. "The winning strategy is to go to lunch."

THE RESULTS DESCRIBED are so counterintuitive that some readers may want to get a sense of how they are derived. Label the four outcomes of the previous round 1, ..., 4 in the order ( $cc$ ,  $cd$ ,  $dc$ ,  $dd$ ), where  $c$  denotes cooperate,  $d$  denotes defect, and the order is Alice–Bob. Alice's strategy can be written as  $\mathbf{p} = (p_1, p_2, p_3, p_4)$ , her probabilities for cooperating on the current round given each of the four previous outcomes. Similarly, Bob's strategy is  $\mathbf{q} = (q_1, q_2, q_3, q_4)$ .

The strategies  $\mathbf{p}$  and  $\mathbf{q}$  imply a Markov process that advances any mixture of outcomes one round at a time. The Markov transition matrix is

$$M = \begin{bmatrix} p_1 q_1 & p_1(1 - q_1) & (1 - p_1) q_1 & (1 - p_1)(1 - q_1) \\ p_2 q_2 & p_2(1 - q_2) & (1 - p_2) q_2 & (1 - p_2)(1 - q_2) \\ p_3 q_3 & p_3(1 - q_3) & (1 - p_3) q_3 & (1 - p_3)(1 - q_3) \\ p_4 q_4 & p_4(1 - q_4) & (1 - p_4) q_4 & (1 - p_4)(1 - q_4) \end{bmatrix}.$$

Because every Markov matrix has a unit eigenvalue, the matrix  $M' \equiv M - I$  is singular and has zero determinant. The stationary vector, i.e., the long-term probability mix of outcomes of the game, satisfies

$$\mathbf{v}^T \mathbf{M} = \mathbf{v}^T, \text{ or } \mathbf{v}^T \mathbf{M}' = 0.$$

The significance of the stationary outcome probability vector is that its dot product with Alice’s prison sentences  $\mathbf{S}_A = (1, 0, 6, 3)$  or Bob’s  $\mathbf{S}_B = (1, 6, 0, 3)$  gives the expectation value of their respective times in jail. Close attention should be paid to the extent that each may be able to unilaterally influence  $\mathbf{v} \cdot \mathbf{f}$  for a given  $\mathbf{f}$ , where  $\mathbf{v}$  is the stationary vector and  $\mathbf{f}$  is a given four-vector.

Now the promised high-school algebra: Cramer’s rule for calculating a determinant, applied to the matrix  $\mathbf{M}'$ , is

$$\text{Adj}(\mathbf{M}')\mathbf{M}' = \det(\mathbf{M}')\mathbf{I} = 0,$$

where  $\text{Adj}(\mathbf{M}')$  is the adjugate matrix—what most of us in high school learned as the matrix of minors. This equation implies that every row of  $\text{Adj}(\mathbf{M}')$  is proportional to  $\mathbf{v}$ . Choosing the fourth row, we see that the components of  $\mathbf{v}$  are, up to a sign, the determinants of the  $3 \times 3$  matrices formed from the first three columns of  $\mathbf{M}'$ , leaving out each one of the four rows in turn. These determinants are unchanged if the first column of  $\mathbf{M}'$  is added into the second and third columns.

The result of these manipulations is a formula for the dot product of an arbitrary four-vector  $\mathbf{f}$  with the stationary vector  $\mathbf{v}$  as a single  $4 \times 4$  determinant:

$$\mathbf{v} \cdot \mathbf{f} = \det \begin{vmatrix} -1 + p_1 q_1 & -1 + p_1 & -1 + q_1 & f_1 \\ p_2 q_3 & -1 + p_2 & q_3 & f_2 \\ p_3 q_2 & p_3 & -1 + q_2 & f_3 \\ p_4 q_4 & p_4 & q_4 & f_4 \end{vmatrix}.$$

It is now possible to see the remarkable fact that the second column is entirely under Alice’s control (the  $\mathbf{p}$ ’s), while the third column is under Bob’s control. Given a particular  $\mathbf{f}$ , each has the possibility of choosing a strategy that will make the determinant zero.

From here, it is only a short step to the results already mentioned. Alice chooses for  $\mathbf{f}$  a desired linear combination of the score vectors,  $\mathbf{f} = \alpha \mathbf{S}_A + \beta \mathbf{S}_B + \gamma$ , and then calculates a strategy  $\mathbf{p}$  that zeros the determinant. That linear combination of scores is made zero. These are exactly the ZD strategies, including the specializations of extortionate or ultimatum.

Our paper was published in *Proceedings of the National Academy of Sciences* in mid-2012. Despite a simultaneously published, glowing commentary by evolutionary biologists Alex Stewart and Josh Plotkin, few people seemed to care about the theory of mind.<sup>15</sup> What attracted attention was the extortion business. People reacted to it emotionally. Poundstone posited links to abusive marriages, terrorism, the current US Congressional deadlock, and income inequality.<sup>16</sup> Commenting on an article about the paper, a reader remarked, “This and similar ‘quant’ nonsense is precisely what has led to the ... too-big-to-fail banking disaster we are currently confronting. ... These

studies are worse than useless, they are parasitic cancers on society.”<sup>17</sup> Elsewhere, a kinder commenter added, “Once again, physicists invade a field and add value.” Even that seemed barbed.

All this from a simple equation of high-school algebra.

Emotion tainted even the soberest responses to our work. Mathematically speaking, we had shown that a broad class of two-person games—not just IPD—had an unsuspected hidden algebraic structure, our ZD. It turned out that there were other, quite different, hidden structures. Mathematician Ethan Akin discovered a set of strategies with the twin properties that Alice and Bob both got off with minimal misdemeanors, and that neither could gain by unilaterally changing strategy.<sup>18</sup> He named these *good* strategies. Some of our ZD strategies were *good*, but others were not. By implication, they were evil. Stewart and Plotkin wrote an elegant paper focusing on the subset of ZD strategies that were *generous*, meaning that one player voluntarily ceded a greater share of reward to the other.<sup>19</sup> I was relieved that some of our ZD strategies turned out to be both *good* and *generous*, though not including, of course, our extortionate strategies—which, by the way, would still beat all the good and generous strategies in head-to-head competition.

Alice and Bob play against each other over and over, but that is not how populations work. Population biologists study the dynamics and evolution of a mix of strategies, where an Alice must sometimes compete against a Bob, or another Alice, or the mutant Alicia—like Alice, but subtly different. Additional concepts come into play. Is the mutant able to *invade* the population, so that, by its greater fitness over many generations, its Alicias become dominant? The Bobs may have a strategy that is successful when played against the Alicies, but so mutually destructive when played against each other that a population dominated by Bobs is not an *evolutionary stable strategy*. In a finite population, there is a related kind of evolutionary inertia: a favorable mutation must exceed a certain threshold to avoid, on average, dying out just by chance. Bob’s, or Alice’s, strategy may win but still not be evolutionarily robust.

To the seeming delight of many, the extortionate ZD strategies were found to be neither stable nor robust, nor able to invade a population to any significant degree. In a population, extortion was, roughly speaking, self-limiting. Extortioners would tend to mutate into generous players because, most of the time, generous players would be playing against other generous players. Christoph Adami and Arend Hintze later quipped that this proved mathematically that winning isn’t everything.<sup>20</sup> Dyson liked these results. I thought they were interesting and advanced the field, but I was bothered by the emotional coloration that seemed to accompany them. Freeman, I concluded, had gone over to the side of the sentimentalists.

We can ponder whether anything of value can be learned from the events I have described. That a couple of physi-

cists—the more senior, eighty-eight years old—could invade a field and, over a holiday vacation, find an undiscovered nugget capable of attracting such attention may say something about serendipity, or about the genius of Freeman Dyson; or it may suggest that subfields of science can easily become too set in their ways, and that the scientific enterprise should seek institutional mechanisms that encourage more cross-fertilization over scientific boundaries.

Also worth pondering is the human tendency to label scientific findings with emotive words like good, generous, and, yes, extortionate. Most of the time this surely does no harm. It makes the science livelier and helps communication to each other and the public. Occasionally, though, the labels in a field become a mythos that can color it with subjectivity. The application of game theory to evolutionary biology led to a mythos, according to Poundstone, that you can't fool evolution and the most successful strategies are fair ones.<sup>21</sup> Neither assertion is a scientific truth. In the nineteenth century, one particular mythos attached to natural selection was that it wasn't just a description of the way things *were*, but a description of the way things *ought to be*. That led to social Darwinism and its misuse of science in justification of racism, imperialism, eugenics, and other horrors. Evolution by natural selection *is what it is*. Everyone should be on guard against labeling it with either moral virtues or moral failings. A decade before the publication of Darwin's *On the Origin of Species*, Alfred Lord Tennyson's poem "In Memoriam A. H. H." gave a startlingly good description of natural selection. This is the poem where "Nature, red in tooth and claw" ultimately loses out to "... love, Creation's final law." But it is poetry, not science.

If I ever find myself fighting for survival in a population of Alices, Bobs, Darwins, and Alicias, I plan to go with ZD extortion. And I would advise you to do the same—but just not too many of you.



1. Carl Sagan, *Contact* (New York: Simon and Schuster, 1985).
2. Cixin Liu, *The Three-Body Problem*, trans. Ken Liu (New York: Tor Books, 2014).
3. An elegant review by Martin Nowak ("Five Rules for the Evolution of Cooperation," *Science* 314, no. 5,805 (2006), doi:10.1126/science.1133755) compares five mechanisms: direct reciprocity, indirect reciprocity, kin selection, network reciprocity, and group selection.
4. Richard Dawkins, *The Selfish Gene* (New York: Oxford University Press, 1976).

5. Nowak, "Five Rules," 1,560.
6. Charles Darwin, Chapter VIII, in *On the Origin of Species* (London: John Murray, 1872 [1859]).
7. The blind watchmaker metaphor was introduced by William Paley in his 1802 book *Natural Theology* (London: R. Faulder), long before Darwin.
8. William Poundstone, *Prisoner's Dilemma* (New York: Anchor/Doubleday, 1992).
9. Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).
10. Axelrod, *Evolution of Cooperation*. Poundstone, *Prisoner's Dilemma*.
11. William Press and Freeman Dyson, "Iterated Prisoner's Dilemma Contains Strategies that Dominate Any Evolutionary Opponent," *Proceedings of the National Academy of Sciences* 109, no. 26 (2012): 10,409–13, doi:10.1073/pnas.1206569109.
12. Hessel Oosterbeek, Randolph Sloof, and Gijs van de Kuilen, "Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis," *Experimental Economics* 7 (2004): 171–88, doi:10.1023/B:EXEC.0000026978.14316.74.
13. Joseph Henrich and Joan Silk, "Interpretative Problems with Chimpanzee Ultimatum Game," *Proceedings of the National Academy of Sciences* 110, no. 33 (2013): E3049, doi:10.1073/pnas.1307007110.
14. William Poundstone, "On 'Iterated Prisoner's Dilemma Contains Strategies that Dominate Any Evolutionary Opponent' by William Press and Freeman Dyson," *Edge.org*, June 18, 2012.
15. Alexander Stewart and Joshua Plotkin, "From Extortion to Generosity, Evolution in the Iterated Prisoner's Dilemma," *Proceedings of the National Academy of Sciences* 110, no. 38 (2013): 15,348–53, doi:10.1073/pnas.1306246110.
16. Poundstone, "On 'Iterated Prisoner's Dilemma'."
17. Although the original article posted by *MIT Technology Review* in 2012 is no longer available, this response can be found in an archived discussion thread on the website *Hacker News*, dated August 17, 2012. See the comment posted by the user ramblerman.
18. Ethan Akin, "The Iterated Prisoner's Dilemma: Good Strategies and Their Dynamics," *arXiv* (2017 [2012]), arXiv:1211.0969v3.
19. Stewart and Plotkin, "From Extortion to Generosity."
20. Christoph Adami and Arend Hintze, "Evolutionary Instability of Zero-Determinant Strategies Demonstrates That Winning Is Not Everything," *Nature Communications* 4, no. 2,193 (2013), doi:10.1038/ncomms3193.
21. Poundstone, "On 'Iterated Prisoner's Dilemma'."

DOI: 10.37282/991819.21.1

BOOK REVIEWS

# The Secret Lives of DNA

*Neeraja Sankaran*



*Published on June 21, 2021*

*Unravelling the Double Helix: The Lost Heroes of DNA*

by Gareth Williams

Pegasus Books, 528 pp., \$35.00.

**D**NA FIRST CAME to the public's attention in 1953 when, as James Watson later claimed, Francis Crick "winged into The Eagle to tell everyone within hearing distance" that the two of them had discovered the secret of life.<sup>1</sup> This secret was the double helical structure of DNA. The discovery was an achievement in its own right, but it should not be conflated with either the discovery of DNA itself or the discovery of its function as the material carrier of life's instructions.<sup>2</sup> These discoveries were made eighty-five and nine years earlier, respectively. Without knowledge of the existence of DNA and its vital function, no one would have thought to tackle the problem of its structure.

In *Unravelling the Double Helix*, Gareth Williams details the way in which Watson and Crick came up with their double-helix model only in the last quarter or so of his book, which comprises twenty-six chapters running to nearly 500 pages. The first three-quarters of the book traces the stories of several scientists from the mid-nineteenth century onward.<sup>3</sup> All these people were, in one way or another, involved in some aspect of research on DNA and "were variously enthralled, seduced or infuriated" by it.<sup>4</sup>

**I**N A CHAPTER aptly titled "In the Beginning," Williams describes the first seduction enacted by DNA.<sup>5</sup> It began in 1868, in a venue apropos for romance: a medieval castle in the German town of Tübingen. But any romantic images dissipate upon learning what went on inside. The castle had been acquired by the University of Tübingen in the 1830s, and during the icy winter of 1868, in an unheated laboratory occupying the space of its former kitchens, a young Swiss student named Friedrich Miescher was painstakingly washing pus off surgical bandages obtained from a nearby hospital.

Why pus? Unlikely as it may now seem, pus was the most easily obtainable and abundant source of the cells Miescher needed for his investigations. His intention in coming to Tübingen had been to study the chemical composition of cells under the pioneering physiological chemist Felix Hoppe-Seyler, best known for coining the name hemoglobin.<sup>6</sup> Hoppe-Seyler steered Miescher toward white blood cells (leucocytes) as the best source material for his studies of cell chemistry. A nearby surgical clinic provided a steady supply of surgical bandages, the pus in which yielded sufficient quantities of leucocytes for his experiments.<sup>7</sup> From these cells, Miescher extracted a phosphate-rich substance that seemed distinct from other substances such as proteins, lecithins, and the products of their breakdown.<sup>8</sup> He called this new material *nuclein*, the substance we now recognize as DNA, for its specific location in the cells' nuclei.

There was nothing easy about Miescher's work, as Williams makes clear with vivid descriptions of the

tasks at hand: comparing, for instance, the extraction of nuclei from the leucocytes to "pitting cherries less than one-thousandth of their usual size."<sup>9</sup> It took Miescher about a year to conduct his analyses and produce sufficient data to submit his write-up to Hoppe-Seyler. The latter was at first suspicious of Miescher's conclusions regarding the uniqueness of nuclein. Hoppe-Seyler did not immediately publish the paper. Instead he waited nearly two years, repeating Miescher's experiments himself and assigning further investigations to another of his students. The results of these experiments convinced him so thoroughly that when Miescher's paper was eventually published, Hoppe-Seyler contributed an introduction. He freely admitted his own prior reservations, before going on to offer fulsome praise: "When I turn to describing the results of these investigations, I can only emphasize that, insofar as I have checked Miescher's information, they must be considered as being fully confirmed."<sup>10</sup>

Williams describes Miescher's initial discovery as a "revelation that fell like a bolt from the blue into a mind that was totally unprepared, because this was the very beginning and, as with the Big Bang, nothing existed before this moment."<sup>11</sup> For an unprepared mind, Miescher proved remarkably prescient about the significance of his discovery. He speculated that the nucleins constituted an entire family of phosphorus-containing substances, whose role in effecting cellular functions would prove equal in importance to that of proteins. He also predicted that, "understanding the relationships between nuclear materials, proteins, and their metabolic products will gradually help to lift the veil, which still so completely obscures the internal processes of cell growth."<sup>12</sup>

**I**N THE THIRD and fourth chapters of *Unravelling the Double Helix*, Williams describes several other origin stories. Landmarks include the first description of the cell's nucleus as a distinct entity by the British botanist Robert Brown in 1833;<sup>13</sup> the German anatomist Walther Flemming's microscopic studies of cell division and his speculations that chromatin—a term he coined—was identical to Miescher's nuclein;<sup>14</sup> and Gregor Mendel's famed experiments on the hybridization of pea plants, which allowed him to lay out the fundamental principles of the inheritance of various traits.<sup>15</sup> Mendel's presence in this roster might seem somewhat incongruous. In contrast to Miescher, Brown, and Flemming, Mendel's name is well-known among scientists and historians, and to a lesser extent among the broader public. But initially, Mendel's contributions were unsung. Save for certain specialist references to his 1866 work, it was not until 1900 that three botanists in Germany and the Netherlands published papers describing his findings and their significance.<sup>16</sup>

Rather than going into the hybridization experiments, Williams begins chapter 4 in the summer of 1878, when C. W. Eichling, a young German traveling salesman of

botanical novelties, spent a day at the monastery in Brunn where Mendel lived and maintained his garden. At an earlier stop in Erfurt, Eichling, whose trip combined “propaganda and [industrial] espionage,” had learned of Mendel quite by chance from a plant merchant, Ernst Benary.<sup>17</sup> Since Brunn was on his itinerary, Eichling followed up on Benary’s lead and paid a visit to this “prominent academic customer’ who had done ... some strange experiments by peas.” Although Eichling and Mendel spent considerable time together as they walked in the abbey’s garden where the hybridization experiments had taken place, peas do not seem to have figured prominently in the conversation. Except for mentioning to Eichling that he had “‘reshaped’ the peas to serve the abbey ‘to better advantage,’” Mendel proved reticent on the details when questioned further, saying only that it was too long a story and changing the subject. Decades later, Eichling admitted that

by failing to draw the master out with a sympathetic question I missed a priceless chance, in that garden sixty-four summers ago, to hear from the lips of the founder of Genetics how he made the discovery which today is recognized as marking an epoch in the study of life.<sup>18</sup>

**N**OT ALL THE stories Williams tells in *Unravelling the Double Helix* are as pleasant as Eichling’s meeting with Mendel. Particularly poignant is the account of the Russian botanist and pioneer of wheat genetics, Nikolai Vavilov. A target of the infamous Lysenko affair, Vavilov suffered a truly harrowing end. “His gravest error,” Williams observes, “was to be a man of principle.”<sup>19</sup> For his pursuit of Mendelian genetics and evidence-based science, Vavilov earned Trofim Lysenko’s enmity and was arrested in 1940 by the Soviet secret police while on a plant-collecting expedition. A 1945 obituary in *Nature* speculated about his fate: “News has recently been received of the death in the Soviet Union of Nikolai Ivanovich Vavilov. The circumstances are not precisely known, but the time was after December 1941 and the place probably Saratov.”<sup>20</sup>

It was not until 1965, when the USSR Academy of Sciences launched an investigation, that the full story of Vavilov’s end came to light.<sup>21</sup> Mark Popovskii, a science journalist who took part in the investigation, faced KGB harassment and had to smuggle material out of the country before finally publishing a biography of Vavilov in 1984.<sup>22</sup> “After 400 sessions of interrogation, some lasting thirteen hours, to make him confess to being an English spy sent to destroy Soviet agriculture,” Vavilov’s fate was decided at a ten-minute trial. He was deemed guilty of being a saboteur and spy, and sentenced to twenty years’ imprisonment.<sup>23</sup> Vavilov took ill and died in a prison hospital in Saratov in January 1943, but not before boosting his fellow prisoners’ morale by delivering whispered lectures on science, agriculture, and trees.<sup>24</sup>

**I**N THE MANY strands of DNA’s secret life, Williams points out that there are only two instances in which scientists “broke entirely new ground.” Otherwise, every actor in this book “had something already established to work on.”<sup>25</sup> The first, of course, was Miescher. The second was the English bacteriologist and public health officer Frederick Griffith. His 1927 discovery initiated a line of investigation culminating in what was perhaps the most important breakthrough for twentieth-century biology. Williams described Griffith’s contribution as a discovery which “lit a slow-burning fuse that, twenty-five years later, detonated the biggest bang of twentieth-century biology.”<sup>26</sup>

In Griffith’s case, as in Miescher’s, the saga began with a similarly unlikely and obscure source of DNA: small, spherical bacteria called pneumococci, which cause pneumonia. Griffith was a specialist in mapping various infections, and, while examining specimens of lobar pneumonia in the early 1920s, he encountered an odd phenomenon. Rather than just one single antigenic type or strain of pneumococcus, he found more than one type in the same patient, appearing several days after the infection took place.<sup>27</sup> Considering it unlikely that the patients could be suffering from multiple infections, he posited that all the later-appearing types were derived from the initial pneumococcal type and devised a series of experiments to test his idea.<sup>28</sup> These experiments led to his groundbreaking discovery of bacterial transformation: the ability of a live avirulent, meaning noninfectious, pneumococcal type to transform into a virulent type when injected into a mouse together with a heat-killed virulent pneumococcus.<sup>29</sup> The specificity of the transformed type always matched the dead strain used in the experiment.

What was the reason one pneumococcal type transformed into another? The difference in virulence between the different types of virulent pneumococci was mapped easily enough to the chemical makeup and antigenic properties of the bacterial capsules. Avirulent strains of bacteria lacked a capsule altogether. But scientists found the capsules could not, by themselves, induce the bacteria to transform; something else was controlling which capsular antigen was produced. This something else, dubbed the “transforming principle,” took more than a dozen years of intermittent work to find. Most of it was conducted in the laboratory of the bacteriologist Oswald Avery at the Rockefeller Institute for Medical Research. These efforts culminated in 1944, with a paper that stopped just shy of announcing that the DNA-containing transforming principle was a gene: “The evidence presented supports the belief that a nucleic acid of the desoxyribose type is the fundamental unit of the transforming principle of *Pneumococcus* Type III.”<sup>30</sup>

Those dozen years from Griffith to Avery, as well as the aftermath, are covered ably by Williams, and readers may consult other accounts for further details and insights into the discovery and its reception.<sup>31</sup> I would like to mention the work of the junior member of Avery’s team, Maclyn

McCarty, whose follow-up contribution merits him a place among the lost heroes of DNA.<sup>32</sup> It was McCarty who designed the experiments that clinched the evidence for DNA's role in pneumococcal transformation. With simple expediency, his experiments compared the transformative abilities of pneumococci when subjected to treatment by enzymes that specifically degraded proteins, DNA, or RNA.<sup>33</sup> Only the enzyme DNase, degrading DNA, could inactivate the transforming principle. But even this evidence that DNA was key failed to attract sufficient notice. Avery's Rockefeller colleague Alfred Mirsky was strongly opposed to his ideas, and the unfortunate consequence was that other biologists were slow to accept DNA as the material of heredity.<sup>34</sup>

WILLIAMS BEGINS *Unravelling the Double Helix* with a glimpse of the end point before narrating the events that led up to it. I end this review by drawing attention to the book's beginning, namely the material that prefaces the narrative. It includes a five-page timeline and an eight-page cast of the major investigators whose work is described in the book.<sup>35</sup> Both are useful features for a tale as long as this one. There is also a preface, in which Williams describes his motives for embarking on this particular project: "ignorance, curiosity and a couple of chance encounters."<sup>36</sup>

In the epigraph, Williams cites Rosalind Franklin. In 1953, upon hearing that Watson and Crick had deduced the double helical structure of DNA, Franklin apparently reacted by paraphrasing Sir Isaac Newton: "We all stand on each other's shoulders."<sup>37</sup> Although Newton's statement has come to represent the importance of giving credit to one's predecessors, or as Williams puts it, "ancestor worship" of a sort,<sup>38</sup> it may not have been as gracious as it seems. Newton alluded to the shoulders of giants in a letter to Robert Hooke, with whom he had a rather fractious relationship over, among other things, priorities in certain discoveries. The statement has been interpreted by some as a sarcastic dig by the upper-class Newton at his lower-class rival's diminutive stature and deformity—Hooke suffered from a humpback.<sup>39</sup> My thought on reading the epigraph was to wonder if Franklin was aware of the intended irony in Newton's statement. If she was in fact aware of the barb, her allusion to it is all the more ironic because of her own openly difficult relationship with Maurice Wilkins, who had, without her permission, shared with Watson her crucial X-ray photograph that led them to propose DNA's helical structure.<sup>40</sup> It is tempting to wonder whether Franklin considered herself the intellectually superior Newton to Wilkins's Hooke.

*Unravelling* is so sprawling, and its cast of characters so wide, that it would be easy to write several reviews focusing on different episodes and figures than I have here. It is not a book to be read in a single sitting, but it needs to be read from beginning to end, at least the first time around, otherwise the reader could become quickly lost. It is to

Williams's credit that he retains control of all the narrative threads without ever getting himself or the readers hopelessly entangled.<sup>41</sup> He found a good story, or rather many stories, and got about the business of telling them.



1. James Watson, *The Double Helix: A Personal Account of the Discovery of the Structure of DNA* (New York: Simon & Schuster; London: Weidenfeld & Nicolson, 1968), 235.
2. Friedrich Miescher, "Über die chemische Zusammensetzung der Eiterzellen (On the Chemical Composition of Pus Cells)," *Medicinish-chemische Untersuchungen* 4 (1871): 441–60; Oswald Avery, Colin MacLeod, and Maclyn McCarty, "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III," *The Journal of Experimental Medicine* 79 (1944): 137–58, doi:10.1084/jem.79.2.137.
3. Gareth Williams, *Unravelling the Double Helix: The Lost Heroes of DNA* (London: Weidenfeld & Nicolson, 2019).
4. Williams, *Unravelling the Double Helix*, xxvi.
5. Williams, *Unravelling the Double Helix*, 9–22.
6. Felix Hoppe, "Über das Verhalten des Blutfarbstoffes im Spectrum des Sonnenlichtes (On the Behavior of Blood Pigments in the Spectrum of Sunlight)," *Virchows Archiv* 23 (1862): 446–49, doi:10.1007/bf01939277; Felix Hoppe-Seyler, "Über die chemischen und optischen Eigenschaften des Blutfarbstoffs (On the Chemical and Optical Properties of Blood Pigments)," *Archiv für pathologische Anatomie und Physiologie und für klinische Medizin* 29 (1864): 597–600, doi:10.1007/bf01926067. Note: Orphaned at a relatively young age, Hoppe-Seyler attached the name Seyler to his family name, Hoppe, in 1864 when he was formally adopted by his brother-in-law Georg Seyler.
7. Ralf Dahm, "The First Discovery of DNA," *American Scientist* 96, no. 4 (2008): 320, doi:10.1511/2008.73.320.
8. Miescher, "Über die chemische Zusammensetzung der Eiterzellen."
9. Williams, *Unravelling the Double Helix*, 12.
10. Felix Hoppe-Seyler, "Über die Chemische Zusammensetzung des Eiters (On the Chemical Composition of Pus)," *Medicinish-Chemische Untersuchungen* 4 (1871): 486 (translation courtesy of Kersten Hall).
11. Williams, *Unravelling the Double Helix*, 8.
12. Miescher, "Über die chemische Zusammensetzung der Eiterzellen," 459–60. A complete scholarly translation of Miescher's original contribution, together with a contextualizing commentary is now available: Kersten Hall and Neeraja Sankaran, "DNA Translated: Friedrich Miescher's Discovery of Nuclein in Its Original Context," *British Journal for the History of Science* 54 (2021): 99–107, doi:10.1017/S000708742000062X.

13. Williams, *Unravelling the Double Helix*, 24–26; Robert Brown, “On the Organs and Mode of Fecundation in Orchideæ and Asclepiadeæ, with Additional Observations,” *Transactions of the Linnean Society of London* 16 (1833): 685–745, doi:10.1017/cbo9781107775473.017.
14. Williams, *Unravelling the Double Helix*, 28–35; Walther Flemming, *Zellsubstanz, Kern und Zellteilung* (Cell Substance, Nucleus, and Cell Division) (Leipzig: Vogel, 1882).
15. Williams, *Unravelling the Double Helix*, 36–51; Gregor Mendel, “Versuche über Pflanzenhybriden (Experiments on Plant Hybridization),” *Verhandlungen des naturforschenden Vereines in Brünn*, Bd. IV für das Jahr 1865 (1866).
16. Robert Olby and Peter Gautrey, “Eleven References to Mendel before 1900,” *Annals of Science* 24 (1968): 7–20, doi:10.1080/00033796800200021. The three publications marking the rediscovery of Mendel are Carl Correns, “G. Mendel’s Regel über das Verhalten der Nachkommenschaft der Rassenbastarde (Mendel’s Law on the Behavior of Progeny of Variable Hybrids),” *Berichte der Deutschen Botanischen Gesellschaft* 18 (1900): 156–68, doi:10.1007/978-3-642-52587-2\_2; Hugo de Vries, “Sur la loi de disjonction des hybrides (On the Law of Hybrid Disjunction),” *Comptes Rendus de l’Academie des Sciences* 130 (1900): 845–47; and Erich von Tschermak, “Über künstliche kreuzung bei *Pisum sativum* (On the Artificial Crossing of *Pisum sativum*),” *Berichte der Deutschen Botanischen Gesellschaft* 18 (1900): 232–39.
17. Williams, *Unravelling the Double Helix*, 36–38.
18. C. W. Eichling, “I Talked with Mendel,” *Journal of Heredity* 33 (1942): 246, doi:10.1093/oxfordjournals.jhered.a105181.
19. Williams, *Unravelling the Double Helix*, 171.
20. Sydney Harland and Cyril Darlington, “Prof. N. I. Vavilov, For.Mem.R.S.,” *Nature* 156 (1945): 621, doi:10.1038/156621a0.
21. Allan Wilson, “Review: *The Vavilov Affair* by Mark Popovskii,” *New Zealand Slavonic Journal* (1987): 246–49.
22. Wilson, “Review: *The Vavilov Affair*,” 246–47; Mark Popovskii, *The Vavilov Affair* (Hamden: Archon Books, 1984).
23. Williams, *Unravelling the Double Helix*, 231.
24. Williams, *Unravelling the Double Helix*, 232; John Fincham, “The Fall of Nikolai Vavilov,” *Nature* 316 (1985): 769–70, doi:10.1038/316769a0.
25. Williams, *Unravelling the Double Helix*, 372.
26. Williams, *Unravelling the Double Helix*, 138.
27. Frederick Griffith, “Types of Pneumococci Obtained from Cases of Lobar Pneumonia,” *Reports on Public Health and Medical Subjects* 13 (1922): 1–13.
28. There is no need to rehash the details of Griffith’s experiments to test his idea. Readers may find them not only in Williams’s beautifully clear account and other historical treatments, but also in Griffith’s own publications. Williams, *Unravelling the Double Helix*, 138–42; Pierre-Olivier Méthot, “Bacterial Transformation and the Origins of Epidemics in the Interwar Period: The Epidemiological Significance of Fred Griffith’s ‘Transforming Experiment’,” *Journal of the History of Biology* 49 (2016): 311–58, doi:10.1007/s10739-015-9415-6; Frederick Griffith, “The Significance of Pneumococcal Types,” *Journal of Hygiene* 27, no. 2 (1928): 113–59, doi:10.1017/s0022172400031879.
29. Griffith, “Significance of Pneumococcal Types.”
30. Avery, MacLeod, and McCarty, “Studies on the Chemical Nature,” 156. Note: The phrase “nucleic acid of the desoxyribose type” stands for DNA. The term “desoxyribose” has since been completely replaced by the more familiar “deoxyribose” which denotes the D in DNA (deoxyribonucleic acid).
31. Williams, *Unravelling the Double Helix*, 143–53; Horace Freeland Judson, *The Eighth Day of Creation: Makers of the Revolution in Biology* (New York: Simon and Schuster, 1979); Lewis Thomas, “Oswald Avery and the Cascade of Surprises,” *Technology in Society* 6, no. 1 (1984): 37–40, doi:10.1016/0160-791X(84)90016-2; Nicholas Russell, “Oswald Avery and the Origin of Molecular Biology,” *British Journal for the History of Science* 21, no. 4 (1988): 393–400, doi:10.1017/S0007087400025310; Olga Amsterdamska, “From Pneumonia to DNA: The Research Career of Oswald T. Avery,” *Historical Studies in the Physical and Biological Sciences* 24, no. 1 (1993): 1–40, doi:10.2307/27757711; Robert Olby, *The Path to the Double Helix: The Discovery of DNA* (Mineola: Courier Dover Publications, 1994); F. Kohl, “Konzepte und Experimente zum Nachweis der DNA als Vererbungssubstanz: Der historische Weg von Mendel und Miescher zum Schlüsselversuch der Avery-Gruppe und dessen Bestätigung am Phagenmodell (The Concepts and Experiments to Demonstrate DNA as the Substance of Heredity: The Historical Path of Mendel and Miescher to the Key Experiment of the Avery Group and Its Confirmation in the Phage Model),” *Deutsche medizinische Wochenschrift* (1946) 121, no. 34/35 (1996): 1,066–69, doi:10.1055/s-0029-1233824; Ute Deichmann, “Early Responses to Avery et al.’s Paper on DNA as Hereditary Material,” *Historical Studies in the Physical and Biological Sciences* 34, no. 2 (2004): 207–32.
32. McCarty’s achievements deserve wider recognition for their scientific merit alone, but it should be added that he also qualifies for the label that Williams bestowed upon Walther Flemming, as “one of the few genuinely nice people in the history of DNA.” Williams, *Unravelling the Double Helix*, 29. I can attest to this same quality in Mac McCarty, whom I met with and interviewed a few times in 1996 and 1997, more than half a century after he had performed the transformation experiments.
33. Maclyn McCarty and Oswald Avery, “Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types,” *Journal of Experimental Medicine* 83 (1946): 89–96, doi:10.1084/jem.83.2.89; Maclyn McCarty, “Chemical Nature and Biological Specificity of the Substance Inducing Transformation of Pneumococcal Types,” *Bacteriological Reviews* 10 (1946): 63–71, doi:10.1128/br.10.1-2.63-71.1946.
34. Williams, *Unravelling the Double Helix*, 217.
35. Williams, *Unravelling the Double Helix*, xiii–xvii, xviii–xxv.
36. Williams, *Unravelling the Double Helix*, xxvi–xxx.
37. Williams, *Unravelling the Double Helix*, ix.

38. Williams, *Unravelling the Double Helix*, xxviii.
39. Robert Crease, *The Great Equations: Breakthroughs in Science from Pythagoras to Heisenberg* (New York: W. W. Norton & Company, 2010, Kindle Edition), 82.
40. Williams, *Unravelling the Double Helix*, 8, 327–29.
41. Although writing is not Williams’s primary profession, *Unravelling* is by no means his first book; he has written others about aspects of the history of medicine and disease, and is the author of a delightful book about the Loch Ness monster. Gareth Williams, *A Monstrous Commotion: The Mysteries of Loch Ness* (London: Orion Books, 2015).

DOI: 10.37282/991819.21.20

BOOK REVIEWS

# Secrets

*Jeremy Bernstein*



*Published on June 7, 2021*

***Restricted Data: The History of Nuclear Secrecy in the United States***

by Alex Wellerstein

University of Chicago Press, 528 pp., \$35.00.

**N**UCLEAR WEAPONS have been shrouded in secrecy from the very beginning. After plutonium was discovered at the University of California in December 1940, researchers led by Glenn Seaborg submitted a pair of letters to the *Physical Review*. The details of their discovery were withheld from publication until after the war.<sup>1</sup> Once the project to make a nuclear weapon got underway, secrecy became a very serious matter indeed. The story of these efforts and how they evolved after the war is the subject of Alex Wellerstein's *Restricted Data: The History of Nuclear Secrecy in the United States*. It is an extremely detailed study. I can tell he has been working on this book for many years because in his acknowledgments he refers to sources who have long been dead. In his introduction, Wellerstein points out that he has never worked on a classified project or held a security clearance. "This no doubt leaves many additional gaps in the story," he observes, "but it also allows me to share what I have found with impunity."<sup>2</sup>

**B**EFORE I TURN to a discussion of the book, let me explain why this subject is of special interest to me. In the winter of 1954, it had become clear that by the following year I would obtain my PhD from Harvard. I had passed the qualifying examinations, taken all the required courses, and was well underway with my thesis. I had not given much thought to what I might do next. As it turned out, an assistant professor I knew had heard about a new laboratory opening at the University of California at Berkeley and passed on my details to one of its founders, Edward Teller. I arranged to meet with Teller in Washington at the spring meeting of the American Physical Society. It was only after our meeting that I came to understand why he had started the Lawrence Livermore National Laboratory. Around 1952, Teller began to feel that the researchers at Los Alamos were not working hard enough on nuclear weapons to confront the Soviet Union. He decided to found a rival weapons laboratory. Even if Teller had offered me a job, I doubt that I would have accepted.<sup>3</sup>

After obtaining my degree, I was offered a job that would keep me in Cambridge for at least another year. One year became two and at the end of my second year I was accepted at the Institute for Advanced Study in Princeton. It was around this time that the chairman of the physics department at Harvard, Kenneth Bainbridge, came to me with an offer. Bainbridge had been an important figure at Los Alamos during the war. Robert Oppenheimer had put him in charge of the site in New Mexico where the Trinity test had taken place.<sup>4</sup> Bainbridge told me that the laboratory was offering summer jobs to young PhDs and asked

if I was interested. I was very interested. Los Alamos had an almost mystical significance for me due to its history and remote location. During our conversation, Bainbridge noted that I would have to pass a security clearance before I could work there.

Wellerstein discusses the history of these clearances in his book and describes how they evolved from informal agreements to the highly structured and rigorous arrangements that I confronted. By 1957, an access authorization known as a Q clearance was required to work in any of the technical divisions at the laboratory.<sup>5</sup> As the highest level of security clearance, it was a prerequisite for anyone who needed access to classified information connected with nuclear weapons research. The FBI handled all the interviews for the clearance process and contacted many of my friends and neighbors as part of its background checks.<sup>6</sup> I was relieved when I received my clearance.

**W**HEN I ARRIVED at Los Alamos in the summer of 1957, I shared an office with another Harvard postdoc named Kenneth Johnson. There may have been a security notice to the effect that we would only receive classified information on a need-to-know basis. After a couple of days, we realized that no one was giving us any information about anything—including what we should be doing. I decided that we might as well work on our own physics problems and proposed one to Johnson that I had begun to think about. He was a whiz when it came to calculations and soon had the whole thing worked out. I went to the head of the theoretical division, J. Carson Mark, and asked what we should do with our result. Mark suggested we write up the work and circulate it as a Los Alamos preprint. A paper like this would show that we were capable of working on non-weapons-related physics at the laboratory.

Not long after I got there, some distinguished theorists began arriving at Los Alamos to work on controlled fusion. This research program was aimed at fusing light elements, such as the isotopes of hydrogen, under controlled conditions to form helium and yield an energetic neutron. It is this kind of reaction that powers the hydrogen bomb. At the time, controlled fusion research was classified. It was only after I obtained my clearance that I was able to listen to a seminar. A little more than a year later, the subject was declassified and the research presented at the Atoms for Peace conference in Geneva, alongside similar work from Great Britain and the Soviet Union.<sup>7</sup> As it turned out, the Russians had more to teach us than we them. This area of research is discussed in detail by Wellerstein, and his account includes many details I did not know.

Among the theorists who descended on Los Alamos was Francis Low, one of the most brilliant theoretical physicists of his generation. Low also liked to play tennis. We began to play together on a regular basis and even became part of the Los Alamos team. Sometime in late August, Low was forced to skip a match because he was going to Nevada

to observe some bomb tests. I knew that Los Alamos was heavily involved in a series of nuclear tests that summer, which I later discovered were known as Operation Plumb-bob. No fewer than twenty-nine tests were scheduled to take place between May 29 and October 7. Among them were projects from both Los Alamos and Livermore; the Los Alamos tests were named after scientists and the Livermore tests, after mountains. It had never occurred to me that someone outside the weapons program could watch these tests. Low told me he had been invited by Mark. When I asked Mark if I could accompany Low, he agreed on the condition that I pay for the airfare between Albuquerque and Las Vegas.

On the afternoon of August 30, we boarded a plane for Las Vegas. The Nevada Test Site is located about sixty-five miles north of the city. We were met upon arrival at the airport by a government car that took us at once to a casino.<sup>8</sup> In hindsight, it seems crazy that above-ground nuclear tests were taking place only sixty-five miles away from a major city. I was told that the mushroom clouds were clearly visible from the casinos. Later that evening we were informed that the test scheduled for the following morning at sunrise was going ahead. After a few hours of sleep we made our way to a concrete bunker to observe the explosion. In the distance I could see the tower that held the device. We were given dark lenses to hold over our glasses and instructed to look away and count to ten before turning around. When I turned around, the clouds were glowing and the nearby hill was lit up with flame. Next came the shockwave and then finally the sound of the explosion, which almost seemed like an afterthought. I learned later that the device being tested, Smoky, had a yield of forty-four kilotons, around three times the yield of the bomb dropped on Hiroshima.

After a short break, Mark took us on his rounds. The first stop was the tower that had been erected for a Los Alamos device, Galileo. From memory, the tower appeared to be several hundred feet tall. The platform that held the device was accessed by taking a flimsy-looking open elevator most of the way up, and then climbing the rest of the way on an equally flimsy metal ladder. The device itself looked more like a science experiment than a bomb. There were wires protruding from it in all directions and a couple of men working on it attaching yet more wires. I noted a familiar *cluc-cluc* sound in the background. It was the sound that a pump makes when it creates a vacuum. I wondered why I could hear it on the platform, but since I had no need to know, I did not ask.

After Mark was finished at the tower, we got back in his car and drove to a blockhouse at the edge of the site. When he opened the door I recoiled. On a shelf inside was an array of metal pits. Despite my limited knowledge of nuclear weapons, I was still able to recognize the plutonium cores. Mark took one off the shelf and handed it to me. It was about the size and weight of a bowling ball and warm to the touch. If I had known that the density

of plutonium was greater than lead, I would have realized that these were hollow pits. Even so, it would never have occurred to me to ask why.

I had no need to know.

The next morning, we watched the Galileo device explode with a yield of only eleven kilotons. After the initial atomic bombs had been assembled, the designers realized that they could make a much more efficient device if they combined fission and fusion. This is not what is typically referred to today as a hydrogen bomb, but something in between. Before detonation, a partial vacuum is created in a hollow pit. A mixture of deuterium and tritium gas is then injected. When the fission device detonates, these nuclei fuse to form helium, emitting a neutron. This neutron is sufficiently energetic that it can fission uranium-238, and this is what boosts the bomb. The Galileo test was a new version of this device. After the test, I began the long drive back east with Low. We never spoke about the experience again, either during that journey, or in the years that followed. Low died in 2007, so I will never have the chance to ask him what he knew when he was watching these tests.

I had no need to know.

As part of the security measures in wartime Los Alamos, Wellerstein describes how a “system of colored badges distinguished the different categories of knowledge one might be entitled to.”<sup>9</sup> If the badge was white, the bearer was permitted access to restricted data on a need-to-know basis.<sup>10</sup> Oppenheimer introduced a weekly colloquium for all white badge holders. There was also an afternoon tea. The latter tradition was still being observed during my summer at Los Alamos. I remember a British physicist who had been there during the war looking at the people in attendance and remarking that “the days of the great Los Alamos teas are over.” There had once been an abundance of present and future Nobel Prize winners at these events. Those days were long gone.

**A**S WELLERSTEIN OBSERVES, secrecy has been part of the nuclear ethos almost since the discovery of the neutron by James Chadwick in 1932. A year later, the Hungarian-born polymath and physicist Leo Szilard sensed the implications of the changing political climate in Germany and moved to England. Szilard was out for a walk in London when he was struck by an idea. Suppose that there were a process in which, say, two energetic neutrons were produced. Fission is just such a process, but it would not be discovered until the end of 1938. Once emitted, the two neutrons could produce more neutrons as part of a chain reaction. Szilard realized that the reaction could be explosive and immediately filed a secret patent application, which was granted in 1936. After moving to the United States several years later, Szilard drafted a confidential letter to President Franklin D. Roosevelt, which was signed by Albert Einstein, warning that the Germans might attempt to make an atomic bomb.<sup>11</sup>

As Wellerstein notes, the whole question of nuclear secrecy took on a different character after the creation of the United States Atomic Energy Commission (AEC) in 1946. The Smyth Report is a good example of the somewhat random character of information releases prior to the AEC. Written by the physicist Henry DeWolf Smyth and dated July 1, 1945, the report preceded not only the attacks on Hiroshima and Nagasaki, but also the first Trinity test. It was compiled at the request of General Leslie Groves, head of the Manhattan Project. He realized that a proper accounting would be needed to show how the vast sums of money allocated to the project had been spent. Smyth was chosen to write it. In the end, the report was not released until several days after the bombing on Nagasaki. It quickly created controversy and there were even claims that it revealed the secret of the atomic bomb. What Smyth really revealed was just how hard it was to produce the materials to make one. The report makes no mention of implosion, for example, but there is a great deal of discussion about the vast enterprise needed to produce plutonium.

Until I read Wellerstein's book, I had forgotten that Hans Bethe wrote an article in the spring of 1950 for *Scientific American* on the hydrogen bomb.<sup>12</sup> I knew Bethe well: no one was more scrupulous about restricted data. His article was no exception, but the AEC decided that a figure like Bethe writing on the topic was a revelation in itself. They insisted that several thousand copies of the magazine be destroyed, along with the printing plates.

WELLERSTEIN'S ACCOUNT OF A technological transfer that took place around 1956 was of particular interest to me. It has to do with what is usually known as the Zippe centrifuge, although its designer, Gernot Zippe, once told me that he preferred to call it the Russian centrifuge. Zippe was born in 1917 in what was then Austria-Hungary. After obtaining a degree in physics from the University of Vienna in 1938, Zippe joined the Luftwaffe as a flight instructor and researcher. In the early 1940s, he began studying mechanical engineering at the University of Munich and joined a team dedicated to isotope separation, eventually becoming the group's leader. Zippe was captured by the Russians at the end of the war and interned at a prisoner-of-war camp for German scientists in the outskirts of Moscow. In the summer of 1946, he was sent to a research facility near Sukhumi on the Black Sea where he was tasked with designing gas centrifuges that could be used to separate uranium isotopes in uranium hexafluoride gas.<sup>13</sup>

Before arriving in Sukhumi, Zippe had never seen a centrifuge and knew nothing about them. All he had at his disposal was an antiquated Russian model and some equally outdated literature. During the war, the American physicist Jesse Beams successfully separated uranium isotopes using a centrifuge he had developed while working on the Manhattan Project. Despite considerable invest-

ment, the project ultimately failed to deliver sufficient quantities of highly enriched uranium. Starting from scratch, Zippe came up with a design that was orders of magnitude more efficient.

In 1956, Zippe was released by the Russians, and the following year in Holland he attended a meeting on centrifuge design. Realizing that his own design was far better than any of the models on display, Zippe joined the centrifuge industry. As Wellerstein notes, the classification of centrifuge designs has always been rather murky. All the gas centrifuges being used for uranium enrichment in Iran, for example, are modified versions of Zippe's original design.

Wellerstein's chapter on the Zippe centrifuge is followed by another dealing with someone else I knew, the theoretical physicist Keith Brueckner, who died in 2014. Brueckner was the most ambitious person I have ever met and his struggles with the AEC make for entertaining reading. Brueckner had some novel ideas about confining plasma using lasers that had potential applications in the efforts to harness controlled fusion for power generation. At the time, Brueckner was working in collaboration with KMS Industries, a private sector organization unconnected to the AEC.

In 1969, Brueckner applied for three patents and was informed that his work was considered classified. His ideas had, in fact, already been found in connection with the hydrogen bomb program. "All of that previous work was classified," Wellerstein notes, "and the secrecy surrounding the topic meant that someone like Brueckner could argue that his own invention was not only independent, but arguably had been first."<sup>14</sup> Undeterred, Brueckner filed a dozen more patent applications. This led to a very entertaining confrontation, one that is well described by Wellerstein. A compromise was eventually reached, but whatever these ideas were, they do not appear to have led to any breakthroughs in controlled fusion research.

The following chapter examines the spread of nuclear secrets and how difficult it might be for a terrorist group, or even a student, to assemble their own device.<sup>15</sup> As Wellerstein notes, there have been a number of well-publicized incidents involving students. The most celebrated case is that of John Aristotle Phillips, who came up with a design for a bomb while he was an undergraduate at Princeton in the mid-1970s. Phillips was inspired by John McPhee's book *The Curve of Binding Energy*, a profile of the bomb designer Ted Taylor.<sup>16</sup> Phillips attempted to fill in some of the deliberate omissions in Taylor's account with the help of Freeman Dyson, who really should have known better. When Bethe saw the design, he dismissed it as worthless. Dyson subsequently claimed that he had only gotten involved to show how easy it was to obtain declassified material in Washington.<sup>17</sup>

All of this, and indeed the rest of Wellerstein's book, confirms what I have long believed to be true: the infor-

mation needed to make a nuclear weapon has been in circulation for a long time. In his book, Wellerstein discusses the Iranian and Israeli weapons programs. Israel certainly has this information and my guess is that Iran does too. The fissile material needed for such a device is much harder to obtain. In my view, controlling the spread of these weapons involves controlling the production of these materials. “From a technical standpoint,” Wellerstein writes, “nuclear weapons should have been very easy to control.”

As Oppenheimer understood in 1945, the material pipelines to acquiring nuclear weapons are relatively large, and controlling the uranium enrichment facilities means controlling the spread of the bomb, even in the face of possibly incomplete or non-existent secrecy. That nuclear control has been elusive should give us pause. The problem of controlling nuclear weapons has never been a purely technical one—it is rather, a political problem. And technical solutions to political problems are rarely adequate.<sup>18</sup>



1. John Kennedy et al., “Properties of  $^{94}(239)$ ,” *Physical Review* 70, no. 555 (1946), doi:10.1103/PhysRev.70.555. The following explanation was included when the paper was finally published in 1946:

This letter was received for publication on the date indicated but was voluntarily withheld from publication until the end of the war. The original text has been somewhat changed, by omissions, in order to conform to present declassification standards.

2. Alex Wellerstein, *Restricted Data: The History of Nuclear Secrecy in the United States* (Chicago: University of Chicago Press, 2021), 9.
3. Our meeting was taken up, for the most part, by Teller rehearsing a talk he had prepared for the following afternoon. I found his talk largely incomprehensible—it was something to do with the meson theory of nuclear forces—and not particularly interesting. Just before I left, Teller made the odd remark that it was nice to be doing physics again. It took me a while before I put two and two together. He had been testifying in the hearing which resulted in Robert Oppenheimer losing his security clearance. Teller’s carefully crafted negative testimony had been decisive. When the physics community found out about what he had done, Teller became a pariah. In his book, Wellerstein provides an excellent account of Oppenheimer’s trial, which also turned on his lack of enthusiasm for making a hydrogen bomb as well as his personal morality. Oppenheimer did not defend himself very well. He once told me that during the trial he felt like they were discussing someone else.

4. After the successful test it was Bainbridge who made the oft quoted remark, “Now we’re all sons of bitches.”
5. The “Q” was short for “Personal Security Questionnaire.”
6. While these checks were taking place, I was worried about my great-aunt May. If not an actual member of the Communist party, she was a devoted fellow traveler. Either they never came across her, or they found her too batty to be of interest. Years later when I was writing for the *New Yorker*, I thought that I would use the Freedom of Information Act to get my clearance document and then write something I was going to call *Friends and Neighbors*. When I received the file, it was almost entirely redacted and useless for my purposes.

7. The main attraction [was] the promised revelation of secret fusion research by the United States, Great Britain and the Soviet Union, i.e. the talks of the leading fusion scientists Hannes Alfvén, Lev Artsimovich, Ludwig Biermann, Peter Thonemann and Edward Teller on “the possibility of controlled fusion.”

Sabina Griffith, “Two Weeks in September, 1958: Atoms for Peace Conference in Geneva,” *ITER Newslines*, Sept. 1, 2008.

8. It turned out that some mathematically inclined soldiers had analyzed casino blackjack and found a way to beat the dealer. It was written up and had been analyzed by the people at Los Alamos by playing hundreds of thousands of hands on a computer. A little card had been issued which you could bring to the casino and try your luck.
9. Wellerstein, *Restricted Data*, 57.
- 10.

Those with yellow badges, such as security details, could enter the technical areas of the lab but were not to be told any classified information at all. Blue badges were for clerks and warehouse employees who would need to know some classified information, like schedules and rosters and names, but would not be allowed access to technical information. Red badges were for technicians and secretaries whose access to information could be quite deep within the scope of their own job but could never exceed it.

Wellerstein, *Restricted Data*, 57.

11. An interesting story note about Einstein is that he had not heard of fission when Szilard gave him the letter in 1939. It had been discovered a year earlier and was a sensation among physicists, including Einstein’s colleagues John Wheeler and Niels Bohr at Princeton, who had created the first general theory of fission. In his later life, Einstein had little interest in ordinary physics.
12. Alfred McCoy, “How an Article about the H-Bomb Landed Scientific American in the Middle of the Red Scare,” *Scientific American*, September 1, 2020, doi:10.1038/scientificamerican0920-73.
13. For further details, see my essay “A Nuclear Supermarket” in *Physicists on Wall Street and Other Essays on Science and Society* (New York: Springer, 2008), 52. For a more detailed

account of Zippe's transfer to Sukhumi and the work in which he was involved, see Pavel Oleynikov, "German Scientists in the Soviet Atomic Project," *The Nonproliferation Review* 7, no. 2 (2000): 1–30, doi:10.1080/10736700008436807. The weapons research program at Agudzera near Sukhumi led to the creation of the Sukhumi Institute of Physics and Technology. The institute split in two after the conflict between Georgia and Abkhazia in the early 1990s. Some of its people and programs were relocated to the Georgian capital, Tbilisi, to form the Ilia Vekua Sukhumi Institute of Physics and Technology.

14. Wellerstein, *Restricted Data*, 327.

15. I once advised a high school student who was trying to make a small fusion reactor in his garage. He was certainly not trying to make a bomb, but he did get some neutrons out of

his device, showing that there had indeed been fusion.

16. Taylor was my boss when I consulted on the Orion spaceship project. For more about this project, a brief biographical sketch of Taylor, and some remarks from Freeman Dyson, see my essay "Reflections on Project Orion," *Inference: International Review of Science* 5, no. 2 (2020), doi:10.37282/991819.20.4.

17. When my friend, the Pakistani physicist Riazuddin (he had only one name), was put in charge of the attempt to make a nuclear weapon, he took a sabbatical at the University of Maryland so that he could be close to Washington and take advantage of the declassified material.

18. Wellerstein, *Restricted Data*, 410.

DOI: 10.37282/991819.21.15

EXPERIMENT REVIEWS

# Reconstructing Ancestral Proteins

*Chase Nelson*



*Published on August 11, 2021*

Lucas Wheeler and Michael Harms, “Were Ancestral Proteins Less Specific?” *Molecular Biology and Evolution* 38, no. 6 (2021): 2,227–39, doi:10.1093/molbev/msab019.

**M**EMBERS OF A FAMILY resemble one another. Such similarities are most often noticed in contemporaries, as when siblings have the same smile. But they also provide insight about individuals long departed. Absent a historical record, the best bet for inferring the traits possessed by a forebear is to examine all their descendants in the present. It matters neither how long ago the ancestor lived nor how copious their progeny; if 90% of the offspring are tall, it seems reasonable to suppose the primogenitor was too. As the aphorism suggests, the present is the key to the past.

What is true of people is also true of proteins, and for the same reasons. Proteins are molecular machines built from strings of amino acids, typically a few hundred residues in length, chained together with covalent bonds like beads on a necklace. Most organisms employ an alphabet of twenty amino acids in their proteins, with an overall function defined by their precise order—leucine at position 5, lysine at position 26. If two proteins contain identical amino acids at a sufficient number of positions, common ancestry can be inferred and they can be grouped into a family. If every currently known family member contains, say, lysine at position 5, chances are high that the same was true of the ancestral protein.

But proteins aren’t always what they used to be. This is true in the sense that one amino acid occasionally gets swapped for another over the course of evolution. But it is also true in terms of general properties. Each amino acid has its own size, hydrophobicity, and charge. As a result, neighboring residues along the chain undergo biophysical interactions, giving rise to a series of twists and turns. The chain progressively folds into a particular three-dimensional shape and thereby assumes a particular function. Swapping an amino acid can disturb this shape and thereby alter the function, most commonly by changing the suite of molecules with which the protein has the ability to interact.

Inferring the exact amino acids in ancient proteins has long been a topic of interest among theoreticians. Writing in 1963, Linus Pauling and Emile Zuckerkandl introduced the concept of chemical paleogenetics. They reasoned that sequencing enough proteins in the present could allow sufficient numbers to be grouped into families, thus allowing their extinct protein ancestors to be inferred and synthesized in the lab.<sup>1</sup> These proteins could then be used to test a wide range of evolutionary hypotheses about ancient organisms, even those that left no trace in the fossil record. The field blossomed in the 1980s with leaders such as Steven Benner, followed later by Gina Cannarozzi, Betül Kaçar, Joseph Thornton, and many others.

An influential example of ancestral protein reconstruction, conducted by Eric Gaucher and colleagues, involved bacterial elongation factors of the Tu family.<sup>2</sup> Focusing on

bacteria that are currently mesophilic—organisms that grow optimally at temperatures between 20 and 40°C—these researchers found the ancestral protein to function optimally between 55 and 65°C, suggesting that the ancestor was instead thermophilic. This example highlights the necessity of protein reconstruction: it may not be enough to examine the average properties of a protein’s modern descendants. The ancestral molecules themselves must be constructed.

Other applications of ancestral protein reconstruction have involved testing general hypotheses about genome evolution and evaluating specific gene functions to gain a better understanding of the connection between genotype and phenotype.<sup>3</sup> Proteins can even be compared directly to their ancestors over an evolutionary tree, circumventing the historical, and therefore statistical, dependencies that have long plagued comparisons among modern forms alone.<sup>4</sup>

**G**IVEN THAT PROTEINS are associated with repertoires of target molecules, a particular protein can be categorized as relatively specific, having few targets, or general, having many targets. From a theoretical standpoint, the existence of an overall historical trend in the direction of, say, general-to-specific could allow a principle of protein evolution to be established. Indeed, Roy Jensen proposed just such a principle in 1976, reasoning that early cells must have possessed a small collection of genes encoding proteins with relatively general activities.<sup>5</sup>

Protein specificity may seem a mere theoretical curiosity, but it is perhaps the most far-reaching line of research to spring from ancestral protein reconstruction and for one lucrative reason: protein engineering. A protein binding its target is not the end but the beginning of its functional story. The binding of a single molecule might set off a change in the protein’s conformation that ultimately stretches the target until it rips in two. Or, the binding of multiple molecules might stimulate the protein to smash them together until a new product is forged. In either case, the binding of targets by proteins speeds up perhaps every chemical reaction carried out by the cell.

If ancestral proteins tended to be more general than their modern forms, they would have been capable of binding a wider range of targets and driving a greater diversity of potential reactions. They might also have exhibited other valuable features, including increased stability at high temperatures and the ability to fold without the aid of modern so-called chaperone proteins.<sup>6</sup> In such cases, efforts to engineer new protein applications would be much better spent investigating variations on ancestral proteins instead of their more limited descendants. There have indeed been glimmers that ancestral proteins were more general.<sup>7</sup>

**I**N A 2018 STUDY, a team including evolutionary biochemists Lucas Wheeler and Michael Harms addressed the question of protein generality by

focusing on the S100 protein family, so named due to their solubility in 100% saturated ammonium sulfate.<sup>8</sup> This family comprises just over twenty members in the human genome, most of them encoded by a cluster of genes located on the long arm of chromosome 1. For the vast majority of S100 proteins, the binding of Ca<sup>2+</sup> ions transiently exposes a hydrophobic binding pocket which, if occupied by a well-fitting target molecule, forms a stable structure that holds both the Ca<sup>2+</sup> and the target in place.<sup>9</sup> Members of the S100 protein family function as calcium sensors, maintaining homeostasis by responding to intracellular Ca<sup>2+</sup> fluctuations in real time. Many play key roles in cell replication and immune detection.<sup>10</sup>

Wheeler and colleagues focused on just two S100 family members: S100A5 and S100A6. This pair is thought to have resulted from the duplication of a single ancestral gene, dubbed *ancA5/A6*, around 318 million years ago in the common ancestor of mammals, birds, and reptiles.<sup>11</sup> The initial result was two identical and therefore redundant copies, but such tandem gene duplication events eventually allow each descendant gene to specialize in just a subset of the ancestor's tasks.<sup>12</sup> Both S100A5 and S100A6 are known to be overexpressed in certain cancers and to bind some of the same targets, but have slightly different, if poorly defined, functions.<sup>13</sup>

S100 is a particularly well-chosen family for investigating the evolution of protein specificity: this family binds a set of targets so diverse that no one pattern or motif can be used to summarize them. It can at least be said that the targets are typically short protein fragments, or peptides, twelve or more amino acids in length and occurring within larger proteins. Binding probably occurs through a combination of shape complementarity and hydrophobic interactions.<sup>14</sup> Known peptide targets include regions of the proteins sodium/calcium exchanger 1 (NCX1), Siah-interacting protein (SIP), and two commercially available peptides.

Focusing on just these four targets, the team first determined which were bound by each human S100 copy: S100A5 bound the two commercial peptides and NCX1, but not SIP, while S100A6 bound only one commercial peptide and SIP, but not NCX1 or the other commercial peptide. The team then used maximum likelihood methods to reconstruct the ancestral *ancA5/A6* protein and assess its own binding: remarkably, all four targets were bound by the ancestor.<sup>15</sup> It seemed that in the time since their duplication from *ancA5/A6*, S100A5 and S100A6 had partitioned binding partners as one might split belongings after a divorce. What was one had become two, and neither quite amounted to what it once was. These results hinted at a trend toward increasingly specific proteins in evolution.

**D**ESPITE THESE RESULTS, a sample size of four targets was a limited basis on which to draw any conclusions. As the authors themselves cautioned,

[I]t could be that the proteins both acquired more peptides that we did not sample ... while becoming more specific for the chosen set of targets. ... Particularly given the large number of targets for these proteins, distinguishing these possibilities will require an unbiased, high-throughput approach to measuring specificity.<sup>16</sup>

In a new study, Wheeler and Harms increase the sample size of targets from four to approximately 100,000.<sup>17</sup> To do this, they first note that S100 proteins might conceivably bind peptide targets not produced by a present-day organism. Thus, rather than limiting to known interaction partners, they instead examine a random sample drawn from all 20<sup>12</sup>—that is, 4.096 × 10<sup>15</sup>—possible peptides that are twelve amino acids in length. Their final sample is a commercial library of some 10<sup>9</sup> unique peptides.

Not each of the 10<sup>9</sup> potential targets is actually bound by S100A5 or S100A6. To home in on those that are, the authors turn to quantitative phage display, in which gene fragments encoding random peptide targets are inserted into the surface protein gene of bacteriophages. Each phage expresses one target. The whole pool of phages is then mixed with S100 to allow binding, after which the S100 is isolated. Those phages expressing a target bound by an S100 will remain stuck, while the others are washed away. Finally, the bound phages undergo sequencing of their genes en masse to identify which 12-amino acid peptides were bound. Another experiment is run in parallel with a competitor peptide already known to occupy S100's binding pocket, allowing the results to be normalized in the event that some of the targets bind indiscriminately to other parts of the S100s.

The results are something of a surprise: instead of increasing specificity with fewer targets, the authors observe a pattern of shifting specificity. S100A5 and S100A6 still bind just a subset of the ancestor's targets; but, when assaying this large sample, it becomes clear that the loss of old targets is offset by a gain of new targets. The specificities of the S100s have not, in fact, changed in comparison to their ancestor. Using alternative, and less likely but still plausible, ancestors yields generally consistent results.<sup>18</sup>

Limiting studies to a handful of known interaction partners had given the wrong picture. To more accurately characterize a protein's intrinsic specificity, unbiased sampling of targets—including those not known in modern organisms—is necessary.

**T**HE STUDY OF molecular evolution seeks to reveal how the building blocks of organisms have changed over time, both in pattern and in process. Wheeler and Harms have charted new territory in describing the pattern of evolution at the protein level, albeit for a single pair of closely related proteins. Can any implications be drawn from their work about what drives protein evolution?

One of the most impressive answers to this question comes from a 2008 study by Shozo Yokoyama and colleagues.<sup>19</sup> These researchers resurrected the protein ancestors and intermediates of thirty-eight modern-day rhodopsins, visual pigments used by vertebrates to see dim light. Such light is most prevalent at dusk and has a wavelength between 400 and 500 nanometers (nm), with lower wavelengths penetrating to deep sea environments. Because the wavelength of maximal absorption,  $\lambda_{\text{max}}$ , can be easily assayed in the lab, its biological function is straightforward to assess quantitatively. As might be expected, deep-sea fish tend to have rhodopsins with a lower  $\lambda_{\text{max}}$  than shallow water fish and terrestrial animals, ~480 nm versus ~500 nm.<sup>20</sup> The researchers therefore sought to characterize each amino acid change leading to these different  $\lambda_{\text{max}}$  values.

Across the rhodopsin's length of 354 amino acids, 203 positions underwent an amino acid change but only twelve of these affected  $\lambda_{\text{max}}$ . The team's more recent work on visual pigments yields similar results: "[T]he effects of the small proportion of adaptive sites on the evolutionary rates are buried among those of the neutral changes."<sup>21</sup> Most amino acid changes that are accepted during visual pigment evolution have little or no discernible effect on function. Statistical methods of sequence analysis generally fail to predict those that do.<sup>22</sup> With notable exceptions, the same is true of most proteins—neutral changes dominate.<sup>23</sup>

The researchers made another stunning observation: similar shifts in  $\lambda_{\text{max}}$  and specific amino acid changes both recurred multiple times in evolution. The independent recurrence of the same evolutionary substitutions could result from at least two mechanisms. In the first, all mutations occur at similar rates, but those causing functional differences are promoted by natural selection. In the second, specific mutations recur at disproportionately high rates. Although selection is sure to have played an important role in protein evolution, the fact that specific mutations recur, sometimes multiple times, points to the contribution of the second option: there is a bias in the mutational input itself.

Take the most important example in mammals: the CpG dinucleotide, that is, a C followed by a G on the same DNA strand. In most of the genome, C → T mutation rates are ten times higher at CpG sites than at Cs in other contexts.<sup>24</sup> Because they are more likely to occur, they are also more likely to recur, substantially biasing the pool of variation upon which selection can act. Although the C → T mutation rate is not quite so high in protein-coding genes, it is still elevated. Given that new amino acids arise due to changes in the underlying DNA, it is conceivable that the known CpG bias could give rise to a bias in protein evolution. The DNA triplet ACG, for example, encodes the amino acid threonine. But because ACG contains a CpG dinucleotide, it is subject to an elevated rate of C → T transitions resulting in ATG, which encodes methionine. The extent

to which protein evolution is driven by directionality in the mutational input itself remains to be systematically evaluated, but work by Jay Storz and colleagues suggests it is substantial.<sup>25</sup>

**A**NCESTRAL PROTEINS CAN only be inferred for modern proteins similar enough to be grouped into families, of which S100s are one example. This leaves out deeper protein relationships between families. As a rule of thumb, a pair of proteins matching at fewer than ~30% of their positions cannot be confidently aligned. This is because such levels of similarity are likely due to chance alone.<sup>26</sup> As a result, it is only possible to scratch the surface of evolutionary history—only those proteins which diverged relatively recently remain similar enough to compare with confidence. The deepest questions about the origins of novel gene families remain shrouded in mystery.

This includes protein specificity. The ancestor of S100A5 and S100A6 may not have been more general than its descendants—but is the same true of other proteins and protein families? If so, how far back can this trend, or lack thereof, be extrapolated? Which targets were actually present in their environments? How did the primordial archetypes which gave rise to the modern protein families evolve, and had their own ancestors been more general? And, are more general proteins easier to chance upon in sequence space, that is, could they be reasonably expected to have arisen as evolutionary starting points?

To answer these questions, a lot more work just like that of Wheeler and Harms will need to be done.<sup>27</sup>



1. Linus Pauling and Emile Zuckerkandl, "Chemical Paleogenetics: Molecular 'Restoration Studies' of Extinct Forms of Life," *Acta Chemica Scandinavica* 17 (1963): S9–16.
2. Eric Gaucher et al., "Inferring the Palaeoenvironment of Ancient Bacteria on the Basis of Resurrected Proteins," *Nature* 425, no. 6,955 (2003): 285–88, doi:10.1038/nature01977.
3. David Liberles, ed., *Ancestral Sequence Reconstruction* (New York: Oxford University Press, 2007).
4. Joseph Felsenstein, "Phylogenies and the Comparative Method," *The American Naturalist* 125, no. 1 (1985): 1–15, doi:10.1086/284325.
5. Roy Jensen, "Enzyme Recruitment in Evolution of New Function," *Annual Review of Microbiology* 30, no. 1 (1976): 409–25, doi:10.1146/annurev.mi.30.100176.002205.
6. Valeria Risso, Jose Sanchez-Ruiz, and S. Banu Ozkan, "Biotechnological and Protein-Engineering Implications of Ancestral Protein Resurrection," *Current Opinion in Structural Biology* 51 (2018): 106–15, doi:10.1016/j.sbi.2018.02.007.

7. Olga Khersonsky and Dan Tawfik, "Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective," *Annual Review of Biochemistry* 79, no. 1 (2010): 471–505, doi:10.1146/annurev-biochem-030409-143718. For a detailed survey of the work undertaken by Dan Tawfik and laboratory, see Tyler Hampton, "Dan S. Tawfik Group: The New View of Proteins," *Inference: International Review of Science* 1, no. 1 (2014), doi:10.37282/991819.14.8.
8. Lucas Wheeler et al., "Conservation of Specificity in Two Low-Specificity Proteins," *Biochemistry* 57, no. 5 (2018): 684–95, doi:10.1021/acs.biochem.7b01086; Anne Bresnick, David Weber, and Danna Zimmer, "S100 Proteins in Cancer," *Nature Reviews Cancer* 15, no. 2 (2015): 96–109, doi:10.1038/nrc3893.
9. Rosario Donato et al., "Functions of S100 Proteins," *Current Molecular Medicine* 13, no. 1 (2013): 24–57, doi:10.2174/156652413804486214.
10. Bresnick, Weber, and Zimmer, "S100 Proteins in Cancer."
11. Danna Zimmer et al., "Evolution of the S100 Family of Calcium Sensor Proteins," *Cell Calcium* 53, no. 3 (2013): 170–79, doi:10.1016/j.ceca.2012.11.006; Sudhir Kumar et al., "Time-Tree: A Resource for Timelines, Timetrees, and Divergence Times," *Molecular Biology and Evolution* 34, no. 7 (2017): 1,812–19, doi:10.1093/molbev/msx116.
12. Susumu Ohno, *Evolution by Gene Duplication* (New York: Springer-Verlag, 1970); Austin Hughes, "The Evolution of Functionally Novel Proteins after Gene Duplication," *Proceedings of the Royal Society B: Biological Sciences* 256, no. 1,346 (1994): 119–24, doi:10.1098/rspb.1994.0058.
13. Bresnick, Weber, and Zimmer, "S100 Proteins in Cancer."
14. Lucas Wheeler et al., "Learning Peptide Recognition Rules for a Low-Specificity Protein," *Protein Science* 29, no. 11 (2020): 2,259–73, doi:10.1002/pro.3958.
15. Ziheng Yang, "PAML 4: Phylogenetic Analysis by Maximum Likelihood," *Molecular Biology and Evolution* 24, no. 8 (2007): 1,586–91, doi:10.1093/molbev/msm088.
16. Wheeler et al., "Conservation of Specificity in Two Low-Specificity Proteins," 693.
17. Lucas Wheeler and Michael Harms, "Were Ancestral Proteins Less Specific?" *Molecular Biology and Evolution* 38, no. 6 (2021): 2,227–39, doi:10.1093/molbev/msab019.
18. Wheeler and Harms, "Were Ancestral Proteins Less Specific?"; Geeta Eick et al., "Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty," *Molecular Biology and Evolution* 34, no. 2 (2017): 247–61, doi:10.1093/molbev/msw223.
19. Shozo Yokoyama et al., "Elucidation of Phenotypic Adaptations: Molecular Analyses of Dim-Light Vision Proteins in Vertebrates," *Proceedings of the National Academy of Sciences* 105, no. 36 (2008): 13,480–85, doi:10.1073/pnas.0802426105.
20. Austin Hughes, "The Origin of Adaptive Phenotypes," *Proceedings of the National Academy of Sciences* 105, no. 36 (2008): 13,193–94, doi:10.1073/pnas.0807440105.
21. Shozo Yokoyama et al., "A Simple Method for Studying the Molecular Mechanisms of Ultraviolet and Violet Reception in Vertebrates," *BMC Evolutionary Biology* 16, no. 1 (2016): 64, doi:10.1186/s12862-016-0637-9.
22. Hughes, "Origin of Adaptive Phenotypes."
23. Masatoshi Nei, *Mutation-Driven Evolution* (Oxford: Oxford University Press, 2013).
24. Alan Hodgkinson and Adam Eyre-Walker, "Variation in the Mutation Rate across Mammalian Genomes," *Nature Reviews Genetics* 12, no. 11 (2011): 756–66, doi:10.1038/nrg3098.
25. Jay Storz et al., "The Role of Mutation Bias in Adaptive Molecular Evolution: Insights from Convergent Changes in Protein Function," *Philosophical Transactions of the Royal Society B: Biological Sciences* 374, no. 1,777 (2019): 20180238, doi:10.1098/rstb.2018.0238.
26. Burkhard Rost, "Twilight Zone of Protein Sequence Alignments," *Protein Engineering, Design and Selection* 12, no. 2 (1999): 85–94, doi:10.1093/protein/12.2.85.
27. I am indebted to Michael Harms for discussion and Zachary Ardern for feedback.

DOI: 10.37282/991819.21.36

BIOGRAPHIES

**John Horton Conway**

**The Game  
of Life**

*Daniel Kleitman*



*Published on May 12, 2021*

**T**HE BRILLIANT mathematician John Horton Conway died of complications from COVID-19 in April 2020. He was eighty-two years old. His career included significant contributions to many domains and earned him some of the most prestigious prizes awarded in mathematics.

Most of us experience mathematics as a collection of rules and results, and it is widely thought that a mathematician is someone who has expert knowledge of those rules and results. What he does with that knowledge is a mystery. This view is not entirely wrong, but it is misleading. Mathematicians achieve success through their research, and mathematical research has two aspects. The first involves the creation of problems—often, but not always, conjectures about why two dissimilar mathematical structures have the same size. The second involves proving or disproving such conjectures. Conway was a great master of the second kind. He preferred to ignore both the received wisdom and the received opinion about a mathematical field in order to strike out on his own. He would reinvent the subject from scratch, developing his own ideas about it, and in Conway's hands, this often led to important progress.

As a small child, Conway was fascinated by numbers. He wondered where the integers went, and calculated the values of very many powers of two for his own amusement. These early interests and experiences seem to have caused the parts of his brain that involved mathematical reasoning to develop extensively. His mathematical talents were always extraordinary, and his academic success brought him to the University of Cambridge, where he continued to excel. During his time as a graduate student, Conway solved a problem proposed by his advisor, determining that all integers could be represented as the sums or differences of 37 fifth powers. As it turned out, another researcher published the same result before Conway wrote up his own work.

Cambridge retained Conway as a postdoctoral fellow, and he began to study infinite ordinal numbers, a subject that never really appealed to him. Over time, his personality developed, and a shy introvert gradually became a gregarious extrovert and a spellbinding lecturer. He spent a great deal of time with graduate students, playing games, analyzing game strategies, and inventing games, activities that some considered time wasting.

In 1967, Conway was almost thirty, the age at which mathematicians are said to start losing their powers. He had not yet published a research paper, and he was starting to feel the pangs of depression. Suddenly, his life changed. A mathematician named John Leech visited Cambridge and tried to interest the mathematicians in a problem related to his own work. Most were busy with their own activities; they listened politely then promptly forgot Leech's problem. Conway was seeking an outlet for his talents and vowed to solve it. And he did.

Almost everyone knows that seven pennies can be arranged so that six of them surround the seventh, touch-

ing it and their two neighbors. With an infinite number of pennies, the plane can be filled so that each penny touches six of its neighbors.

The analogous arrangement cannot be made in three dimensions, using marbles instead of pennies.

Why not?

In three dimensions, up to twelve marbles of the same size can be placed around a given marble, but they will not all touch their neighbors. Six marbles, for example, can be placed around that given marble so that all seven centers lie in one plane, as in two dimensions. On either side of these six marbles, three more marbles, all touching one another, can be placed, for a total of twelve marbles around that given marble. This is the most that there is room to accommodate. But all of the additional marbles cannot, at the same time, touch two of the first six.

Surprisingly, there exists a set of neighbors—all meeting the first marble, or ball, and also meeting all their neighbors—that can be extended indefinitely, as in one or two dimensions, in eight and twenty-four—spaces described by eight or twenty-four numbers—but in no other dimensions at all. Leech had found a way to describe the locations of the centers of all the balls that met all their neighbors in twenty-four dimensions, now known as the Leech lattice, and he wanted to verify that it worked. He asked if someone at Cambridge could describe the symmetries of his construction.

The construction itself is quite large, with each ball having 196,560 neighbors, all of which it touches. As far as its symmetries go, each of these neighbors can be mapped into each of the others. If two neighbors are fixed, there are on the order of 196,560 ways of arranging the others to produce a symmetry, making the total number of symmetry operations roughly  $8 \times 10^{18}$ , all told.

Conway resolved to find a comprehensible description of these symmetries, which form what mathematicians call a group. To determine a group, it is sufficient to find its generators, members from which all the other members can be produced by means of repetitions and combinations. In a short time, he had an idea that worked, and he was able to find such a description of the group.

At the time, group theory was a hot topic in mathematics. Many mathematicians were working to determine all of the finite simple groups, the very atoms of finite group theory. There were a number of infinite families of such groups that were symmetries of standard mathematical constructs that existed in all dimensions. There were also a few others, known as sporadic groups, and these, it was hoped, were all of them. But nobody could prove that, and for good reason.

The group that Conway described, and others closely related to it, were new simple groups that no one had previously encountered. In a relatively short time, mathematicians closed this story by demonstrating that Conway's groups were the end of it. Old and new simple groups included all possible simple groups.

His result was not a mere curiosity. It provided the missing link that allowed the resolution of the big problem of group theory. That resolution opened a new way to prove things about groups: if something holds for all simple groups, it has to be true for all groups.

Conway quickly rose from being a time waster to being a “magical genius,” as his colleague Simon Kochen would later refer to him.<sup>1</sup> Even Conway’s time wasting became exciting. During the 1950s, Stanislaw Ulam and John von Neumann introduced mathematicians and computer scientists to cellular automata. Consider an infinite grid that contains an initial configuration of 0s and 1s. For each cell, the rule that determines its successor depends on its neighbors. It was thus that Conway introduced his famous Game of Life. The rule governing the Game of Life was very simple: a cell gets a 1 only if either the sum of the entries in the eight surrounding cells is three, or the sum of the entries in the cell and the surrounding cells is three. All sorts of wonderful things happen, depending on the starting configuration. Anyone can set this up on a spreadsheet, and there are websites that can display the development over time. Fiddling with this set of rules has been immensely popular and has made Conway’s name well known outside the mathematics community.

The Game of Sprouts is another of his popular time wasters. It is a two-person game in which some number of points are drawn in the plane. Players alternately draw an arc between two points and add a new point somewhere in the middle of the arc. A point can meet at most three lines, two of which may be both ends of the same line.

Another of Conway’s concoctions is the concept of surreal numbers. In the nineteenth century, Richard Dedekind defined real numbers in terms of the rational ones by identifying a real number with the set of rational numbers less than it. The real numbers between 0 and 1 can be described as a decimal point followed by all infinite sequences of 0s and 1s, in binary notation. The rational numbers less than any real number include all finite-length prefixes of that real number, and the set of all of these defines the real number.

Conway went a step further than Dedekind. For any two sets of numbers  $A$  and  $B$  with the largest in  $A$  less than the smallest in  $B$ , he defined  $\langle A | B \rangle$  as follows. When  $A$  and  $B$  are both the empty set, then  $\langle | \rangle$  is zero. If  $B$  is the empty set and the largest defined number in  $A$  is  $n$ , then

$\langle A | B \rangle$  is  $n + 1$ . When the empty set lies to the right of the vertical bar and  $n$  is the largest number to the left of it, then the result is  $n + 1$ . When the largest number in  $A$  is  $x$  and the smallest in  $B$  is  $y$ , then  $\langle A | B \rangle$  is the number halfway between them. If the process of definition is continued infinitely in all possible ways, one gets all the integers, all the fractions whose denominators are powers of 2, and eventually all the fractions and all the real numbers. So far, this construction is somewhat like Dedekind’s. Conway then took a step further to include  $\langle A | B \rangle$  when  $A$  consists of zero and  $B$  consists of all positive rational, or real, numbers. This is something new. It has been called *epsilon*, because it is just like the symbol for an infinitesimal quantity that tortures calculus students. The resulting numbers are termed *surreal*. Rules for adding, multiplying, and otherwise combining surreal numbers are consistent with those we expect, and many things can be proven about them. This was a way to introduce an infinitesimal into number systems in a natural sort of way.

These subjects represent only a small portion of Conway’s contributions to mathematics. Each led to collaborations with other mathematicians who saw ways to use his approaches in the problems of interest to them.

Conway left Cambridge and went to Princeton in 1987, where he remained until his death. He had a gift for making complicated mathematics appear simple, both in speaking and in the many popular books he wrote.<sup>2</sup> The breadth and the depth of his contributions to mathematics were extraordinary.



1. Catherine Zandonella, “Mathematician John Horton Conway, a ‘Magical Genius’ Known for Inventing the ‘Game of Life,’ Dies at Age 82,” *Princeton.edu*, April 14, 2020.
2. Among his volumes are *The Book of Numbers* (New York: Copernicus, 1996); *On Numbers and Games* (New York: Academic Press, 1976); *Regular Algebra and Finite Machines* (London: Chapman and Hall, 1971); *Sphere Packings, Lattices, and Groups* (New York: Springer Verlag, 1988); as well as several co-authored books.

DOI: 10.37282/991819.21.4

# Letters to the Editors

*Sheldon Lee Glashow*

*Kenneth Ford with Jeremy Bernstein*

*Hany Farid*

*Viktor Toth with Jean-Pierre Luminet*



## A Brief History of the Muon

*To the editors:*

My old and dear friend Eduardo de Rafael is a renowned Spanish-French theoretical physicist and a talented amateur pianist. He and his charming English wife live and work in Provence. His superb review of the controversy concerning an apparent departure between observed and calculated values of the muon's magnetic moment leaves little room for criticism, so instead I offer a bit of back-ground.

Carl D. Anderson and Seth Neddermeyer, at Caltech, discovered muons in 1936 by examining the tracks left by cosmic rays as they traversed a cloud chamber flown at high altitude. Anderson, four years before, had discovered positrons, the antiparticles of electrons. Muons were the sixth of what were then regarded as elementary particles, after electrons, positrons, photons, neutrons, and protons. Neutrinos had been postulated by Wolfgang Pauli in 1930, but they were not observed until 1956.

Electrons and muons, generically referred as leptons, are electrically charged particles, as are their neutral counterparts called neutrinos. Today three varieties of charged leptons are known, each with its own neutrino. Leptons do not interact strongly among themselves or with other particles; they seem to be truly elementary. Protons and neutrons, as well as all other strongly interacting particles, are called hadrons. They are not elementary at all, but are made up of quarks which are held together by gluons.

When first discovered, muons were thought to be the particles that had been imagined by Hideki Yukawa in 1934 as the mediators of nuclear forces. Some physicists, including Niels Bohr, dubbed them yukons. But muons were soon found not to interact strongly. Yukawa's hypothetical particles would be discovered among cosmic rays in 1947. They are now called pions. The pion is a hadron made of one quark and one antiquark.

The magnetic moment of any charged lepton is specified by a dimensionless number called  $g$ , its gyromagnetic ratio. The equation Paul Dirac famously invented in 1928 implies that  $g = 2$ , but physicists soon realized that Dirac's "zeroth order" value of  $g$  must be modified by calculable quantum electrodynamical radiative corrections. The lowest order contribution to the anomaly was first computed by my thesis advisor Julian Schwinger in 1947:  $g - 2 = \alpha / 2\pi$ , for both electrons and muons. The anomaly was a mere tenth of a percent, but its theoretical and measured values agreed.

Since then, theoretical and measured values of the muon magnetic moment anomaly have been determined with ever increasing precision, remaining in lockstep agreement with one another until now. Today they seem to have grown slightly apart. Is this due to experimental error, or computational error, or to physics lying beyond the Standard Model? A recent measurement of  $g - 2$  at Fermilab

confirms the earlier CERN result, but uncertainty reigns among the most precise attempts to compute  $g - 2$  from theory. Because both leptons and hadrons are subject to electroweak forces, higher-order contributions to the muon anomaly which involve virtual hadron loops must be evaluated to achieve computational precision comparable to today's experimental precision, and there's the rub! The strong force, quantum chromodynamics or QCD, cannot be addressed perturbatively, as can the electroweak force. De Rafael mentions two attempts to achieve the required theoretical precision: one is based on lattice QCD, the other on the extrapolation of indirectly relevant experimental data. In the former case, experiment and theory are not in severe conflict; in the latter case they are. Until and unless this theoretical conflict is resolved, we will not know whether or not our Standard Theory needs revision. De Rafael will tell us as soon as the veil is lifted.

*Sheldon Lee Glashow is a Nobel Laureate, Higgins Professor of Physics, emeritus, at Harvard University, and University Professor, emeritus, at Boston University.*

DOI: 10.37282/991819.21.57

## More Secrets

*To the editors:*

In May of 1950, when I was a 24-year-old physics graduate student at Princeton, I accepted the invitations of my mentor John Wheeler and Wheeler's friend Edward Teller to join an accelerated effort in Los Alamos to design an H-bomb. Wheeler had given me a soft sell. Teller came to town and gave me a hard sell. The force propelling the acceleration was the Soviet Union's first successful A-bomb test less than a year earlier.

I applied for a Q clearance and it was granted in a few weeks. By late June I was at work in Los Alamos.

Adhering to the lab's security requirements was not hard. Most challenging for me was the requirement that all of one's work be recorded in bound notebooks. There were to be no notepads, no loose pieces of paper, and no wastebaskets containing half-baked thoughts on nuclear matters. I mostly adhered to this requirement. So now, on some shelf in Los Alamos, sit bound notebooks containing my thoughts and calculational efforts. Then, as now, I used a fountain pen.

Taking classified information out of the lab was not hard. Briefcases were not checked at the guard gate, nor were they examined at the airstrip. On the few trips east that I took during my year there, I only put secret information in my briefcase, never in checked luggage nor in parcels to be stowed under the seat ahead. When the stewardesses did not notice or care, I put the briefcase behind

my lower legs next to the front of my seat—otherwise it was under my arm on one side of the seat. In either location, if the briefcase were touched while I was napping, I would wake.

Conversation at the lab was free and open. Only once did I come up against need-to-know. I was present in a small conference room where Hans Bethe, on one of his fairly frequent visits to the lab, was meeting with some of the lab officials. When the conversation turned to the analysis of air samples collected by US aircraft in order to learn more about Soviet atomic explosions, I was shooed from the room. I had no need to know. But that incident was the exception.

The next year, 1951–52, back at Project Matterhorn in Princeton, we had 24-hour guard service in the shack where we worked. Of course we talked outside of that building and even wrote things on pieces of paper. But we were never casual about secrets and did our best to shelter them. In January 1953, Wheeler lost a classified document on an overnight train trip to Washington. It was quite an important one, containing information on the use of lithium deuteride as a thermonuclear fuel. Despite herculean efforts by Amtrak and the FBI, it was never found. President Eisenhower, learning of the incident, was outraged. Wheeler was exceedingly embarrassed.

One of my early jobs at Matterhorn in the summer of 1951 was hiring young women—yes, they were all women and all young—to perform secretarial chores, carry out numerical calculations, and draw graphs. All were swiftly granted Q clearance. I spent part of the next year living in a rented room in Washington by day and running what we considered a powerful computer, the SEAC at the Bureau of Standards, by night. We used a program coded by John Toll, Wheeler, and me that followed the “burning” of the deuterium fuel in the forthcoming test of the first H-bomb, codenamed “Mike.” For this work I needed an assistant. I hired a young man to operate the computer and deal with its hexadecimal output. There was no time to consider seeking a security clearance for him, so he worked without a clearance and with no real idea of what we were calculating. I could only assure him that it had to do with national security and was very important. Perhaps John von Neumann could have unraveled our machine-language code. I do not think anyone else could have. In any event, we operated without safes, without guards, and without bound notebooks.

Sixty years later, I began work on a book that, when it was published in 2015, bore the title *Building the H Bomb: A Personal History*.<sup>1</sup> Its purpose, as stated in the preface, was to inform interested nonscientists about the history of the world’s first H-bomb, provide a memoir of a two-year slice of my life, and serve as a mini-textbook on nuclear physics.

As I dug up recollections and records, did fact-checking—with the help of Anthony Eames, a very helpful graduate student at George Washington University—and wrote, now

with a laptop rather than a fountain pen, it never occurred to me that there might be anything in the book that would catch the attention of the nation’s guardians of secrets. Everything in the book, I was sure, was by that time in the public domain. Yet, in a conversation in early 2014 with the Department of Energy’s historian, Terry Fehner, who had supplied me with some facts, Fehner said, “You know, Ken, since you held a Q clearance and were dealing with secret information in the time period you are writing about, you really should submit your manuscript to DOE for a security review.”

“I hadn’t thought about that,” I replied. “How long would such a review take?”

“Probably about a month.”

“OK, I can afford that. I’ll submit it for a pro forma review.”

In June 2014, I did so, and in August I heard back: “Our team is quite taken with your manuscript,” was the message. “However, some concerns have been identified.” After some back and forth, I agreed to meet with these keepers of the secrets at their office in Germantown, Maryland, for a discussion.

On the appointed day, I showed up around 9:00 a.m., and was told that for me to participate in a discussion of my manuscript, I needed Q clearance. The necessary photographing, fingerprinting, and form-filling was completed expeditiously. By 11:00 I was again a cleared scientist, a status that would last several hours. Then I sat down in the office of Andrew Weston-Dawkes, the head of the security section. A couple of other people were present. One of them had in his hands a copy of my manuscript with a forest of yellow Post-it notes sticking out from its pages. I was alarmed.

The message to me at this meeting ran along the following lines:

We like your book and would be happy to issue it as a DOE report, classified secret, so that it could reach people qualified to read it. Also, we would be glad to provide you with an office here in which you could complete and polish the manuscript, and work with our publishing team. If you want to publish the book commercially, you will have to make many changes and delete much material.

I declined the kind offer and said I wanted to publish the book commercially. We agreed that I would be sent a detailed list of everything in the book that the DOE officials wanted removed or changed. We parted on good terms and I went home, no longer in possession of a Q clearance.

About two months later, in November 2014, I received the list. It contained 60 items. Some of them called for a change, but most called for a deletion. In total, I was told to delete about 5,000 words, or about 10 percent of the book. In my response, I wrote, “Were I to follow all—or even most—of your suggestions, it would destroy the book.”

I worked my way through the 60 items, to see if any had merit. I found none that did. In some cases I could cite unclassified sources—although I acknowledge the

argument that by stating something that has appeared elsewhere, I am authoritatively confirming its truth. Some were descriptions of events, say, in 1945 or in 1960, of which I had no direct knowledge. One was a statement that the main cylinder of the “Mike” device was about 20 feet tall and about 7 feet in diameter. This does indeed sound like classified information, but the DOE officials had no objection to including in the book a photograph of this device next to people, a Jeep, and a forklift that provided a scale of size.

I sent back my rebuttal. After a little more back and forth, I said I would replace “in fact” by “reportedly” in two places. In January 2015, we agreed that we were at an impasse and that there was no basis for further negotiation.

Going ahead required that I feel safe and that the publisher, World Scientific, felt protected. By good fortune, a law firm in Washington, DC, after hearing my story in detail and examining the evidence, agreed to defend me pro bono if it came to that. That took care of my safety. Every publishing contract, of course, assures the publisher that the author alone is responsible for any errors, false or libelous statements, and so on. After some amendment to that clause of the contract to make clear that any breaches of national security were my responsibility, World Scientific agreed to go ahead. And go ahead we did. The book was published in March 2015.

A little before the publication date I contacted Dennis Overbye, a senior science writer at the *New York Times* whom I had met earlier, telling him that I didn’t have a story for him yet but wanted to alert him to the possibility of a story if the DOE came after me. He relayed my message to his colleague William Broad, who said I was wrong, that in fact I had a story already. Broad came down to Philadelphia to spend some time with me, hear my account, and look at the written evidence, including my exchanges with the DOE officials. His article appeared in the science section of the *Times* on March 23, 2015.<sup>2</sup>

Nothing happened. No reverberations. No communication. Just an uptick in sales.

What do I conclude from all of this? I continue to think that secrets are necessary and that people entrusted with them do, with rare exceptions, take very seriously their responsibility to safeguard those secrets. Yet secrets are like helium. They escape from almost any container. They can’t be bottled for long. It is foolish to try to keep secrets for many decades. Andrei Sakharov, in a 1990 memoir that covers his work on the first Soviet H-bomb, refers to the first idea, the second idea, and the third idea that made possible the Soviet Union’s successful H-bomb program.<sup>3</sup> In a 2004 book about Sakharov, Gennady Gorelik identifies these ideas, in effect, declassifying them.<sup>4</sup> So these ideas originated in the 1950s, were first publicly named in 1990, and first publicly explained in 2004. That is a very long shelf life indeed for a secret.

For the record, the first idea was to alternately layer fission and fusion fuel, the second idea was to use lithium

deuteride as a fusion fuel, and the third idea was to use radiation implosion to compress the fusion fuel. Sakharov used the term “layer cake” for the first idea, and it was implemented. Teller had the same idea and called it an alarm clock. It was not implemented. All three ideas circulated in our Los Alamos group in 1951. Historians agree that they arose independently in the Soviet Union without the help of espionage.

What about the keepers of the secrets? I bear them no ill will. They are doing a job that has been defined for them. But secrets erode with time and they leak. There’s not much these servants of secrecy can do about it.

**Kenneth Ford**

**Jeremy Bernstein** replies:

Reading Ken Ford’s comments brought back some of my own encounters with clearance. As I noted in my essay about the Orion project,<sup>5</sup> I spent the early parts of the summer of 1958 at the RAND Corporation in Los Angeles. The headquarters was set inside what looked like a small university campus in Santa Monica not far from the beach. There were two levels of security. You needed a pass to get into the building and a full Q clearance to get past the guard for the physics group. At the time, the group was working on matters related to the hydrogen bomb. No one had anything for me to do until one day I was handed a long list of figures and asked to add them up. I was told that it had something to do with high altitude testing with a possible connection to intercontinental missile defense. It was around this time that Freeman Dyson rescued me with an invitation to move to La Jolla to work on the Orion.

Since it was going to be powered by nuclear weapons, a Q clearance was needed to work on the Orion project. But, as far as I was concerned, this was an irrelevance since what I did with Dyson had nothing to do with these weapons. But it did affect Richard Feynman. He refused to become a consultant. He said that he could not guarantee that he would not talk about what he learned about the project because he always talked about anything that excited him. What I worked on with Dyson was opacity—the absorption of radiation by plasmas, and the like. Dyson had proven a remarkable theorem on the maximum possible opacity using quantum mechanics. My job was to compare it with data. I knew that RAND was working on opacity for things like uranium and that this work was classified. We stuck to light elements. I persuaded Dyson to visit RAND with me to present his theorem. The lecture became something of a farce because Dyson did not know the clearance status of all the people attending and was unable to say anything about Orion. Although we had confined our detailed work to light elements, which avoided all the issues of classification, Dyson’s derivation applied to all elements.<sup>6</sup> If he had been allowed to present this material, the attendees could have

used it to check the validity of their approximation methods and perhaps saved themselves from a good deal of grief.

This connection with RAND led to an even more farcical situation. A few years later, I took a job at the Brookhaven National Laboratory. There was some classified work being done at the lab. I never knew what this was, but it meant that there was some level of security. One day I got a phone call from security letting me know that a classified document had arrived from RAND and asking for my permission to open it. I replied that they could open it, but only if I could read it. I was sure that it had something to do with the high altitude nuclear tests. I was told that I was not authorized to read it because my Q clearance had lapsed. The document was sent back to RAND unopened. I was listed on the cover as one of the co-authors.

**Kenneth Ford** is a retired physicist and writer, and was the founding chair of the Physics Department at the University of California, Irvine.

**Jeremy Bernstein** is Professor Emeritus of Physics at the Stevens Institute of Technology.

1. Kenneth Ford, *Building the H Bomb: A Personal History* (Singapore: World Scientific, 2015).
2. William Broad, “Hydrogen Bomb Physicist’s Book Runs Afoul of Energy Department,” *The New York Times*, March 23, 2015.
3. Andrei Sakharov, *Memoirs* (New York: Alfred A. Knopf, 1990).
4. Gennady Gorelik, *The World of Andrei Sakharov: A Russian Physicist’s Path to Freedom* (New York: Oxford University Press, 2004).
5. Jeremy Bernstein, “Reflections on Project Orion,” *Inference: International Review of Science* 5, no. 2 (2020), doi:10.37282/991819.20.4.
6. Our work was finally published in December 2003. Jeremy Bernstein and Freeman Dyson, “Opacity Bounds,” *Publications of the Astronomical Society of the Pacific* 115, no. 814 (2003): 1,382–87, doi: 101086/380420.

DOI: 10.37282/991819.21.21

## On Algorithmic Amplification

To the editors:

“If we expect no more from citizens than that they gratify their appetites,” Adam Garfinkle warns, “they will be enticed by spectacle and easily taken in by lies. Disinformation will thrive. As it is thriving now.” It is a justifiable concern. Yet the culprits for this plunge into delusion are not the deepfakes and generative adversarial networks (GANs) implicated in Garfinkle’s essay. The real culprit is more mundane and pernicious: the recommendation

algorithms that curate the information surfaced by search engines and social media.

By mid-May of 2020, in the midst of the global pandemic, 28% of Americans believed Bill Gates was planning to use COVID-19 to implement a mandatory vaccine program with tracking microchips.<sup>1</sup> Belief in this conspiracy is not unique to Americans. In our own global surveys, we find that across Central and South America, the Middle East, Northern Africa, the United States, and Western Europe, 20% of the public believes this bizarre claim.<sup>2</sup> Machine-synthesized audio, images, and video—so-called deepfakes—were not part of the creation or spread of this conspiracy. Instead, the conspiracy spread through simple social media posts.

The far-reaching, far-right QAnon conspiracy claims, among other things, that a cabal of Satan-worshipping cannibalistic pedophiles and child sex-traffickers plotted against Donald Trump during his term as US President. A recent poll finds 37% of Americans are unsure whether QAnon is true or false, and 17% believe it to be true.<sup>3</sup> Deepfakes were not part of the creation or spread of this conspiracy. This conspiracy was also created and spread through simple social media posts—with Trump’s tacit endorsement.<sup>4</sup>

A pair of videos depicting Speaker Nancy Pelosi looking inebriated during public appearances were not, as Garfinkle states, GAN-generated deepfakes.<sup>5</sup> The low-tech videos were created by simply slowing the original footage by 25% to simulate slurred speech. Sophisticated technology was not needed to convince millions that Pelosi was “blowed [sic] out of her mind,” as the accompanying social media claimed.

Deepfakes and technologies designed to deceive are not the common thread that connects Bill Gates’s COVID microchips, QAnon’s Satan-worshipping cannibals, and the litany of conspiracies and outright lies polluting the internet. The common thread is the recommendation algorithms that aggressively promote the internet’s flotsam and jetsam onto news feeds and watch lists, plunging users into increasingly isolated echo chambers devoid of reality.

The classic thought exercise asks, “If a tree falls in a forest and no one is around to hear it, does it make a sound?” The modern equivalent should ask, “If information is on Facebook and an algorithm doesn’t promote it onto a news feed, does it exist?”

Every day, four petabytes—more than four million gigabytes—of data are uploaded to Facebook. But not all of this content is equal in the eyes of Facebook’s management. In 2009, the platform removed the ability for users to chronologically sort their news feed, turning over editorial control to algorithmic curation. Facebook’s own internal researchers found their algorithms “exploit the human brain’s attraction to divisiveness.” The researchers went on to conclude that, if left unchecked, the recommendation algorithms will promote “more and more

divisive content in an effort to gain user attention and increase time on the platform.” A separate internal study found that 64% of people who joined an extremist group on Facebook did so because of the company’s recommendation. Facebook’s leadership has chosen to ignore these findings.<sup>6</sup>

Every minute of every day, more than 500 hours of video footage are uploaded to YouTube. The likelihood that any video is widely seen depends largely on recommendation algorithms. Seventy percent of watched YouTube footage is recommended by the company’s algorithms.<sup>7</sup> By 2016, Twitter and Instagram had joined Facebook and YouTube in unleashing attention-grabbing recommendation algorithms to control what users read, see, hear, and—ultimately—believe.

Tech companies implemented recommendation algorithms because they proved remarkably successful at manipulating users. By maximizing clicks, likes, and shares, recommendation algorithms also helped maximize profits. These algorithms also succeeded in creating vicious feedback loops. If a user searches for content about QAnon on these platforms, whether innocently or otherwise, the algorithms will recommend additional QAnon-related content. A few clicks here, a few clicks there, and the user will be quickly ushered down a rabbit hole from which escape will prove difficult.<sup>8</sup> This algorithmic amplification is the root cause of the unprecedented speed and reach with which misinformation is spreading online.

The true power of these algorithms could be seen when Facebook tweaked their recommendation algorithms to favor authoritative news organizations in the days before the 2020 US presidential election.<sup>9</sup> The changes worked exactly as intended, reducing election-related misinformation on the platform. Nonetheless, Facebook reversed course just days later when it was found that the new algorithm also reduced the average time users spent on Facebook and, in turn, ad revenue.

Garfinkle is justified in calling out the potential dangers of deepfakes, which have been used to commit fraud and have been weaponized against women in the form of non-consensual, sexually explicit material.<sup>10</sup> But the more significant risk associated with deepfakes comes in the form of the liar’s dividend.<sup>11</sup> The increasing ease with which images, audio, and video can be convincingly manipulated means that they can also be far more easily dismissed as deepfakes, regardless of their truthfulness or authenticity. More than being duped by a deepfake, it is this excuse to selectively deny the truth that is perhaps the most serious threat posed by deepfakes.

The larger issue of misinformation can be laid squarely at the feet of the tech companies and the recommendation algorithms they have adopted to titillate and outrage users, so that they spend as many hours as possible clicking, liking, sharing, and tweeting.

The current crisis, and the specter of an impending infocalypse,<sup>12</sup> was both foreseeable and preventable. This

is maddening. As Facebook demonstrated, if only for a few days, recommendation algorithms can be adjusted to favor the authoritative, trustworthy, and civil. All that is required is the moral fortitude—or federally mandated legislation<sup>13</sup>—to recognize the devastating impact to individuals, societies, and democracies that social media is having, and to put aside corporate indifference and unchecked greed in favor of even a modicum of decency and social responsibility.

*Hany Farid is a Professor at the University of California, Berkeley, with a joint appointment in Electrical Engineering & Computer Sciences and the School of Information.*

1. Linley Sanders, “The Difference between What Republicans and Democrats Believe to Be True about COVID-19,” *YouGov*, May 26, 2020.
2. Sophie Nightingale and Hany Farid, “Examining the Global Spread of COVID-19 Misinformation,” *arXiv* (2021), arXiv:2006.08830v2.
3. Mallory Newall, “More Than 1 in 3 Americans Believe a ‘Deep State’ Is Working to Undermine Trump,” *Ipsos*, December 30, 2020.
4. Brandon Carter, “Trump, Addressing Far-Right QAnon Conspiracy, Offers Praise for Its Followers,” *NPR*, August 19, 2020.
5. Saranac Hale Spencer, “Viral Video Manipulates Pelosi’s Words,” *Factcheck*, 2020.
6. Jeff Horwitz and Deepa Seetharaman, “Facebook Executives Shut Down Efforts to Make the Site Less Divisive,” *Wall Street Journal*, May 26, 2020.
7. Marc Faddoul, Guillaume Chaslot, and Hany Farid, “A Longitudinal Analysis of YouTube’s Promotion of Conspiracy Videos,” *arXiv* (2020), arXiv:2003.03318.
8. Julia Carrie Wong, “Down the Rabbit Hole: How QAnon Conspiracies Thrive on Facebook,” *The Guardian*, June 25, 2020.
9. Kevin Roose, “Facebook Reverses Postelection Algorithm Changes That Boosted News from Authoritative Sources,” *The New York Times*, December 16, 2020.
10. Lorenzo Franceschi-Bicchierai, “Listen to This Deepfake Audio Impersonating a CEO in Brazen Fraud Attempt,” *Vice News*, July 23, 2020; Mary Anne Franks and Ari Ezra Waldman, “Sex, Lies, and Videotape: Deep Fakes and Free Speech Delusions,” *Maryland Law Review* 78, no. 4 (2018): 892.
11. Robert Chesney and Danielle Citron, “Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security,” 107 *California Law Review* 1,753 (2019), doi:10.2139/ssrn.3213954.
12. The term infocalypse has been attributed to the technologist Aviv Ovadaya. Charlie Warzel, “Believable: The Terrifying Future of Fake News,” *BuzzFeed News*, February 11, 2018.
13. Tom Malinowski and Anna G. Eshoo, “Reps. Malinowski and Eshoo Introduce Bill to Hold Tech Platforms Liable for Algorithmic Promotion of Extremism,” press release, October 20, 2020.

## On Modified Gravity

To the editors:

Jean-Pierre Luminet's essay brilliantly summarizes the current state of affairs concerning dark matter. He also examines a number of alternatives, including modified gravity. The latter discussion is mostly limited to Mordehai Milgrom's modified Newtonian dynamics (MOND), with only a brief mention of Erik Verlinde's emergent gravity model.<sup>1</sup> In this singular focus, Luminet is not alone. Throughout much of the literature, MOND is presented as the archetypal modified gravity theory, despite its many shortcomings. In its original form, MOND is no more than a phenomenological modification of Newton's second law and as such does not even respect basic conservation laws.

In actuality, the literature of modified gravity is far richer, and the search for modified theories of gravitation began far earlier. Foremost among the modern theories of modified gravity is the scalar-tensor theory of Pascual Jordan, Carl Brans, and Robert Dicke.<sup>2</sup> This theory promotes the Newtonian gravitational constant to a dynamical scalar field, the value of which may vary in space and time. In its unmodified form, Jordan–Brans–Dicke theory runs afoul of precision tests that have since been performed using spacecraft in the solar system. Nonetheless, it still serves as a prototype for many new modified gravity theories.

The family of theories known as  $f(R)$  is closely related to Jordan–Brans–Dicke theory.<sup>3</sup> This nomenclature refers to the presence of the Ricci curvature scalar  $R$  in the Lagrangian formulation of general relativity first presented by David Hilbert, i.e., the Einstein–Hilbert Lagrangian. It is possible to replace the scalar  $R$  with a function  $f(R)$ , e.g.,  $f(R) = R^2$  or  $f(R) = 1/R$ , and construct a sensible theory of gravitation. Most such theories can be shown to be fundamentally equivalent to Jordan–Brans–Dicke theory, but there are subtle differences.

Entire families of modified gravity theories, including Jordan–Brans–Dicke theory and  $f(R)$  theories, can be investigated using the parameterized post-Newtonian (PPN) framework. The roots of this approach go all the way back to the pioneering work of Arthur Eddington, who first used such a parameterization to measure how a relativistic theory of gravitation deviates from the Newtonian prediction.<sup>4</sup> Modern versions of such parameterization also cover more exotic families of theories, including theories that introduce unusual features for the metrical field, such as torsion or nonmetricity.<sup>5</sup> As many of the PPN parameters are measured using observations such as precision radio navigation or lunar laser ranging, a number of theories can be excluded, but some tantalizing candidates remain.

Another strong constraint on modified theories of gravitation comes from the recent multispectral observation of GW170817.<sup>6</sup> The simultaneous arrival of a gravitational

wave signal and signals in the electromagnetic spectrum from a neutron star merger yield strong constraints on so-called bimetric theories. These are theories in which not all constituent fields couple to the same metrical field.<sup>7</sup> One such theory is Jacob Bekenstein's relativistic generalization of MOND, known as tensor–vector–scalar gravity (TeVeS).<sup>8</sup>

Other challenges may be less stringent. The apparent absence of dark matter in the dwarf galaxy NGC 1052-DF2 is consistent with the view that most observed dwarf galaxies are tidally disturbed satellites of their hosts.<sup>9</sup> As such, their velocity dispersions will appear much greater than warranted by their observed visible mass, thus mimicking dark matter. If this view is valid, NGC 1052-DF2 may be one of the few tidally undisturbed dwarf galaxies out there. This possibility is consistent with modified gravity theories that predict little or no deviation from Newtonian dynamics for such small galaxies.

At the same time, a viable modified gravity theory must do a lot more than account for galactic dynamics. Extremely precise measurements now exist for minute temperature deviations of the cosmic microwave background from its mean value.<sup>10</sup>

Computerized surveys that involve millions of galaxies offer statistics on the distribution of matter at various cosmic scales.<sup>11</sup> These data sets are modeled with exquisite accuracy in the context of the standard, so-called concordance model of cosmology.<sup>12</sup> In this collision, less cold dark matter plays a fundamental role governing structure growth. A modified gravity theory that does away with dark matter must offer a suitable alternate mechanism.

Constructing a modified theory of gravitation that satisfies all these constraints is incredibly challenging. It is easy to add terms to the Einstein–Hilbert Lagrangian to produce an alternative theory. But it is very difficult to construct an alternative theory that does not run afoul of all the observational data collected about galaxies and the universe since the days of Georges Lemaître and Edwin Hubble. Yet the continuing absence of any direct observation of dark matter provides a strong incentive for research in this direction.

A modified theory of gravitation may be able to tackle issues that challenge the concordance model. One such issue is the Hubble tension, the apparent discrepancy between values of the Hubble parameter, which measures cosmic expansion, derived from data obtained from our local neighborhood versus data derived from cosmological observations.<sup>13</sup> Many modified theories of gravitation yield a different relationship between these values, which may be more consistent with observation. It is, of course, entirely possible that the Hubble tension is rooted in another issue altogether: the notion that the universe is not only homogeneous and isotropic on average, but that the Milky Way region is representative of the average density. If it turns out that the Milky Way is situated in a large cosmic void, the Hubble tension may vanish. Other res-

olutions, such as the presence of systemic bias in certain families of measurement, have also been offered by many authors.

There is also the issue of the cosmological constant, originally proposed by Einstein to create a stationary solution of his field equations. It was reintroduced after the discovery from the observed luminosity-distance relationship of type Ia supernovae that the expansion rate of the universe is increasing.<sup>14</sup> But how is this constant interpreted? It has been known for more than a hundred years that it can be viewed as a medium with strong negative pressure, which in turn shows up as negative effective mass density in the Newtonian approximation of Einstein's theory. Could such a medium exist? An obvious candidate, with the right equation of state, would be the zero-point energy of the quantum fields of the Standard Model of particle physics. But its calculated value is dozens of orders of magnitude higher than the value deduced from the observed acceleration. It would count as a major success for any modified gravity theory if it could offer an alternative explanation for cosmic acceleration, resolving or eliminating this embarrassing discrepancy, known as the cosmological constant problem.<sup>15</sup>

Unambiguous direct detection of dark matter would settle the question decisively. Until such a detection is confirmed, the possibility remains open that the dynamics attributed to dark matter is, in fact, due to deviations from the predictions of Einstein's theory of relativity on the scale of galaxies and beyond. The search for a viable modified theory of gravity is a difficult challenge but it may not be futile.

## Viktor Toth

### Jean-Pierre Luminet replies:

I would like to thank Viktor Toth for his brilliant commentary. He brings additional theoretical and observational insights to the dark matter problem and, more specifically, to the modified theories of gravity that aimed at solving this problem without necessitating the existence of hypothetical particles of nonbaryonic matter. These particles, while predicted by various high energy physics theories, are yet to be detected experimentally.

Toth is right in pointing out that my paper essentially reduces the question of modified gravity to the primitive model—in both senses of the word—of Mordehai Milgrom. I can offer at least two justifications for what may appear to be an oversimplification of the problem.

The first is that my essay remains only a succinct summary of a vast problem that arises when analyzing the apparently abnormal movements of luminous matter. To explain these movements, I invoke either certain exotic forms of dark matter, or a modification of the laws of gravity, or—as I briefly mention at the end of my essay—a possible combination of the two. Much more development

would be needed to treat the full problem exhaustively, as evidenced by the abundant literature on the subject.

The second reason is a personal bias. As a result of my training as a physicist in general relativity, I have always felt somewhat reticent toward the MOND model proposed by Milgrom. As Toth rightly points out, this model is purely phenomenological and does not respect even basic conservation laws. The theory has certainly been improved through its relativistic generalization proposed by Jacob Bekenstein and other researchers. But as far as I know, there is still no convincing MOND theory, i.e., one capable of accounting for the dynamics of celestial objects at different scales in a universal model. The change in gravity that could explain the kinematics of stars within galaxies is not the same as that to explain the dynamics of galaxies within their clusters. MOND models also do not provide a detailed account for gravitational lenses. Former colleagues from the Paris Observatory have pointed out that it is quite possible to reach the same success as MOND through a new kind of matter, dipolar dark matter, retaining general relativity for the law of gravity.<sup>16</sup>

Toth's reply offers a remarkable panorama of the different theories of modified gravity. As such, his letter could have been the subject of a full-fledged essay. Among the theories he cites, the popularized version of Milgrom is undoubtedly the weakest.

As a relativistic physicist, I am obviously aware of the existence of the Jordan–Brans–Dicke tensor-scalar theory,<sup>17</sup> in which the gravitational interaction is transmitted by a scalar field as well as the tensor field of general relativity. The gravitational constant  $G$  is no longer assumed to be constant, and its inverse,  $1/G$ , is replaced by a field that can vary in space and time. Given the extremely small deviations from Albert Einstein's classical general relativity predicted by Jordan–Brans–Dicke's theory, it is generally considered that the two theories cannot be distinguished from each other by astronomical observations. Jordan–Brans–Dicke's theory thus represents a minority point of view in modern theoretical physics. For this reason I did not mention it in my short essay. The same goes for the so-called  $f(R)$  theories pointed out by Toth.

The recent advent of gravitational astronomy allows for the detection of gravitational waves resulting from the coalescence of compact objects such as black holes and neutron stars. This development also creates new possibilities for testing the different theories of gravity with more precision. Toth mentions the multispectral observations of GW170817—an event attributed to the coalescence of a pair of neutron stars that resulted in a double gravitational and electromagnetic signal. Measuring the arrival times of these two signals—which seem to coincide with a high degree of precision—strongly constrains not only the classical theories of gravity referred to as bimetric,<sup>18</sup> of which the MOND theory is a very partic-

ular case, but also some theories of quantum gravity, such as superstrings.<sup>19</sup> I recently discussed these constraints as part of a book on the various approaches in quantum gravity.<sup>20</sup> In a long chapter, I examine another essential problem mentioned by Toth: dark energy. Within the scientific community, dark energy is debated no less fiercely than dark matter. There is an enormous discrepancy between the experimental value of dark energy deduced from the acceleration of cosmic expansion and the theoretical value of quantum vacuum energy calculated within the framework of supersymmetry theory. I express the opinion, shared by some prestigious colleagues, that this discrepancy is a clear sign of the failure of the latter. This viewpoint is an alternative to the more common position, which is to dismiss the interpretation of dark energy in terms of the cosmological constant, originally introduced by Einstein for the wrong reasons. Georges Lemaître—the real founding father of modern relativistic cosmology<sup>21</sup>—was the first to associate the constant to the energy of the quantum vacuum in its fundamental state.

Toth also mentions a subject that has been hotly debated for the last few years known as the Hubble tension. Namely, the apparent discrepancy in measurements of the expansion rate of the universe as derived from local observations versus cosmological observations. Rather than posing a fundamental challenge to general relativity, which forms the basis for the current standard model of cosmology, I share Toth's opinion that the tension may arise from some simplifying assumptions associated with the current model. Toth mentions the hypothesis of homogeneity and isotropy. This has only been tested on a very large spatial scale, but is already known to be violated on the scale of galaxy clusters. There is good reason to think that this tension might arise artificially from the hypothesis of a strictly zero spatial curvature. The tension could be eliminated by a slightly positive space curvature, which would be perfectly compatible with the latest observations of the Planck telescope.<sup>22</sup> It would also have a better basis on a physico-mathematical level.<sup>23</sup>

What I particularly appreciate in this letter is that, while advocating theories of modified gravity to the detriment of nonbaryonic dark particle models, Toth admits it is difficult to obtain a viable modified gravity theory that is both mathematically coherent and experimentally verified. But this certainly does not render futile the research done in this subfield of fundamental physics, a body of work that my short essay may not have done justice to.

**Viktor Toth** is a part-time theoretical physicist and a Senior Research Fellow at the Department of Systems and Computer Engineering at Carleton University.

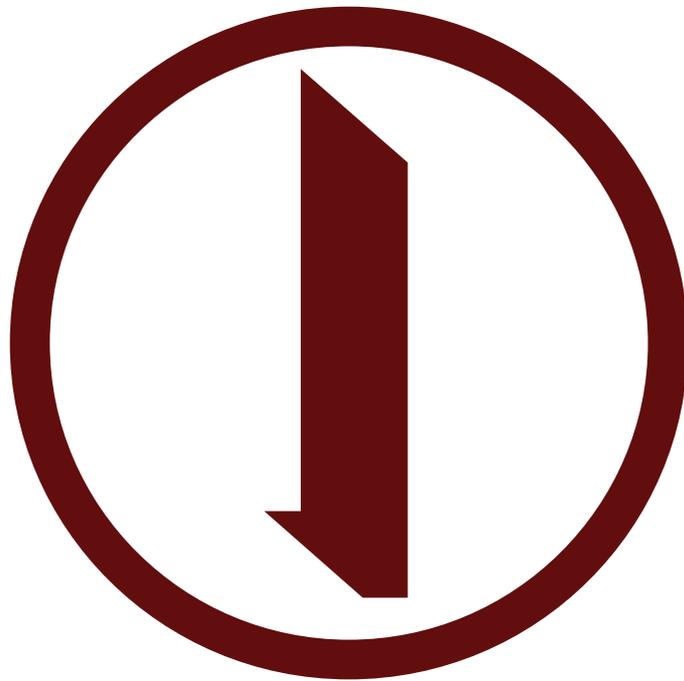
**Jean-Pierre Luminet** is Director of Research at the CNRS Astrophysics Laboratory in Marseille and the Paris Observatory.

- Mordehai Milgrom, "A Modification of the Newtonian Dynamics as a Possible Alternative to the Hidden Mass Hypothesis," *Astrophysical Journal* 270 (1983): 365–70, doi:10.1086/161130; Erik Verlinde, "Emergent Gravity and the Dark Universe," *SciPost Physics* 2, no. 016 (2017), doi:10.21468/SciPostPhys.2.3.016.
- Pascual Jordan, *Schwerkraft und Weltall* (Brunswick: Vieweg und Sohn, 1955); and Carl Brans and Robert Dicke, "Mach's Principle and a Relativistic Theory of Gravitation," *Physical Review* 124, no. 925 (1961), doi:10.1103/PhysRev.124.925.
- Salvatore Capozziello and Mariafelicia De Laurentis, "Extended Theories of Gravity," *Physics Reports* 509 (2011): 167, doi:10.1016/j.physrep.2011.09.003.
- Arthur Eddington, *The Mathematical Theory of Relativity* (Cambridge: Cambridge University Press, 1924).
- Clifford Martin Will, *Theory and Experiment in Gravitational Physics* (Cambridge: Cambridge University Press, 1993).
- Benjamin Abbott et al. (LIGO, Virgo, and other collaborations), "Multi-Messenger Observations of a Binary Neutron Star Merger," *The Astrophysical Journal Letters* 848, no. 2 (2017): L12, doi:10.3847/2041-8213/aa91c9.
- Nathan Rosen, "General Relativity and Flat Space, I & II," *Physical Review* 57, no. 2 (1940): 147–50, doi:10.1103/PhysRev.57.150.
- Jacob Bekenstein, "Relativistic Gravitation Theory for the Modified Newtonian Dynamics Paradigm," *Physical Review D* 70, no. 8 (2004): 083509, doi:10.1103/PhysRevD.70.083509.
- Pieter van Dokkum et al., "A Galaxy Lacking Dark Matter," *Nature* 555 (2018): 631, doi:10.1038/nature25767.
- Planck Collaboration, "Planck 2018 Results: VI. Cosmological Parameters," *Astronomy & Astrophysics* 641, no. A6 (2020), doi:10.1051/0004-6361/201833910.
- Romina Ahumada et al., "The Sixteenth Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra," *Astrophysical Journal Supplement Series* 249, no. 1 (2020): 3, doi:10.3847/1538-4365/ab929e.
- Steven Weinberg, *Cosmology* (Oxford: Oxford University Press, 2008).
- Eleonora Di Valentino et al., "Cosmology Intertwined II: The Hubble Constant Tension," *arXiv* (2020), arXiv:2008.11284.
- "The Nobel Prize in Physics 2011," *NobelPrize.org*, Nobel Media AB 2020.
- Ronald Adler, Brendan Casey, and Ovid Jacob, "Vacuum Catastrophe: An Elementary Exposition of the Cosmological Constant Problem," *American Journal of Physics* 63, no. 7 (1995): 620–26, doi:10.1119/1.17850.
- Luc Blanchet and Françoise Combes, "MOND with or without Dark Matter," *arXiv* (2009), arXiv:0910.5204.
- I briefly mention the theory in my popular essay *Black Holes* (Cambridge: Cambridge University Press, 1992).
- Sabine Hossenfelder, "A Bimetric Theory with Exchange Symmetry," *Physical Review D* 78, no. 4 (2008): 044015, doi:10.1103/PhysRevD.78.044015; Jean-Pierre Petit and Gilles d'Agostini, "Cosmological Bimetric Model with Interacting Positive and Negative Masses and Two Different

- Speeds of Light, in Agreement with the Observed Acceleration of the Universe,” *Modern Physics Letters A* 29, no. 34 (2014): 1450182, doi:10.1142/S021773231450182X.
19. Kris Pardo et al., “Limits on the Number of Spacetime Dimensions from GW170817,” *Journal of Cosmology and Astroparticle Physics* 7 (2018), doi:10.1088/1475-7516/2018/07/048.
  20. Jean-Pierre Luminet, *L'écume de l'espace-temps* (The Foam of Space-Time) (Paris: Odile Jacob, 2020).
  21. Jean-Pierre Luminet, *L'invention du big bang* (The Invention of the Big Bang) (Paris: Le Seuil, 2014).
  22. Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk, “Planck Evidence for a Closed Universe and a Possible Crisis for Cosmology,” *Nature Astronomy* 4 (2019): 196–203, doi:10.1038/s41550-019-0906-9.
  23. Jean-Pierre Luminet, *The Wraparound Universe* (New York: AK Peters, 2008), chapters 10 and 39.

DOI: 10.37282/991819.21.12





**INFERENCE-REVIEW.COM**

**@INFERENCEREVIEW**