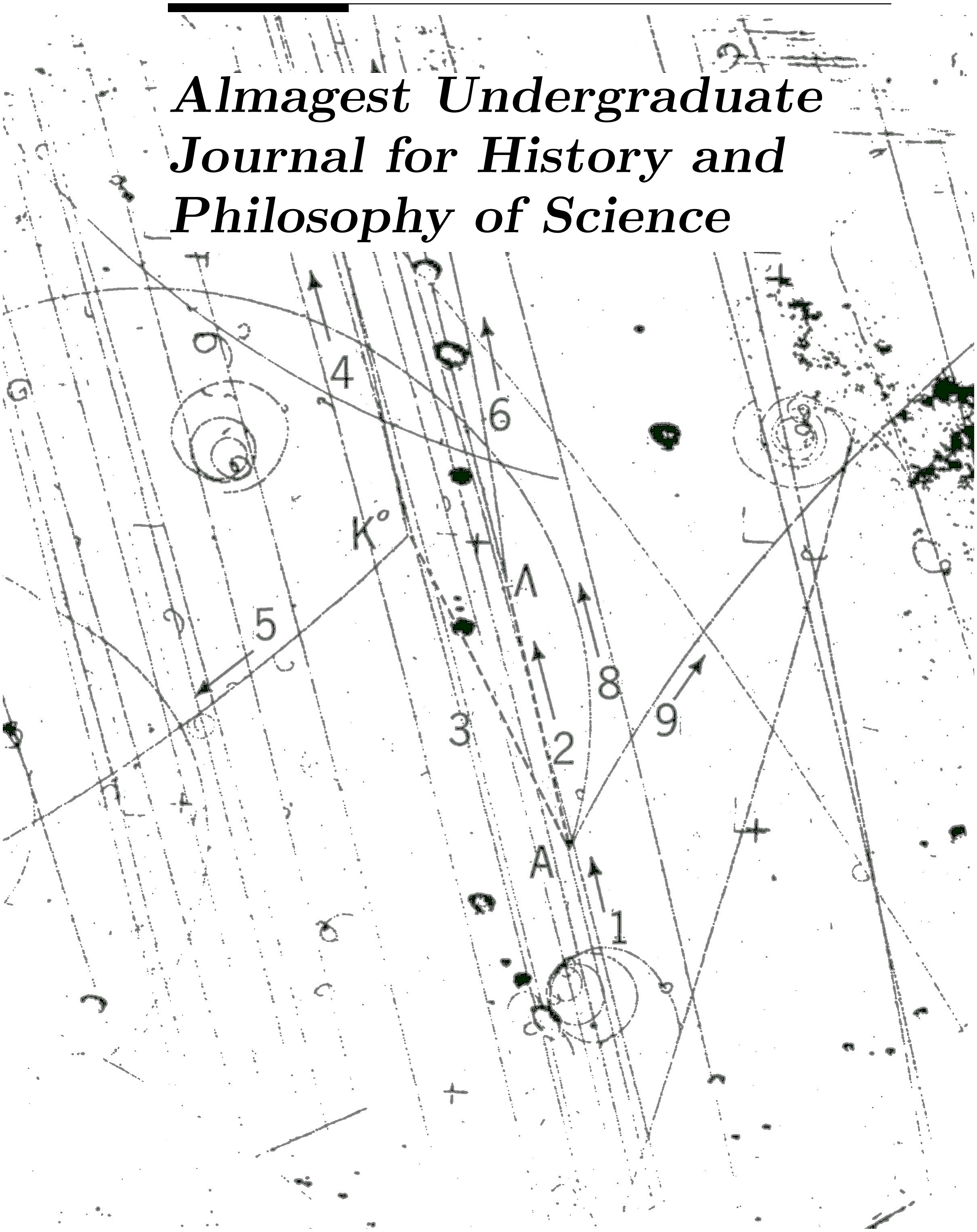
# ***Almagest Undergraduate Journal for History and Philosophy of Science***



Published by Almagest

www.almagestjournal.com

Cover photo taken from the U.S. National Archives, item identifier 22122886. Citation: Department of Energy. Lawrence Berkeley National Laboratory. Public Affairs Department. Strategic Resources Office. Photography Services. 2012.



*First printing, November 2019*

We are indebted to our sponsors, without whom this publication would not be possible. The creation of this issue was supported by the History and Philosophy of Science Undergraduate Society, The Institute for History and Philosophy of Science and Technology, Victoria College, St. Michael's College, and the Arts and Science Student Union. If you are interested in supporting future issues of *Almagest*, please contact us at **almagestjournal@gmail.com**.

# *Foreword*

Founded in 2014, the *Almagest Undergraduate Journal for History and Philosophy of Science* is a student-run, peer-reviewed academic journal which features high-quality original research from students. Created in close connection with the University of Toronto's HPS Undergraduate Society (HPSUS), and based at the Institute for History and Philosophy of Science and Technology (IHPST), *Almagest* is proud to have been able to present students in Toronto with a platform to refine and publish their own creative work. Recently, *Almagest* has undergone significant expansion, posting open-access copies of all past issues online, as well as extending beyond Toronto by opening the call for papers to other institutions world-wide. The work presented in these pages is the culmination of our editorial team's efforts over the past academic year.

In addition to student-written articles, we also feature interviews with leading academics. This year, these include theoretical physicist Professor Freeman Dyson, historian of science Professor Peter Galison, and philosopher of science Professor Ian Hacking, all of whom we graciously thank for their time and enthusiasm.

At *Almagest*, it is our belief that students and researchers need not be mutually exclusive categories, and that any motivated student is capable of producing original research if properly guided. The articles below accurately illustrate the originality and precision of thought which every undergraduate student is capable of. With this in mind, we hope you enjoy reading these articles and interviews, and if you are a student or early-career researcher in any area connected to the History and Philosophy of Science and Technology, we encourage you to submit your own work for consideration for publication in a future issue.

**Patrick Fraser**

Senior Editor and HPSUS President
November 2019

# *Editorial Team*

# *Contents*

# 1

# *Data Ombudsman for the New International Data Bank*

## *Surveillance Capitalism, Privacy, and Big Data*

**Michael Bervell**
*Harvard University*

## Introduction

"Read The Privacy Invaders and The Naked Society together during the same weekend," challenges Lewis Nichols in the last paragraph of his April 1964 New York Times book review, "and youll not dare go to bed without double-bolting the door, pulling down the shades, throwing the switch on the electric-light system, yanking out the telephone wires, pulling all the bedding including the mattress over your head and  not breathing until morning" [16]. In the year after writing his review, Nichols would see the "The Naked Society" by Vance Packard at the top the best-seller list for weeks [18]. In it, Vance argues that changes in technology encroach on privacy and could create a future society with radically different privacy standards [5]. He focuses his attention on one of the largest users of this new technology: the United States government via Lyndon B. Johnsons "great society" programs [17].

Just a year after the publishing of Vances book, President Lyndon Johnsons administration announced their plan to create a National Data Bank, which had a principal aim of helping economists, social scientists, and government specialists investigate major economic or social problems [7]. Even though a score of federal agencies collected data (e.g. the Census Bureau, the Internal Revenue Service, and the Social Security Administration), researchers lacked a way to connect all of these unlinked data points. A National Data Bank, Johnson argued, could help alleviate the publics burden of completing overlapping forms and the governments duplication of these forms across departments. However, Congress, the watchdog of the time, was not so easily sold.

Fearing that the establishment of the National Data Center as a statistical center would lead to abuses, the House Government Operations Committee appointed Cornelius Gallagher, a Congressman from New Jersey, to head the newly created Special Subcommittee on the Invasion of Privacy [7]. In July 1966, Gallaghers committee held three days of hearings with government officials, statisticians, and privacy experts of the day (including Vance Packard). These hearings instigating headlines in the New York Times such as "Data Bank: Peril or Aid?" and "The U.S. Central Data Bank: Would It Threaten Your Privacy?" [19]. Less than two years later, the movement to establish a national data bank was over. In March 1968, amidst public outcry, Congress was able to stop Johnson and his administration from creating a centralized database of information. Senator Edward V. Long stated the significance of the win best in 1968, claiming, "Our privacy today depends on decentralization of information on individuals" [19].

50 years after the original National Data Bank, there is a new International Data Bank. These International Data Banks are the huge databases of information collected, stored, and utilized by companies such as Apple, Google, and even by tech-savvy individuals. While the notion of a National Data Bank is not uniquely modern, the ways data banks operate today have completely shifted. Today, private entities can act as "governments" to pull together the scattered data points of individuals from around the world. In this paper, I hope to ask and answer one question: given the prevalence of surveillance capitalism in the modern, big data era, what should be the "congress" or ombudsman that protects common consumer interests? I hope to shed light on this question through analyzing cases of big data and surveillance capitalism alongside interviews I conducted with technology company executives in December 2018. Ultimately, through analyzing specific case studies I hope to describe a potential solution for creating a corporate ombudsman that will ensure privacy and responsibility in the era of big data and surveillance capitalism.

## Terminology and Theory

To understand what data ombudsman should reign in the era of big data and surveillance capitalism, it is important to first define the three key terms assumed throughout this paper: data ombudsman, big data, and surveillance capitalism. Setting out these assumptions will not only contribute positively to the argument, but also will eliminate any ambiguity about these often-contested meanings.

## Big Data

In their paper "Big Datas End Run around Anonymity and Consent," philosophers Solon Barocas and Helen Nissenbaum argue that big data is best understood as a paradigm instead of than a particular technology, method, or practice [1]. Big data, they argue, is a way of thinking about knowledge through data. Fundamentally it is a belief in the power of finely observed patterns, structures, and models drawn inductively from massive data sets. Rob Kitchin builds off this argument in this 2014 paper, writing that big data is not necessarily notable because of its size, but instead because of the continuous generation of dynamic flows of diverse, fine-grained, relational data [12, p. 2]. This real-time observation and continuous collection is what makes the big data of today so different from the big data described in the 1968 National Data Bank. With this new understanding of big data also come new considerations, such as surveillance capitalism.

## Surveillance Capitalism

The era of surveillance capitalism differs from the era of industrial capitalism in a variety of ways. As Shoshanna Zuboff argues, unlike industrial firms of the mid-twentieth century who made their money providing goods and services, contemporary firms profit off of surveillance [25]. These corporations "solicit social behaviors, monitor those behaviors, map social interactions, and resell what they learn to advertisers and others" [24]. The most notable aspect of surveillance capitalism, however, is that users are generally unaware of this hidden economy. The interfaces of firms, such as Google and Facebook, invite their users to openly share personal, real-time information in exchange for the value gained by using the service. However, unlike the era of industrial capitalism where such exchanges would have been managed by a contract, today, user agreements, the modern-day contract substitute, are rarely read and thus are obsolete.

This theoretical claim seems to also be supported by Dennis Crowley (the founder of Foursquare and Dodgeball)[1] and Ryan Graves (the former CEO

[1] Foursquare is a local search-and-discovery service mobile app which provides search

of Uber[2]) both of whom described problems with the current dialogue between society, technology, and privacy. They claimed, respectively, that "I do think most TOS' are very very very dense – too dense for the average user to read/comprehend" (Crowley) and "I don't think users will ever spend the time to read a ToS that also meets the requirements of the attorneys" [4].

Zuboff continues by explaining that surveillance capitalism is also a black box, fueled by big data collected through digital systems. Thus, the concept of surveillance capitalism can be viewed as the careful, deliberate shaping of environments, both digitally and non-digitally, to subtly influence individuals behavior to generate monetizable data points or predictions based on the results of these continuous experiments [23].

## Data Ombudsman

In Swedish, the word "ombudsman" is used to describe an official who stands up for the common man in order hold the powerful are held accountable for their actions. For instance, the United States the Food and Drug Administration (FDA) plays a key role in protecting and advancing the public health [9]. When individuals have health concerns about a restaurant or about the medications they received from their doctor they turn to the FDA, the ombudsman, who can make change. Given this, who should the common folk turn to when they have problems with their electronic data concerns? Since there is no "FDA for data," citizens are left with no ombudsman. Conceivably, we can imagine citizens filing class action lawsuits against data-insensitive entities; however, today, no such standardized system exists for all citizens of the United States.

Obviously, the FDA protects health, thus extending citizens lives and creating a better nation to live in; but why is a "data ombudsman" necessary? Advocates of data ombudsman argue that whether in the form of regulation,

results for its users. The app provides personalized recommendations of places to go to near a user's current location based on users' "previous browsing history, purchases, or check-in history". In reality, this app is one that expertly sells the location data of individuals to other corporations at a profit. They have 60 million users every month. (Summary courtesy of Wikipedia). Dodgeball was a location-based social networking software provider for mobile devices. Users texted their location to the service, which then notified them of crushes, friends, friends' friends and interesting venues nearby. Google acquired Dodgeball in 2005 and discontinued it in 2009, replacing it with Google Latitude.

[2] Uber Technologies Inc. (doing business as Uber) is a peer-to-peer ride-sharing, taxi cab, food delivery, bicycle-sharing, and transportation network company (TNC) headquartered in San Francisco, California, with operations in 785 metropolitan areas worldwide. Its platforms can be accessed via its websites and mobile apps. Uber has been prominent in the sharing economy, so much so that the changes in industries as a result of it have been referred to as Uberisation.

intermediaries, or processes, a data ombudsman would create a more just and fair nation to live in [6]. By enforcing privacy, ombudsmen help individuals maintain autonomy, a fundamental human right. As John Stuart Mill advocates in *On Liberty*, "it really is of importance, not only what men do, but also what manner of men they are that do it" [15]. Even if an individual has no control over anything else, they have a right to have control over themselves. Without privacy, those who are the most powerful will likely manipulate and control citizens for their own gain. For instance, it is easy to imagine scenarios where, without privacy, insurance companies can use large data sets as a means for discrimination [22]. In a world with a data ombudsman, citizens are free to choose how to live their life while being liberated from inappropriate corporate intrusion. Thus, a "data ombudsman" is needed to not only protect consumer data, but also to protect more fundamental rights such as privacy, autonomy, and freedom of choice.

## Data Ombudsman in Practice

In his 1986 PhD dissertation, Colin John Bennett, now a professor of political science at the University of Victoria, proposed three potential models for creating a "data ombudsman." He labels them (1) the self-control model, (2) the subject-control model, and (3) the institutional control model.

Under the self-control model "fair information practice is implemented internally without intervention by external bodies" [2, p. 199]. In such a scenario, the record-keeper is expected to take a position of "self-surveillance," protecting consumers against abuse, providing rights of access, and disclosing the use of personal data. This is the current practice that most data-centric companies in America adopt. Crowley, drawing from his personal experience, states that "the users trust the services with save-guarding their privacy, and it's the services responsibility to honor the trust the users put in them" [3].

Second, Bennett describes the subject-control model of data ombudsman. This form of control requires that those directly affected should actively protect their data [2, p. 200]. Under this "self-help solution" an individual has the ability to control the extent to which corporations access her personal data, and thus she can protect her own privacy. Bennett argues that such a system creates a "right to know" which would discourage misuse and abuse of personal data by corporations. Data subjects would have the right to know about the existence of data banks, information about what is in these data banks, and have privileges to correct or erase data from these data banks.

Finally, Bennett describes the institutional control model, which is dependent upon an established regulatory institution that can both regulate the record-keeper and assist data subjects in the pursuance of their rights [2, p. 203]. Bennett describes three potential approaches: the licensing approach (a strict process to both register and regulate data brokers), the registration approach (an authority that requires data brokers to register activity), and the data commissioner approach (a third-party intermediary that handles all data). Based on his experiences as the CEO of Uber, Ryan Graves, describes a potential benefit of the institutional control model: "I'm all for capitalism but I don't think corporations have a great history of making the right decisions in these types of cases" [4].

While each model of data ombudsman has its benefits and drawbacks, the remainder of this paper will focus only on discussing the subject-control model, which has widely been accepted as the conventional model for the international data ombudsmen of the 21st century.

## Data Ombudsman as the Subject-Control Model

### Optional Ommatidia

In his 2017 feature film, *Dragonfly Eyes*, Xu Bing, a Chinese artist, throws a spotlight on this phenomenon of super surveillance, surveillance realism, and surveillance capitalism. His piece tells an 81-minute dystopian love story through using only publically available footage from surveillance cameras in China that stream 24/7. Over the course of 2 years, his team collected this data by using "over 20 computers to download footage, 24 hours a day" [13]. The consistency of this footage from Chinese streaming sites, dashboard cameras, CCTV cameras, and other publically available sources were fueled by the rise of big data. Bing then edited this footage to create a narrative. The result is a "fiction" story told entirely through raw footage of actual people, not actors.

This voyeuristic and omniscient style of storytelling likens Chinas surveillance cameras to a dragonflys ommatidia and the directors team to the dragonflys mosaic-creating brain. More than that, *Dragonfly Eyes* provides a springboard to discuss issues of how surveillance has changed from a form of oppression to a form of capitalism in the big data era. Most recently, China has invested heavily in growing their network of closed-circuit television (CCTV) cameras with goals of having over 400 million by 2020 [20]. Cities in China are poised to become living dragonflies; however, unlike the traditional dragonfly, the "eyes" in these cities point in all directions, at all times, and at all people [14]. Bings film, *Dragonfly Eyes*, is unique because it could only be created in todays world of big data. For instance, even though Bing originally had the idea in 2013, he could not gather enough publically available surveillance footage to make the film work until 2015. Because of the phenomenon of Big Data, and specifically big public data, *Dragonfly Eyes* came to life.

Bing and his team did not own a camera or hire any actors. Instead, by using publically available big data he and his team were able to circumvent the problems of traditional filmmakers. Moreover, unlike the subjective cameras of traditional filmmakers, which are constantly moved or turned off and on, the surveillance cameras gaze does not judge. The surveillance camera simply records as much information as possible in a given time-frame and dispenses that to any viewer. Unlike previous films of the surveillance genre, *Dragonfly Eyes* is notable because of its form, in addition to its story. The form of the film is built upon using the images of individuals and the individuals themselves as objects. Each person in the film is an involuntary "actor"; they are a means (a character) to creating an end (the story). These unsuspecting "actors" do not have the ability to avoid this objectification. Thus adding to the films commentary on surveillance capitalism.

Luckily, however, Xu Bing subscribes to the subject-control model of data ombudsman and offers his data subjects a way out. During the production phase of the film, Bing and his team sought out the individuals used as actors in order to receive their permission [21]. Using their specific GPS positions, Bing and his team were able to find the subjects they had been filming [11]. As a result of his work, which is one of the first and only films to truly capture the totality of surveillance capitalism, Bing demonstrated that the data ombudsman as the subject-control model is actually feasible. However, this model also comes with its own set of problems. How can subjects control their data if they dont even have a relationship with the data-collector?

## Shadows as Subjects

To answer this question, lets consider the example of Facebooks "Shadow Profiles." In April 2018, Mark Zuckerberg, the founder and CEO of Facebook, appeared before the United States congress. During this hearing, Congressman Ben Lujan asked Zuckerberg about data that Facebook collects on non-users, or shadow profiles [10]. These shadow profiles are collections of data that Facebook collects on people who have not signed up for Facebook, never signed a consent agreement, and certainly have not accepted a privacy policy. To this, Zuckerberg responded that after making a Facebook account "anyone can turn off and opt out of any data collection for ads, whether they use our services or not" [10].

Here, Zuckerberg seems to run into a paradox. Like Xu Bing, Zuckerberg is a proponent of the subject-control model of data ombudsman; however, it seems that Facebook as a platform has not implemented this model effectively. He argues that Facebook users should not only be able to view their data, but they should also be able to manipulate, correct, or delete it. However, when Facebook collects information about non-users they lack the option to be their own data ombudsman.

The problems continue: philosophers Solon Barocas and Helen Nissenbaum argue that even if non-users are able to opt-out Facebook, they will be objectified by the "tyranny of the minority" [1, p. 60]. In the world of big data, the volunteered information of the few can unlock the same information about the many. For instance, "friending" someone on a social networking site can "signal that this is a person with whom we share certain interests, affinities, and history. In associating with this person, we open ourselves up to inferences that peg us as people who share certain qualities with this other person" [1, p. 61]. Whether you are a platform user or not, the subject-control model of data ombudsman is not able to adapt to the problem of tangential networks. This is true both in the case of shadow profiles and in the tyranny of the minority. For these reasons, in comparing the 1968 National Data Bank with the currently muddied waters of 2018 corporate International Data Banks, it seems that the best ombudsman may be Bennetts third option: regulation.

## Conclusion

On March 6, 1967, almost exactly a year before the National Data Bank officially ended, U.S. Senator Sam J. Ervin Jr. gave a speech to American Management Association in New York City. "Thank you for inviting me to share your meeting to discuss computers and privacy with you today," he began [8]. Throughout his speech he continued to warn about the dangers of privacy that are brought to the forefront in the era of new technology. He concludes that it is the responsibility of Congress as "the ultimate guardian of liberties of the people" to ensure that malicious intent does not poison a benign technology [8]. Today, half a century after these remarks were made, its sentiment still reigns true. Despite the rise of big data and the prevalence of the new surveillance capitalism economy, Earvin was correct in claiming that the machine alone does not invade privacy. Instead, privacy is destroyed as a result of man.

It is worth returning, then, to the question I asked in opening this paper: given the prevalence of surveillance capitalism in the modern, big data era, what should be the data ombudsman that protects common consumer interests? Who will keep man in check?

Through using the framework of Colin John Bennett, we are presented with three potential options. First, the data ombudsman could follow a model of self-control, one that permeates the current world of data brokering. Though in reality, the incentives of data protection and the protection of autonomy do not align. Perhaps then the data ombudsman should follow a model of subject-control. While this works in certain scenarios, it is not realistic in the data-heavy world where we interact with data collectors that we may not even know. Finally then, we are left with Bennetts third option  the option that has been implemented, the data ombudsman as an institutional control model. It is too early to tell if this option will overcome the problems presented before it; however, of the three this "return to order" is the most adaptable to the current world of the International Data Bank.

Yes, the government-run National Data Bank of 1968 may have ended due to congress, the ombudsman of the time, but in 2018 the new problems of the corporate-run International Data Banks are alive and in need of a data ombudsman.

## Bibliography


[1] Solon Barocas and Helen Nissenbaum. *Big Data's End Run around Anonymity and Consent*. Cambridge Books Online, 2014.

[2] Colin Bennett. *Regulating the computer: A comparative study of personal data protection policy*. University of Illinois at Urbana Champaign, 1986.

[3] Michael Bervell. Personal interview with dennis crowley. Email, December 2018.

[4] Michael Bervell. Personal interview with ryan graves. Email, December 2018.

[5] Joseph Bishop. Book review: The naked society. *The Yale Law Journal*, 74(194–199), 1964.

[6] Amy Boyer. Why is privacy important.

[7] Luther J. Carter. National data bank: Its advocates try to erase "big brother" image. *Science*, 163(3863):160–162, 1969.

[8] Sam Earvin. The computer and individual privacy, 1967.

[9] U.S. Food and Drug Administration. What we do, 2019.

[10] Taylor Hatmaker. Zuckerberg denies knowledge of facebook shadow profiles. *Tech Crunch*, 2018.

[11] Anthony Huang. Fruits of surveillance: Xu bing interview about 'dragonfly eyes'. *Musée Vanguard of Photography Culture*, 2018.

[12] Rob Kitchin. Big data, new epistemologies and paradigm shifts. *Sage Journals*, 2014.

[13] Edmund Lee. How chinese artist xu bing made an entire film out of surveillance camera footage. *South China Morning Post*, 2018.

[14] Joyce Liu. In your face: China's all-seeing state.

[15] John Stuart Mill. *On Liberty*. Longman, Roberts & Green, 1869.

[16] Lewis Nichols. Still naked. *The New York Times*, 1964.

[17] Margaret O'Mara. The end of privacy began in the 1960's. *The New York Times*, 2018.

[18] Philip Quarles. Beyond 'eggheads': Vance packard pulls back the curtain on advertising, 1958.

[19] Nan Robertson. Data bank: Peril or aid? the u.s. central data bank: Would it threaten your privacy? *The New York Times*, 1968.

[20] Jon Russell. China's cctv surveillance network took just 7 minutes to capture bbc reporter. *Tech Crunch*, 2017.

[21] Dan Schindel. Cobbling a movie together from surveillance footage. *Hyperallergenic*, 2017.

[22] David Schultz. It's legal for some insurers to discriminate based on genes.

[23] Richard H. Thaler and Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, 2008.

[24] Fred Turner. The arts at facebook: An aesthetic infrastructure for surveillance capitalism. *Poetics*, 67:53–62, 2018.

[25] Shoshanna Zuboff. Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30:75–89, 2015.

# 2

# *Can Teleology Explain Biological Events? D.M. Walsh and Purposiveness*

**Sherry Ann Morgenstern**
*Yale University*

## Introduction

Teleological purposes help us understand what biological entities are. For example, the purpose of canine teeth is to grip and tear food. The purpose of teeth helps us understand what they are. . We intuitively distinguish things we encounter based on the thing's apparent end goal. We conceptualize things as whole based on their apparent telos.[1] However, it is less clear whether purposes can also explain. 'Purposiveness' in this paper refers to the final end or telos of an object [4].[2] Since Kant's time, mechanism has been the more popular tool for providing explanations, especially in science [5]. Mechanistic explanations explain in terms of parts. For example, canine teeth are four different kinds of tissues formed in a pointed way and attached to the upper jaw. However, D.M. Walsh argues that purposiveness explains biological events alongside mechanism. His paper "Organisms as Natural Purposes: The Contemporary Evolutionary Perspective" argues for this thesis using the difference-making model of explanation. This paper will consider whether purposiveness is a useful explanatory tool alongside mechanistic explanations by considering possible challenges to Walsh's paper.

I consider two challenges to Walsh's argument and show that Walsh's account

[1] For other intuitive understandings of teleology, especially teleology as a value-based phenomenon, see [1]

[2] Kant gives several definitions of 'purposiveness' throughout the *Critique*, and the definition given here is derived from Walsh's application the concept. For further clarification on the idea of purposiveness, see [2].

of 'purposeness' can overcome them and does ultimately have the capacity to explain biological phenomena. I show that Walsh's understanding of explanation rests on a premise, the difference-making model of explanation, which weakens his argument. By considering an alternative model of explanation, the deductive-nomological view, we can consider a further critique of his thesis. However, I argue that Walsh's argument ultimately shows that purposiveness does have the capacity to explain biological phenomena.

In order to show Walsh's argument is successful, I first outline Walsh's case that purposiveness explains the predictability of biological events where mechanism cannot. Then, I recount Walsh's conception of the difference-making model, which offers a theory of what constitutes a 'good explanation'. I discuss a potential challenge to the view but argue that Walsh's account withstands it. In the next section, I compare the difference-making model to the deductive-nomological model, and conclude that the difference-making model lacks the robustness of exceptionless laws found in the deductive-nomological model. Nonetheless, this shortcoming does not compromise Walsh's primary aim to show that purposes explain biological events.

## The Stakes: Walsh's Argument for Purposiveness's Explanatory Power

Walsh contends that purposiveness has the capacity to explain biological phenomena. Further discussion will mainly focus on the first premise of the argument, but for context it will be helpful to recount the argument in full:

### Walsh's Argument for Explanatory Purposiveness

1. To explain a phenomenon is to single out those conditions that make the difference between the phenomenon's occurrence and nonoccurrence [5, p. 785].
2. Plasticity is the adaptive capacity of organisms to regulate their subsystems to build and maintain a stable organism [5, p. 778].
3. Plasticity is the condition that makes the difference between biological regularity's occurrence and nonoccurrence.
4. Plasticity explains the regularity of biological events.

5. Plasticity is biological purposiveness.
    (a) If plasticity is adaptiveness, and adaptiveness is purposiveness, then plasticity is biological purposiveness.
    (b) Plasticity is adaptiveness; it is a biological phenomenon.
    (c) Adaptiveness (alongside self-organization and self-regulation) defines purposiveness [6, p. 780]; it is a biological phenomenon [6, p. 783].
6. Purposiveness explains the regularity of biological events.
7. If purposiveness explains the regularity of biological events, then purposiveness has the capacity to explain.
8. C: Purposiveness has the capacity to explain.[3]

Walsh's account is appealing because it gives meaning to our conceptions of organisms as wholes. That is, the account validates the intuition that organisms are more than miniscule parts we only call a 'whole' for convenience. It does this by showing that organisms as units explain something meaningful in the natural world: Regularity. In the pure mechanistic account, these conceptions of organisms as meaningful units are often portrayed as heuristics at best. For example, it is intuitive to think that the purpose of a bird's wings is to enable it to fly. However, a mechanistic account would suggest that the purpose of a bird's wings is to balance other biological systems within the birds. Nonetheless, Walsh's argument relies on a contentious first premise. That is, to explain a phenomenon is to single out those conditions that make the difference between the phenomenon's occurrence and nonoccurrence. The next two sections will discuss the plausibility of this premise with reference to the 'difference-making' view of explanation. First, I will show that difference-making may not support Walsh's case - it seems possible that it supports a conclusion opposite to his own. The second section will consider whether Walsh's favored mode of explanation is the best kind available in science. I will argue that Walsh's account survives both attacks, though each objection weakens Walsh's argument.

[3] The second part of this argument is that mechanism does not adequately explain biological regularity. Although that argument is not strictly relevant to this paper, it can be summarized the following way: 1. If a thing cannot be explained by the general nature of matter, then it cannot be explained mechanically [6, p. 785]; 2. Biological regularity cannot be explained by the general nature of matter; 3. Biological regularity cannot be explained mechanically.

## The Difference-Making Model

The difference-making model further suggests that to explain a phenomenon is to single out those conditions that make the difference between the phenomenon's occurrence and nonoccurrence. The model is one of many different kinds of scientific explanation. Before examining the leading counterproposal to the difference-making model, the deductive-nomological view, it will be helpful to examine Walsh's summary of how the model works. His defense relies on a concept called 'invariance':

### Walsh's Difference-Making Model[4]

1. If one can answer a question such as 'what if things had been different?' ($Q$) of phenomenon ($P$), then one can explain $P$.
2. If the invariance relation holds, then one can answer questions like $Q$.
3. The relationship between P and its difference maker ($D$) is usually invariant.
4. Invariance is a relation between a range of properties $D$ might possess, called ($d_n$), and a range of properties P might possess, called ($p_n$).
5. $d_n$ and $p_n$ are invariant iff the relation is change relating and robust.
6. 'Change relating' refers to a difference in the value of $d_n$ that brings about a difference in the value of $p_n$.
7. 'Robust' refers to the insensitivity of $d_n$ and $p_n$ to a range of other factors extrinsic to their relation.
8. The invariance relation holds.
9. One can answer questions like $Q$.
10. One can explain $P$ [5, p. 785].

Walsh does not provide cases that show when the invariance relation holds

[4]This heading was written for ease of reference. However, Walsh did not invent the difference-making model and cites James Woodward in his summary of it. This outline restructures Walshs summary.

and does not hold. [5] Nonetheless, the difference-making model appears to work well in situations that call for collective explanation. For example, Walsh asks us to imagine a room of people who all read at a third grade level. We may want to know why each person individually reads at that level, in which case we would examine their individual lives and, using the invariance model, name the factor that makes the difference between her reading at a third grade level rather than another level. Alternatively:

> "We may want to know why the constitution of the classroom is such that all the children read at that level. Here we advert to those features that make the difference between this class having this particular composition rather than some other. In [the] example, the class is constituted this way because there is selection for reading ability; only those students who can read at the grade three level are allowed into the class" [5, p. 786].

The example implies that an outside actor – perhaps a teacher - determines the selection. Walsh seems to think that plasticity is the analogue to this teacher. The adaptiveness of organisms is the outside actor or phenomenon that explains the regular occurrence of sub-structural traits.

However, the invariance model may not support Walsh's case for collective explanation. It seems possible, at least at first, to explain regularity through individuals using the invariance model. The question "What is the feature of this classroom that makes the difference between (1) all students in it reading at a third grade level and (2) some other kind of classroom composition?" might set us up for a misleading answer. It may be that no feature of the classroom 'selects' for reading ability. Rather, there could be infinitely many reasons that lead students to this particular room, and we just label the room as a 'third grade reading classroom' for ease of reference and understanding. Instead of asking about characteristics of the classroom, we could wonder, "What about being a third grade reader makes one more likely to end up in this classroom than another?" We could answer this question with respect to the classroom, but more usefully, the opponent might argue, we could answer it with respect to its component parts.

We might imagine that each third grade reader possesses a gene that causes her to enter classrooms until she finds one with books labeled 'Grade 3' on the shelf. Assume other students possess similar genes for their own reading levels, so they bypass the third grade classroom. The shelf does not 'select'

[5] It is therefore difficult to know when the model should be applied and when it should be abandoned.The ambiguity is a potential problem as Walsh describes the relation between ($P$) and ($D$) as 'typically' invariant as opposed to 'always' invariant.

the students, but instead the students' select' the shelf. The resulting composition has the appearance of classroom selection, but in fact the regularity is explained by a component of the individuals in the scenario. If we can name those specific qualities that each individual student in the classroom shares, we can predict what kind of student will end up in the classroom. This explanation does not require any reference to the classroom itself or its selection capacity.

Walsh would have a sufficient answer to this objection. He agrees this individualistic account does offer a real explanation for the student's presence in the classroom but that this is not a full explanation. It may be that students in the class share traits beyond their reading ability, like the book-seeking trait, and that those traits predict that the students will end up in this particular classroom. However, Walsh would say that it could also be that none of the students share traits beyond those that the classroom appears to be meant for. Imagine the classroom is meant for all eight-year-old children in a particular county who read at a third grade level. Even though the children may also share the book-seeking trait in addition to those three criteria, they would have still ended up in this particular classroom had they not shared the book-seeking trait. For organisms, this result would be due to their 'adaptiveness' (plasticity). Although book-seeking really does explain why some of the children ended up in the classroom, it does not explain why all of them did with regularity. Another mechanism would have led them to the third grade readers room as long as they met the requirement for age, location, and ability. Walsh could defend this claim empirically by placing an eight-year-old who read at a third grade level and did not have the book-seeking gene in the county. Alternatively, he could defend the claim hypothetically by appealing to the many cases in which such regularity occurs despite different starting mechanisms.

Additionally, one can explain more about the third grade classroom by using the collective explanation rather than the individualistic one. Once one determines that the classroom selects for reading ability, one can ask a further question: why does the classroom select for this trait? The answer will depend on the situation. Whatever the answer is, it will provide more complete information than the individualistic explanation. The classroom might select for reading ability because it is the place where third grade reading will take place because it is the environment best suited to third grade readers, or because the other nearby environments are hostile to third grade readers. On the individualistic picture, the question 'why does the book-seeking gene manifest in third graders?' cannot appeal to a bigger picture in its explanation, but must appeal to a smaller one. Third graders have this gene due to smaller units that make up the gene. The individualistic picture is not inaccurate, Walsh says, but it is incomplete.

Of course, the quality of explanations resulting from the difference-making model depend on what one sets as the counterfactual. The question 'why is the classroom composed of third grade readers rather than fourth grade readers?' will yield a different explanation than the question 'why is the classroom composed of third grade readers rather than Polish farmers?' Walsh's model can technically answer both questions, but the answers may not be useful and the questions may not be worth asking. I believe this aspect of the view is a disadvantage insofar as explanations depend on the legitimacy of the question and our ability to accurately answer the counterfactual. The difference-making model offers a way to combine collective explanation and individualistic explanation to form a fuller picture of a given state of affairs. Although it relies on our ability to accurately predict the outcome of counterfactuals, the resulting outcomes offer a more comprehensive range of explanation. Now, I will turn to another mode of explanation that more directly challenges the first premise of Walsh's argument for biological purposiveness.

## The Deductive-Nomological Model

I have explained Walsh's difference-making model and argued it offers a more comprehensive level of explanation, even if it at times places too much trust in our individual predictive powers. The deductive-nomological ($DN$) model offers a further challenge of the first premise of Walsh's difference-making model. The first premise of that argument states, "If one can answer a question such as 'What if things had been different?' ($Q$) of phenomenon ($P$), then one can explain ($P$)." According to Carl G. Hempel, a scientific explanation makes deductions from exceptionless laws. Two major components make up an explanation: An 'explanandum' which is the phenomenon to be explained, and an 'explanans' which accounts for the phenomenon. While Walsh gestures at what constitutes a prediction through examples, Hempel names what constitutes a prediction explicitly:

> "Prediction in empirical science consists in deriving a statement about a certain future event (for example, the relative position of the planets to the sun, at a future date) from (1) statements describing certain known (past or present) conditions (for example, the positions and momenta of the plants at a past or present moment), and (2) suitable general laws (for example, the laws of celestial mechanics)" [3, p. 13].

Hempel concludes that the logical structure of a scientific prediction is similar to a scientific explanation. Prediction and explanation both "[involve] reference to universal empirical hypotheses" [3, p. 38]. A universal hypothesis is "a regularity of the following type: In every case where an event of a specified kind ($C$) occurs at a certain place and time, an event of a specified kind ($E$) will occur at a place and time which is related in a specific manner to the place and time of the occurrence of the first event" [3, p. 35].

The invariance model's notions of 'change-relating' and 'robustness' imply a cause and effect relation similar to Hempel's universal hypothesis. Nonetheless, Hempel's specifications for what constitutes a prediction are absent from Walsh's summary. The opponent could argue that these specifications hinder Walsh's notion of difference-making, since they demand knowledge of past conditions and general laws. Walsh asserts that even if the mechanics of a given organism had differed from what they were, the organism would have still achieved regularity within a population. That is, we could predict its biological behaviors. However, the laws describing biological regularity, if they exist, are far more nebulous than the celestial laws Hempel refers to. Additionally, we often lack data on the past conditions of organisms. It is difficult to say, at least with the kind of certainty that Hempel demands, what would have happened if a member of *chlamyphorus truncatus*[6] lacked gene $X$ or replicator $Y$.

Walsh somewhat successfully responds to this critique by denying the need to reference a governing law in his analysis:

> "Self-organizing systems offer an example of systematic regularity of phenomena that is not attributable to the unity of governing law. The robustness of self-organizing systems enables them to produce and maintain a trajectory toward a particular endpoint despite significant differences in initial conditions and perturbations in the process. It is not the unity of law, or of underlying causes, that determines the systematic regularity of self-organizing systems; it is something else, a structural property of the system as a whole"'[5, p. 785].

Walsh believes that governing laws 'peel apart' in their attempt to explain biological regularity [5]. Walsh is right that the initial conditions of each biological organism's move towards regularity can differ which Hempel's understanding of a general law cannot account for [3, p. 42].[7] However, Walsh and Hempel's disagreement might be resolved if the differing initial conditions in each organism shared a common trait. If this were true, regularity might

[6] Pink fairy armadillo.

[7] This conclusion is implied in Hempel more than explicitly stated. He describes initial conditions as components of general laws. If initial conditions $A$ differed from initial conditions $B$, then there would be two correspondingly different general laws $A$ and $B$.

still be explainable by general laws. For instance, suppose that each organism starts with a different number of total genes. Despite their different starting points, the organisms' subsystems emerged in a highly regular way. One might still derive a general law from these differing initial conditions by identifying a sameness about the differing number of genes: For instance, noticing that they all fell in a range of 30,000-90,000 total genes.

This solution has potential so long as the initial conditions did not vary too widely. However, even if the differences in initial conditions were small, the solution likely does not meet the rigor that Hempel seeks to establish in his *DN* model. The idea may be closer to what Hempel calls an 'explanation sketch' which consists of "a more or less vague indication of the laws and initial conditions considered as relevant, and it needs 'filling out' in order to turn into a full-fledged explanation" [3, p. 42]. The 'filling out,' Hempel notes, requires empirical research. Such research would be difficult to conduct on organisms with wildly varying initial conditions. In addition, because the initial conditions are so variable, it is likely that the resulting law would result in a 'vague indication' regardless of the amount of research conducted.

Despite the difference-making view's lack of comparative rigor, the view retains the capacity to explain the regularity of biological phenomena. An 'explanation sketch' is all that is needed  at least for now  to show that an organism can exert causal power as a single entity. Although more work may need to be done to explain the function of plasticity, the difference-making view can nonetheless account for regularity in spite of differing initial conditions. It does this by abandoning the *DN*'s requirement for uniform initial conditions altogether, or by accepting the kind of compromise I discussed above. Either solution comes with consequences for the difference-making view: If it abandons the *DN* requirements, it settles for an 'explanation sketch.' A sketch may be sufficient for showing that there is biological purposiveness, but it is not the most systematic kind of explanation available. If it accepts the compromise above, it must accept the disadvantages of that compromise too.

## Conclusion

In this paper, I have argued that Walsh's argument for purposive explanation mostly succeeds. In the first section, I recounted the argument for biological purposiveness. Then, I examined his defense of its first premise: That explanation consists in identifying a difference maker. I concluded that the *DN* model offers a more robust account of prediction than the difference-making view, but that the difference-making view is nonetheless able to defend biological purposiveness. Although mechanical explanations dominate the discourse around biological events, teleological explanations in fact make more sense of our intuitive understanding of biology. Walsh's defense of teleological explanation, despite its weaknesses, offers a philosophical alternative to mechanism. If we accept his view, we gain a means through which we can explain biological events holistically as well as through an examination of biology's ever-smaller parts. Natural questions such as 'What is the purpose of X biological function?', a question which does not have a place in mechanistic explanations, make sense in the context of teleological explanation. The ability to ask such questions is a distinct advantage of Walsh's view.

## Bibliography

[1] Mark Bedeau. Where's the good in teleology. *Philosophy and Phenomenology Research*, 52(4):781–806, 1992.

[2] Angela Breitenbach. Teleology in biology: A kantian perspective. *Teleology*, pages 31–56, 2009.

[3] Carl G. Hempel. The function of general laws in history. *The Journal of Philosophy*, 39(2):35–48, 1942.

[4] Immanuel Kant. *Critique of the Power of Judgement.* Cambridge University Press, 2000.

[5] Denis M. Walsh. Organisms as natural purposes: The contemporary evolutionary perspective. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 37(4):771–791, 2006.

[6] James Woodward. Explanation and invariance in the special sciences. *The British Journal for the Philosophy of Science*, 51(2):197–254, 2000.

# 3

# *Direct Quantum Explanation*

## *Or, Why Bob Discards Keys*

**Stefan Vladusic**
*University of Toronto*

## Introduction

There has been push-back in both the literature of philosophy of physics and philosophy of scientific explanation claiming that physical phenomena can be legitimately causal. This is especially clear when considering quantum mechanics, since many of the field's founding figures believed that causation and determinism were inextricably linked. Therefore, the common view that quantum mechanics is fundamentally probabilistic suggests that there is no way to causally interpret quantum phenomena. As a result, many in the physical sciences continue to believe that causation and quantum mechanics cannot sit well with one another.

However, some philosophers have abandoned the presupposition that causation and determinism are two sides of the same coin. Hans Reichenbach may have been the first philosopher to do so, presenting his statistical "principle of the common cause" in the 1950s. Reichenbach's original position may have fallen out of favour, but the notion of a probabilistic account of causation certainly has not. While many indeterministic accounts of causation have emerged, James Woodward's Interventionist account has proven to be one of the most influential in the recent literature [9]. In brief, Woodward believes that $A$ causes $B$ if aspects of $B$ are controllable or can be manipulated through $A$. That means that even if $A$ can only change the probability of $B$ obtaining some state, $A$ can still be a genuine cause of $B$. For example, suppose that $F$ describes whether a coin flip is heads or tails. Now suppose that there is a very resourceful cheater, who has some way to make the coin rigged to land on heads 70% of the time. Though the cheater cannot deterministically set the outcome of any coin flip, they can nevertheless manipulate or control aspects of $F$. So, whether or not the cheater cheats is a cause of, or is at the very least has a causal influence over, $F$.

Given that Woodward's account of causation is not deterministic, Shrapnel argues that certain quantum phenomena are genuinely causal, and can be explained causally [6]. In particular, Shrapnel argues that the radical pair model of the avian magneto-compass is legitimately causal and explanatory. Expanding on Shrapnel's work, I claim that the BB84 quantum cryptographic protocol provides another example of quantum causal explanation within Woodward's framework, and is different in kind to Shrapnel's example. Namely, it is based on theoretical (rather than experimental) arguments and involves direct intervention on the quantum attributes of a causal system. In section 2, I will briefly summarize Woodward's Interventionist account, and Shrapnel's discussion of the avian magneto-compass. In section 3, I outline the BB84 protocol and argue that it is a genuine example of quantum causal explanation. Finally, I will argue that the causal nature of the BB84 protocol is different in kind from that of the avian magneto-compass, in section 4.

## Interventionism and the Radical Pair Model

Prior to discussing the BB84 protocol, it is important to review Woodward's interventionist account, and Shrapnel's analysis of the avian magneto-compass. Succinctly, Woodward believes there exists a causal relationship between variables $C$ and $E$ if there exists some intervention variable $I$ such that $E$ bears the same relation to $C$ under some range of changes to $C$'s value via $I$. For $I$ to qualify as an intervention on $C$, it must be some nomologically possible cause of $C$ such that: (1) for some range of values taken on by $I$, $C$ is not affected by any other causal variable, and (2) I ensures that there are no correlations between $C$ and $E$ other than a direct causal network spanning the two variables [9]. For a more detailed overview of the Interventionist account, see Woodward and Hitchcock, Frisch, or Shrapnel [4, 6, 10].

It seems that describing certain quantum phenomena causally (in Woodward's sense) is difficult if not nomologically impossible. Quantum phenomena (like indeterminacy, superposition, etc.) suggest that intervening on some variable without affecting causal background conditions may simply be impossible in some cases. Furthermore, it seems that even if an intervention were possible, it need not change the value of a causal variable. Consider an experiment with only two possible outcomes $\alpha$, and $\beta$. Textbook quantum mechanics tells us that there is some quantum system such that if this experiment is performed on this system, then the outcome of the experiment is certain to be $\beta$. Now, suppose that a particle is in this definite state, but is changed (via some intervention, assuredly) so that there is a genuine 50/50 split between the $\alpha$ and $\beta$ outcomes. In this case, it may very well be that the outcome is still $\beta$, despite the intervention. Formally speaking, if $\{|\alpha\rangle, |\beta\rangle\}$ is the eigenbasis of our observable corresponding to the measurement outcomes, such that the incoming state is as $|\beta\rangle$ prior to an intervention, and as after an intervention. It follows that measuring the system will yield a 50/50 split between $\alpha$ and $\beta$. In this case, it may be that the measurement outcome is $\beta$, and the intervention hasn't affected the experimental outcome. Though Woodward's account permits these scenarios [8], these facts, alongside a skepticism of causation's role in physics, help explain the apprehension towards accounts of quantum causation (see Woodward [9] and Suarez [7] for overviews of some skeptical positions).

Despite these weighty complications, recent work from author like Shrapnel, Frisch and Suarez argue against the orthodox (and defeatist) position towards quantum causation [4, 6, 7]. In particular, Shrapnel uses the avian magnetocompass as an example of a causal system which is partially quantum [6]. Recent work by Cai, Guerreschi, and Briegel has suggested that a European robin's navigational system may manipulate quantum systems via the radical-pair model [2]. Their model suggests that within the bird's retina, there is an entangled electron pair in a singlet state. An incoming photon can then change the singlet state into a state between the singlet and triplet state. The model accounting for the electrons' dynamics depends partly on the Earth's magnetic field. Depending on the collapse state of the electron pair when encountering a reactant molecule, different chemical reactions will occur. If repeated many times, these reactions change the visual field of the bird. This change in vision is what helps the robins navigate [2].

Importantly, Shrapnel claims that a magnetic field can act as an intervention variable on the system's spin state [6]. So, the radical-pair model describes a hybrid deterministic/probabilistic causal system in which an intervention variable can act on entangled quantum states. Since Woodward's account permits probabilistic causal networks [8], the radical-pair model satisfies the interventionist requirements for causation in every respect. And just to sweeten the deal, scientific experiments have shown that controlling local magnetic fields creates difficulties in a robin's navigational abilities [6]. So, intervening on a magnetic field is more than a mere nomological possibility – it has been actualized in experiment. The upshot is that Shrapnel seems to have found an example of quantum causation.

## The BB84 Protocol

In the spirit of Shrapnel's project, I propose that the BB84 protocol also serves as an example of quantum causal explanation. Note that this section is devoted to establishing that the BB84 protocol yields quantum causal explanations in the interventionist framework – a discussion of the BB84 protocol vs the radical-pair model is in section 4.

The BB84 protocol is the earliest quantum key distributor and quantum cryptographic protocol [5]. Whereas classical cryptographic protocols utilize purely mathematical principles to create secure cryptographic keys, quantum key distributors use the inherent indeterminacy of quantum measurements to securely distribute cryptographic keys. In BB84, this is achieved using photon polarization [1].

Suppose a photon with a polarization of angle  is sent into a detector oriented at angle $\beta$. The probability of its being transmitted is $\cos^2(\alpha - \beta)$, and of being absorbed is $\sin^2(\alpha - \beta)$. The system will behave deterministically if and only if $\alpha \parallel \beta$ or $\alpha \perp \beta$ [1]. These results are translated into the formalisms of quantum mechanics by considering these vectors:

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \qquad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix} \tag{3.1}$$

Here $|1\rangle$ represents horizontal polarization, and $|0\rangle$ represents vertical polarization. The set $\{|0\rangle, |1\rangle\}$ forms a basis for a 2-dimensional Hilbert space. So, any polarization state can be described by the state vector:

$$|\psi\rangle = \begin{bmatrix}\cos\alpha\\ \sin\alpha\end{bmatrix} = \cos\alpha \begin{bmatrix}1\\0\end{bmatrix} + \sin\alpha \begin{bmatrix}0\\1\end{bmatrix} \tag{3.2}$$

I'll unpack the formalism here, so that the ensuing arguments are easier to understand for those without a technical background in physics. One of the postulates of textbook quantum mechanics is that the state of some quantum system can be described by a vector in a Hilbert space. While I will not provide a formal description of this postulate, it may help to think of it as a generalisation of vectors as understood in high school physics. Vectors are often introduced as arrows of a certain length which lie in space. Vectors in a Hilbert space are like these vectors, except that vectors in Hilbert spaces often encode something other than position and direction. In the BB84 example, for instance, the vectors encode information about the photon's polarization. The idea is that, formally speaking, the polarization of any photon can be described as a vector in this Hilbert space, and can be written as a combination of the $|0\rangle$ and $|1\rangle$ vectors.

It is also clear that the Hilbert space is spanned by the orthonormal basis $\{|+\rangle, |-\rangle\}$ where $|+\rangle$ and $|-\rangle$ are defined as follows:

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|1\rangle + |0\rangle) \\ |-\rangle &= \frac{1}{\sqrt{2}}(|1\rangle - |0\rangle) \end{aligned} \tag{3.3}$$

Consequently, every state vector $|\psi\rangle$ can be re-written in terms of the $|+\rangle$ and $|-\rangle$ vectors. Again, I will unpack the formalism. Here, I have outlined two ways of describing a photon's polarization as a vector in a Hilbert space. The first makes reference to the vectors $|0\rangle$ and $|1\rangle$, the latter $|+\rangle$ and $|-\rangle$. However, to those less familiar with quantum mechanics, it may be unclear what this maneuver has achieved. The idea is that while both bases of the Hilbert space can encode information about a photon's polarization, each basis is more convenient to use when describing certain kinds of measurements. A useful analogy is that these different bases are like everyday speech and philosophical jargon. While it is possible to describe complex philosophical ideas using everyday speech, it is often more convenient to simply use jargon. In the case above, changing between the $\{|0\rangle, |1\rangle\}$ and the $\{|+\rangle, |-\rangle\}$ bases has a clear physical interpretation – it reflects that the orientation of the measurement apparatus has changed by 45° [1]. Thus, if one prepares a photon in the $|+\rangle$ state, a measurement of this state in the $\{|0\rangle, |1\rangle\}$ basis yields the following probabilities:

$$\langle 0|+\rangle = \frac{1}{\sqrt{2}}\langle 0|0\rangle + \frac{1}{\sqrt{2}}\langle 0|1\rangle \tag{3.4}$$

$$= \frac{1}{\sqrt{2}} \tag{3.5}$$

$$\Rightarrow \mathbb{P}(|0\rangle) = |\langle 0|+\rangle|^2 = \frac{1}{2} \tag{3.6}$$

Similarly,

$$\mathbb{P}(|1\rangle) = |\langle 1|+\rangle|^2 = \frac{1}{2} \tag{3.7}$$

Thus, if an incoming photon is characterised by $|+\rangle$ there is a 50/50 chance of measuring the photon's polarization in either the $|0\rangle$ or $|1\rangle$ eigenstate. In fact, preparing a photon in an eigenstate of the non-measurement basis yields a 50/50 split in terms of which state will be measured in the other basis [5].

With this background information, I can describe the BB84 protocol. Note that the following presentation of the protocol is my own (I've simplified the language to try and make it easier to read), but the procedure is fundamentally the same as that suggested in the original paper written by Bennett and Brassard [1].

Consider two agents, Alice and Bob, who wish to send each other a cryptographic key over a public channel. Specifically, they want to make sure that Eve (a malicious third-party) cannot figure out their key. This could be achieved classically, but Alice has a single photon emitter and Bob has a single photon detector. So, they develop a clever protocol:

1. Alice generates a random string of bits and a random sequence of polarization bases (members of the sequence are the 1/0 or the plus/minus basis) of the same length.

2. For each bit, Alice sends either a '1' or '0' photon (encoded by $|1\rangle$, $|+\rangle$ and $|0\rangle$, $|-\rangle$, respectively) in the appropriate basis.

3. Bob then measures each photon's polarization, choosing measurement bases at random.

4. Alice and Bob publicly compare the bases of each measurement and discard measurements made in different bases.

5. Alice and Bob then randomly choose some bits which make up a significant part (say one third) of the reduced key and compare measurements publicly.

6. If the number of discrepancies between Alice and Bob's measurements are less than some predetermined percentage, the remaining measurements become Alice and Bob's key. If not, Alice and Bob repeat 1-5 until they are satisfied.

It may not be immediately clear what this accomplishes, so consider the case where Eve 'listens in' on Alice's transmissions. In order to measure some photon's polarization, Eve must choose a measurement basis for her measurement device. Suppose for argument's sake that Alice and Bob have both selected the 1/0 basis, but Eve has selected the plus/minus basis. Furthermore, suppose Alice has sent over a $|0\rangle$ state. Then Eve's measurement will result in either the $|+\rangle$ or $|-\rangle$ state. When Bob measures the state altered by Eve, there is a 50% chance the state will collapse to $|1\rangle$ – the opposite 'quantum bit' sent out by Alice! So, if Eve measures every photon, she will induce a 25% error rate, as the probability that Eve chooses the basis not used by Alice is 50%, and the chance that Eve's measurement will collapse to the other eigenstate in Alice's basis is also 50%. Hence, if Alice and Bob have an error discrepancy of at least 25%, then they know Eve has been 'listening in' [1].

Consider the protocol within the Interventionist framework. Now, in the literature of scientific explanation, an explanation is generally viewed as an answer to a why question. In this vein, I believe that the BB84 protocol is explanatory because it can answer "Why did Alice and Bob reject this cryptographic key?" We can embed the answer to this question – that Eve tampered with the system – in a network of causal variables. By showing that Eve's listening in is an intervention on the acceptance status of a quantum key, it seems that Eve's tampering explains Alice and Bob's acceptance or rejection of a key. This network and its relationships – I hope to show – are then the explanandum of the why question. And so, finding an explanation boils down to investigating which variables are invariant under some intervention on a causal relation. According to Hitchcock and Woodward, if this can be done, we can explain this why question [10].

When accepting or rejecting a key, there are clearly two salient considerations: Alice's original key, and Bob's measured key. Now, it seems intuitive that Bob and Alice's actions can causally influence the variable corresponding to accepting or rejecting a key (henceforth $A/R$; $A/R = 1, 0$ for accept, reject). However, I don't think this is the case. Rather, I believe that Alice and Bob's actions can be subsumed as background conditions figuring in the stability of the causal network.

Consider Bob's act of measuring some incoming states (let this be represented by the variable B). There are only two ways one of Bob's measurements can differ from Alice's incoming state ($S'$) – by choosing to measure the incoming photons in the basis opposite to Alice's, or by recording an error introduced by Eve [1]. I believe the first case is not genuinely causal. If it were, then we could find some sort of switch so that the values of other causal variables (namely $E$ – the variable corresponding to Eve's measurements) are held fixed. But this switch must hold E to be zero, for if $E = 0$, then $E$ would clearly affect the ensemble of states $S'$, contradicting the assumption.[1] Thus, all Bob's measurements can do is propagate Alice's states ($S$). So, $S$ and $S'$ will turn out identical. But this just means that there is no testing intervention on $A/R$ since Bob's measurements will just agree with $S$. Thus, $B$ is not causally relevant to $A/R$ (recall that Woodward and Hitchcock take a variable with no testing interventions on some other variable $X$ to be causally irrelevant to $X$ [10]).

I have of course assumed there is no way of changing $B$ other than $E$, and this may seem counter-intuitive. However, I believe the other ways of changing $B$ – that is changing the basis of measurement, or by changing the state $S$ – must be subsumed as background conditions. The protocol demands that when comparing $S$ and $S'$ these strings agree on basis measurements when considering the former. So, implicit in the explanation is that Alice and Bob have already eliminated those bits in $S$ and $S'$ differing in basis measurements. However, basis choice is clearly relevant in the explanans – without it accounting for bases, the protocol makes no sense. Hence, basis choice must be what Woodward calls a *background condition* [9] – a variable which does not explicitly factor into the causal network of $A/R$, but is nevertheless important to the stability of said causal network.

The same reasoning shows that Alice's original state must also be a background condition. Assuming that S features explicitly in the causal network of $A/R$, $E$ must equal 0 when intervening on $S$ (per reasoning above). Now, changing S just corresponds to generating a new ensemble of photons. This presents a contradiction. It's clear that modifying $S$ modifies $S'$ (since $S = S'$ if $E = 0$). But this just means that there is no testing intervention on $A/R$ – $S$ and $S'$ being indistinguishable guarantees $A/R = 1$. Thus, $S$ cannot be a cause of $A/R$. Yet, it's obvious that $S$ is salient in the explanans. So, $S$ ought to be considered a background condition.

The upshot is that in the idealized BB84 experiment, the only causal variable *explicitly* acting as a putative cause of $S'$ is $E$ – whether Eve intervenes on the system, and by how much. Since the value of $A/R$ depends on $S$ (a background condition) and $S'$, it follows that the only explicitly causal factors affecting $A/R$ are $E$ and $S'$. Hence, Eve's interventions on the states measured by Bob constitutes a proper cause on the variable $A/R$. It generates a testing intervention, there exists a set of circumstances where $E$ is the only cause of $A/R$ (other than other causes lying on the causal pathway between $E$ and $A/R$), and it trivially leaves the values taken by other causal variables unchanged. Thus, I believe that given Hitchcock and Woodward's account of causal explanation [10], Eve's eavesdropping constitutes a genuine explanation

[1] Note that I assume Eve must induce an error since the key is sufficiently long. This assumption is made in the original paper, and it seems compatible with Woodward's probabilistic account. See [1, 8]

of why Alice and Bob either accepted or rejected some quantum key. Importantly, Eve can only modify the measured state $S'$ by directly modifying the polarization of incoming photons – a quantum phenomenon.

## BB84 vs. the Radical Pair Model

While I believe that the example above does demonstrate genuine explanation through quantum phenomena, there are many possible objections. This section tries to address some possible responses. I will also try to demonstrate that the difference between the BB84 protocol and the radical-pair model is philosophically significant.

There are several concerns I will not address in detail since they are outside of the scope of this paper. I will discuss neither the legitimacy of coarse-graining quantum variables (which was done in section 3), nor the role of causation in quantum physics. All I will say is that Frisch and Shrapnel both take this coarse-graining to be legitimate [4, 6], and since my paper is an expansion on their work, so will I. Similarly, I take the work of Woodward [9] and Frisch [4] to demonstrate that causation in physics is possible. This may not satisfy the skeptic, so if one desires, one can view my argument as depending on legitimacy of these two notions.

Some may argue that the BB84 example is far too idealized to be insightful since it neglects noise, quantum hacking attacks, non-secure communication channels, and so on (see Lo, Curty and Tamari [5] for an overview of some of these topics as they relate to physics). I disagree for two reasons. The first is that the idealized protocol demonstrates how quantum phenomena can be used in a 'ground up' causal explanation – I will discuss this later. The second response is more pragmatic; despite the idealization of BB84 in my argument and the original paper, the protocol has been successfully implemented in real life over relatively large distances (100km) in non-ideal conditions [5]. Furthermore, there are also commercially available quantum key distributors. Even though these commercial systems are not explicitly based on BB84, their criteria of accepting or rejecting a key are still based on the manipulation of quantum systems [5]. And so, we can see that quantum cryptographic protocols can be successfully implemented in practice and are not merely theoretical.

I think a much more compelling objection to my example concerns whether the explanans to "Why did Alice and Bob reject/accept the key?" is legitimately quantum mechanical in nature. Furthermore, some may believe the explanans is quantum mechanical, but fail to see it as being different in kind to the radical-pair model. I believe that addressing the latter, showcases the genuinely quantum nature of the explanans.

As mentioned before, the BB84 example differs from the radical-pair model because it is a 'ground up' explanation. What I mean by this is that BB84 directly utilizes the notions of quantum indeterminacy and quantum mechanical statistics to create a cryptographic framework. From this theoretical framework, actual, practical key distributors have emerged. Compare this to the radical-pair model for the avian magneto-compass. This is a hypothesis which hopes to give theoretical justification to an already observed phenomenon and whose precise mechanisms are not entirely known [6]. Furthermore, there are competing hypotheses explaining the navigational abilities of birds [2]. So, in an important way, the BB84 protocol is built on quantum phenomena 'from the ground up' – fundamental quantum phenomena underpin the whole idea. Meanwhile, the radical-pair model uses quantum phenomena as a possible explanation of macroscopic events. This seems to be a genuine distinction when differentiating these kinds of explanations [3]. This difference is strengthened by the fact that the radical- pair model seems to work despite quantum effects, while the BB84 protocol only works because of quantum phenomena. Shrapnel's example – by her own admission – works because the reactants in the radical pair model are 'blind' to the electrons' entanglement [6]. The BB84 protocol, however, is fundamentally different – the protocol is only sensible because of quantum phenomena. If one were to conduct the BB84 protocol with classical (i.e. not quantum) bits then Eve's measurements would not affect Bob's measurements.[2] What this points out is that any intervention on the photons' polarizations – the relevant factor in determining whether a key is accepted or not – must not only be an intervention on a quantum system, but an intervention which directly affects quantum properties. The radical-pair model is only sensible when the system is 'blind' to some quantum properties (namely, entanglement). The BB84 protocol is only sensible when one directly intervenes on the quantum properties of some particle; in this case, photon polarization. Because of these differences, I claim that the BB84 protocol yields direct quantum explanation, whereas the radical-pair model yields indirect quantum explanation.

I believe that the discussion above demonstrates that the explanans of "Why did Alice and Bob reject this quantum key?" is irreducibly quantum mechanical in nature. Simply put, there is no meaningful way within the context of the BB84 protocol for Eve to 'listen in' on Alice's string for a sufficiently long time without directly influencing some quantum phenomena relevant to Bob's

[2] I mean 'affect' in a very literal sense here. If Eve measured classical bits, then there would be no change between $S$ and $S'$ . So, $E$ couldn't possibly be a testing intervention on $A/R$ for the same reason $B$ cannot be. On classical bits Eve's measurements cannot affect $S$ in the interventionist sense – I'm not using such loaded language willy-nilly!

measurements. As such, any explanation predicated on the causal network introduced in section 3 requires reference to direct quantum interventions on Alice's original photon ensemble; quantum phenomena must factor into such an explanation. As a final note, I believe that considering these two examples together further legitimises the notion of quantum causal explanation. Shrapnel argues that the radical-pair model showcases a form of quantum causal explanation which is based upon actual experimentation and manipulations of a quantum system without intervening on its fundamentally quantum characteristics [6]. The BB84 example shows an example of theoretical arguments stemming from direct interventions on the system's quantum properties. Thus, there are cases of quantum explanations in theory and in experiment; directly and indirectly.

## Conclusion

In this paper, I addressed two candidates for quantum causal explanation in Woodward's interventionist framework – Shrapnel's example of the radical-pair model, and my own example of the BB84 quantum cryptographic key distributor. I argued that Shrapnel's example is explanatory because it avoids direct interference with quantum phenomena, and that its legitimacy is also predicated upon experimental rather than theoretical justification [6]. I first argued that the BB84 experiment can be translated into an interventionist causal network whose testing interventions directly modify quantum phenomena. I then argued that the protocol BB84 is predicated on theoretical, idealised arguments, and is only explanatory when one directly intervenes on the quantum characteristics of Alice's photon ensemble. In this way, the BB84 protocol alongside the radical-pair model demonstrates the robustness of quantum causal explanation – it can be both experimental and indirect; theoretical and direct in nature.

## Bibliography

[1] Charles H. Bennett and Gilles Brassard. Quantum cryptography. *Theor. Comput. Sci.*, 560(P1):7–11, December 2014.

[2] Jianming Cai, Gian Giacomo Guerreschi, and Hans J. Briegel. Quantum control and entanglement in a chemical compass. 2009.

[3] Dennis Dieks. Bottom-up versus top-down: The plurality of explanation and understanding in physics. In Henk De Regt, Sabina Leonelli, and Kai Eigner, editors, *Scientific Understanding: Philosophical Perspectives*, pages 230–248. University of Pittsburgh Press, 2009.

[4] Mathias Frisch. Physics and the human face of causation. *Topoi*, 33(2):407–419, 2014.

[5] Hoi-Kwong Lo, Marcos Curty, and Kiyoshi Tamaki. Secure quantum key distribution. *Nature Photonics*, 8:595–604, 2014.

[6] Sally Shrapnel. Quantum causal explanation: Or, why birds fly south. *European Journal for Philosophy of Science*, 4(3):409–423, 2014.

[7] Mauricio Suarez. Causal inference in quantum mechanics: A reassessment. In Federica Russo and Jon Williamson, editors, *Causality and Probability in the Sciences*, pages 65–106. College Publications, 2007.

[8] James Woodward. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, 2003.

[9] James Woodward. Causation with a human face. In Huw Price and Richard Corry, editors, *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford University Press, 2007.

[10] James Woodward and Christopher Hitchcock. Explanatory generalizations, part i: A counterfactual account. *Nous*, 37(1):1–24, 2003.

# 4

# A Most Unexpected Journey

## An Interview with Freeman Dyson

**Patrick Fraser**
*Senior Editor, University of Toronto*

In 1940, Trinity College at the University of Cambridge had a new student of mathematics. Freeman Dyson, 17 years old at the time, had begun what would prove to be an extremely fruitful scientific career. In 1947, Dyson moved to America where he began studying the up-and-coming theory of quantum fields (which would eventually become the basis for the Standard Model of particle physics). It was here that Dyson made his first breakthrough in theoretical physics, unifying all of the major approaches to the theory at the time, and solving many foundational problems in the field. After these major achievements, Dyson didn't slow down. He continued to make significant discoveries across many subjects, not just physics, including in nuclear engineering, space science, biology, and many others. Dyson's latest book Maker of Patterns, is a collection of personal letters written throughout his life and career. With many unexpected twists and turns, and with numerous cameo appearances by characters such as Ludwig Wittgenstein, Paul Dirac, G. H. Hardy, Kurt Gödel, Albert Einstein, Francis Crick, John von Neumann, Niels Bohr, Richard Feynman, T. S. Elliott, and Stephen Hawking, his life story reads more like a thrilling novel than anything else.

Freeman Dyson is a man of many hats. At times, a theoretical physicist unifying quantum electrodynamics, a biologist studying the origins of life, a space scientist developing novel approaches to space travel, a Cold War diplomat advocating for unilateral nuclear disarmament, an author, and arguably even a philosopher (though he surely would reject this characterization). By his own description, however, he has always been a mathematician first and foremost. He has made many substantial contributions to a wide variety of distinct fields in modern science, and although he is now 95 years old, he is still actively writing books, giving talks, and participating in scientific discourse. I spoke with Professor Dyson in his office at the Institute for Advanced Study in Princeton, New Jersey.

*The following interview has been edited for length and clarity.*

PATRICK FRASER (PF): Thank you very much for meeting with me Professor Dyson. A question that I wanted to start off with: You've managed to go your entire career without a PhD. That's quite unusual. How did that come about?

FREEMAN DYSON (FD): Well, there were two quite separate reasons. First of all, I happened to grow up in England where the PhD wasn't so well established. There were lots of people in England who didn't have PhDs, and it was regarded as essentially a foreign invention designed for Germans who came to England and needed some piece of paper to take home. Then when I came to America, I became convinced that the system was doing a lot of harm; I saw these American graduate students whose lives were being ruined, really, by the system and so I didn't feel any inclination to join them. It's a stupid system; it's designed for German academics in the 19th century and for them it was fine, but for a normal career in the United States, it does not make sense. Especially not for women; it was particularly bad for women because it takes up so much time at a crucial period of their lives when they are starting to raise families and it's extraordinary how the system so screws them up. I think that's a big reason why there are fewer women than men in academic life; it's designed for men, and not for women, and that was stupid of us to have taken it over. I have strong feelings about that, so I'm proud for having not a PhD myself, and I'm proud of having not even any PhD students, and I'm proud of having six kids, none of them having a PhD.

PF: So, when you did begin your studies, you started out studying Mathematics at Cambridge. What lead you, as an adolescent, to want to pursue pure mathematics?

FD: So, it was not really as an adolescent. One of my earliest memories is actually being put for a nap after lunch in my crib, I must have been about three years old, but I started figuring out that if you added one, and a half, and a quarter, and eighth, and a sixteenth, and go on to infinity, the answer is two! And so that was my first venture into pure mathematics. This was a source of great satisfaction, and I don't think I talked about it to anybody. Then I read the book of E. T. Bell which you may know, *Men of Mathematics*, which is biographies of mathematicians, and that's also a very inspiring book. And so that confirmed in me the idea that studying mathematics was a good idea.

PF: So, at this point in your career, you've worn many hats, and so would you still identify yourself as a mathematician at heart?

FD: Yes! I mean, my tools are all mathematics.

PF: Right. That's fascinating, because it seems as though you've gone into many different fields, like biology and astrophysics, and a great wealth of other areas. It's interesting to see that overarching theme carry forward all the way.

FD: Right. No, I've always had a short attention span. So, I had to jump around from one thing to another.

PF: Getting more related to philosophy of science, you've witnessed, and actually played a great role in many different revolutions in contemporary science, and so this puts you in a unique place. From your perspective, are scientific theories approaching some kind of absolute truth about the physical world, or do you see science more as a proliferation of diverse ideas that all describe nature in a unique way, even if they are not necessarily converging to some absolute truth?

FD: We don't know. I mean that's what makes science exciting of course; it's so completely unpredictable. But there have always been these contrary tendencies, especially in physics; we have unifiers, who are people who produce great ideas: we have quantum mechanics, which unifies chemistry and everything else, and it was an amazing time for unifiers during the quantum revolution. And on the other side, there's the diversifiers who are going out and looking at the real world, and finding all sorts of unexpected surprises, and making things more complicated, and that's what makes it so fruitful. You have both of these things going on at the same time, and so which will win is of course totally unpredictable but I sort of hope that neither will win. I would say with my image of the world is an infinite forest with beautiful trees, and all kinds of beautiful things, and we're exploring; we're simply going deeper and deeper into the forest. That doesn't mean that we'll ever get to the other side.

PF: This reminds me of something that Wittgenstein said in his *Tractatus*, which was that the best way to teach philosophy would be to tell students individual facts of natural science, and as soon as they try to interpret them, show them some reason why they're wrong.

FD: Yes. That's what we do, of course, and Wittgenstein talked a lot of nonsense, but he also talked sense and that's what philosophy is. I always feel philosophy is really a branch of literature, and not a branch of science.

PF: Interesting. A lot of scientists have openly expressed disdain towards philosophy, and in particular, towards philosophy of science. People like Steven Weinberg have said that philosophy of science is not actually a useful pursuit, and might hold science back, for instance. How do you think science and philosophy ought to interact?

FD: I don't think they should. I rather agree with Weinberg; I wrote a review of a book about philosophy, where I asked: Why did philosophy lose its bite? There was a time when philosophers were important. They talked to kings, and emperors, and what they said mattered, and somehow, since they've become academic, they become less and less important, and nobody cares any longer what philosophers say; they only talk to each other. And why did that happen? I think it's sort of largely driven by this idea of making philosophy scientific, which turned out to be a mistake. As long as philosophy dealt with life in general, it was far more powerful. When it starts talking about the details of physics and chemistry, then it becomes unimportant.

PF: So how would you characterize your image of science in general?

FD: For me, science is not a grand vision of the universe. Science is a bunch of tools that happen to be remarkably effective in answering a limited group of questions.The tools of science are like the tools of music. Violins and cellos are amazingly effective for producing sounds that move our emotions. Lasers and computers are amazingly effective for exploring how natureworks. I have always loved to play with mathematics just as a musician loves to play a piano.

PF: That's a very elegant way to put it. To switch gears a little bit, you've worked on a variety of different problems throughout your career. What lead you down different pathways?

FD: I'm always looking around for interesting problems and it's a matter of chance whether I happen to find a problem at the same time that I'm not doing something else. But I think it's almost always happening to meet the right people at the right time. The big event in my scientific career was quantum electrodynamics. I came to the United States as a student of Hans Bethe to do theoretical physics, and it just happened by an enormously lucky chance, that there was Dick Feynman. I'd never heard of Dick Feynman, and within a week I'd already found him, and I understood that this was a genius who was really doing great stuff, and I could learn a lot by listening to him. So it just happened, just like that, and within a week or two of arriving in the States, I knew more about physics than almost anybody else, just by luck. And afterwards, I met a lot of other people in the same fashion, just by being with the right people at the right time. I happened to meet Francis Crick during the war in London, and so I got interested in Molecular Biology before it existed.

PF: If we look at the modern environment of scientific institutions, we see that a lot of people are forced to hyper-specialization; a lot of people spend their entire careers working on one niche topic, which is of course in contrast to the wide range of research programs that you've worked on. What advice would you have for younger academics who are likewise interested in many different topics, and who also have a hard time focusing on one thing for too long?

FD: Well, of course, the two subjects which I'm most interested in today are astronomy and biology, and both of them have this characteristic that they are diversifying rather than unifying. Every week, we have the astronomers' lunch, which is, for me, sort of the main event of the week. It's a gathering of astronomers every week, and there's always something new. It's a subject that's advancing so rapidly, just because of the huge number of interesting things in the sky, so there are problems for everybody, and it's not hard to find something that nobody else is doing, so that's what makes it so lively. It's just going ahead like a rocket ship, and of course theoretical physics is slowing down to a snail. I've lost interest in theoretical physics for that reason; it seems to me that it's slowed down to a point where it becomes boring. But the same rapid advancement is true in biology of course, so we have a little group of what they call systems biologists here who are sort of theoretical – many of them are physicists who switched because they found physics was getting boring – and in biology, there's the same enormous wealth of interesting problems. So, I would say that's where you should go. I mean, take your choice, but those two subjects are just much more fun than anything else at the present time. Of course, one of the tragic features of the landscape is fusion. We have here a fusion project in Princeton which has been going on for seventy years or so, and quite a number of talented people went into it, and I had the good sense not too. They went into it with tremendously high hopes; fusion was really going to change the world, but it's been such a dead loss. It turns out they made a bad mistake of switching from science to engineering much too soon,. That was driven by the politics; you could get money to build big machines, but you couldn't get money to do little experiments.

PF: How do you feel about this rise of big science in general?

FD: Well, there are two different problems there. The problem of big science is in the huge projects, where I think it has some sort of a boundary around a billion dollars. A project after that is driven by politics more than by science, and that then becomes very harmful. Of course, the worst example was the space shuttle, which was sort of fake science program. One of my students went up as an astronaut on the shuttle twice and he said it was a wonderful camping trip. He had to pretend to be a scientist, but it was just sort of Mickey Mouse science. It was totally worthless, really, as science, but he had to pretend to be doing science to get a seat on the shuttle. So, it cost us a billion dollars just for him to have a few weeks of happy life in space. Of course, it was a jobs program; it was providing jobs for a lot of people, and that's why it was kept alive. That's a typical example of what happens when you get over a billion dollars. The two countries which have actually done best in particle physics in the last thirty years or so are Japan and Canada. They didn't have too much money. SNOLab of course has really been very good, and I don't remember what the cost figures are, but certainly below a billion, and that was a brilliant piece of work. And the Japanese Kamiokande is a similar thing. I would say that for particle physics, that's the way to go; to big underground detectors.[1] The beauty of the underground detectors is that they can actually discover things you didn't expect, whereas at CERN, you can't do that. CERN's hadron collider background is so enormous you can only see what you've programmed into the detectors; you never make an unexpected discovery.

PF: It's always about confirmation. To that end, do you think a project like CERN is even worth carrying out?

FD: Well, it's hard to say it's not worth carrying out. The question is what else should they have done? I think it would have been much better to have five or six instruments, each of them of modest size so you could change quickly from one to another. They probably would have done better science, but you can't say for sure, I mean the fact is, politically, it's been a huge success; it's brought the European countries together

[1] Both SNOLab and Kamiokande detect subatomic particles using large tanks of water buried deep underground.

in a very happy relationship. I can't say I'd want to abolish CERN, but for science, it certainly wasn't a good idea.

PF: Right. So as you've already said, you are quite critical of manned space flight as well, and the Shuttle program ended a few years ago. Are you sad to see that program end?

FD: No. Oh, no, that was long overdue. Bob Dicke invented this corner cube laser reflector which was taken by the astronauts to the moon. It was wonderful because it was really an honest piece of science that the astronauts could actually help with and it was obviously right for Apollo, but Bob Dicke, being a practical person, actually built the whole apparatus himself. He bought the corner cubes at the Edmond's Scientific Company down the road and the whole thing assembled cost $5000, and he showed it to NASA and they said oh yes, we'll do that, that's wonderful. And so Bob Dicke said: Shall I make 5 of these for you?, and NASA said: No, no, you don't get to do it, we get to do it. So, NASA did; NASA built it themselves. It cost $3 million apiece, so the cost increased by a factor of 600 doing it the NASA way. Now we are seeing some of these commercial operations are really doing a lot better than NASA. There are some things NASA does well, and other things it does badly. It's good they have both of them now. But what is wrong with the manned program was that it was sold to the public under false pretenses as a science program when, of course, it's actually just a sporting event.

PF: Right. Getting back to stuff that's more oriented towards physics, I noticed that Professor Witten is just across the hall which I am somewhat startled by. I recall your description of one of his talks in the 80's about superstring theory and describing it as a definitive moment where physics was pushed two levels of abstraction further. Of course, today, we see this proliferation of a wide range of very abstract theories, which could almost be labelled as purely mathematical theories, so my question is: Do you think that physics has, at this point, departed from the realm of the physical, and has more been subsumed by mathematics?

FD: That certainly is true. An analogy which I like to talk about is Lie algebras, which were invented by Lie, and he was really a mathematician who thought he could provide a basis for classical physics, and in fact it was a failure. Physicists were not interested as it was too mathematical for them, and mathematicians were not interested because it was too concrete for them, and so he died a disappointed man. And then fifty years later you had quantum mechanics where everything is linear and Lie theory turns out to be the language nature speaks, and it fit wonderfully for quantum field theories, so that's a good example where it seemed to be a total failure, and then it turned out to be a huge success a hundred years later, but so it could happen again. So, I think all those abstract theories that are occupying all these young people around here, I would not say you shouldn't do that, but I think it is overcrowded. Last time I looked at the problems, there were about 10,000 string theorists, and they probably had problems for 1000.

PF: So, to what extent should our physical theories be actually grounded in empirical phenomena that they try to describe?

FD: Well, I would say as much as possible, but pure mathematics somehow is deeply embedded in the natural world, and that is of course the miracle; it's quite different in biology. I mean, in biology, mathematics is just a tool, on the very superficial level, but in physics, mathematics really takes you deep.

PF: Of course, in addition to physics and mathematics, you've studied the origins of life quite extensively. Why is this a topic which everyone, irrespective of their scientific background, finds so fascinating?

FD: It just sort of seems so obvious. When we're children, we ask how babies are born and it's one of the most exciting things to learn how babies actually are born, and to me, the origin of life just sort of goes with that, sort of absolutely. Natural curiosity is what we're born with, and it's an obvious question, the biggest question we don't know the answer to in a way. We found out how babies are born, but we still don't know about life.

PF: Right. Another question that occupies many, which seems to be avoided by most people when they try and study the origins of life, is the origins of consciousness. What is your perspective on things like the origins of consciousness and self-awareness?

FD: Yes, it's a wonderful question, and every time I look at a baby, I'm sort of asking that question. I mean, what's going on in that little head? Obviously, babies are conscious, but in a very different way. They are bombarded with this enormous stream of input coming in from all directions. How does a baby learn to sort it out the way they do? Of course, it's a question we could imagine being able to answer if we really understood the brain, so it's a scientific question in the end, but we're nowhere near getting to it.

PF: That's truly amazing. In addition to your scientific research, you have also been a very strong voice in discussions about policy-making for nuclear weapons, and safe nuclear science in general. What are your thoughts on how policies ought to be formed when potentially dangerous technology is developed by scientists?

FD: Well, there's a huge contrast between physics and biology there. Biology has really done well, and physics has done badly. When recombinant DNA was discovered, the biologists saw at once that this was dangerous, and also essential, so they set up an international meeting, stopped all the experiments for a year or so, and agreed on the guidelines as to what was allowed and what was forbidden, and it worked amazingly well. They did that themselves, before the government took any action and so there have been no real health hazards from that technology. And now, of course, they have the CRISPR which is presenting similar problems. Now it is even more acute because CRIPSPR is so powerful, but they do it the same way and so you have an international agreement among the experts, and then the governments are generally quite happy to put it into law, and that's exactly what the physicists did not do.

PF: Right. So, you think then that when scientists arrive upon a development that has potential for both great good and great bad, that they should be the ones to take that first step and bear responsibility for how it should be used?

FD: Yes, absolutely.

PF: Moving back to theoretical physics, with the development of modern physics, we see a lot of times people introduce unobservable entities, gauge fixing for instance. The theories that we use sometimes employ mathematical tricks and tools that disappear when the calculations settle. That's always seemed rather unfounded to me. Could you shed some light on why that's something we are able to do, or if that's something we should be allowed to do at all?

FD: Oh certainly you should be allowed. Setting up rules for how to do science is generally very counter-productive. But why it should work so well I can't tell you. It's an amazing thing, this whole mathematical world; it's so rich and somehow still so real.

PF: Right. But on the other side of things, we also see that, when we build our theories, we often note things like boundary conditions that we impose from our physical understanding, and we'll use these conditions to manipulate the equations that govern how our physical theories behave. So, if we're allowed to bring in these mathematical entities, then it would seem to imply that we are working within mathematics, so why then are we allowed to bring in these physical hypotheses which need not be grounded in any sort of mathematics? Why is it that we're able to just start cancelling stuff because when we go out into the world that's the way it looks?

FD: FD: Yes, well I can't answer that, but I can tell a story. Murray Gell-Mann started working on $SU(3)$,[2] and $SU(3)$ turned out to be what nature does. I remember having a big argument with Murray Gell-Mann; I didn't believe any of that stuff, and I had this argument about the ten-fold representation of $SU(3)$. Murray made that a big part of his symmetry; you had that array of ten particles which all had this symmetry, and I said look at the cross sections,[3] they're so different! I mean check the case of the edge of the decuplet triangle. You had four particles, and if you looked at the cross sections for producing them in a proton-proton collision, each step was 100 times smaller and so the excited state of the proton $P^{++}$ would be produced with a big cross section, and then they would go down by a factor of 100 for the $\Sigma^+$, and a hundred for the more for $\Xi^0$, and finally a factor of 1,000,000 for the $\Sigma^-$. How can you even talk about things being symmetrical when they're so different! They couldn't possibly be right, and Murray said it doesn't matter how big the cross sections are as this is just an abstract symmetry and that's how nature thinks, and of course he was right! So, one Nobel prize to him. It is amazing; nature doesn't care about magnitudes, all it cares about is the abstract structure, which to me was very hard to accept. Murray was really very full of insight.

PF: I have a few questions about some of your early quantum electrodynamics work, and in particular about the paper on large-order diverge in perturbation series in QED. You talked about how, with this divergence problem, either we're going to need additional mathematical content for our physical theories or additional physical content. I would be interested in hearing what your thoughts were at the time when you started looking at that problem, and when you started thinking about how physics would have to change to accommodate these large-order divergences.

FD: Yeah well, of course the joke is that we didn't take it as seriously. I mean that when we were putting quantum electrodynamics together, it all looked sort of make-shift. Obviously, it wasn't the final system, it was a sort of kitchen recipe for calculating which worked up to a

[2] $SU(3)$ refers to a particular abstract mathematical structure which underlies our modern theory of particle physics, and which purportedly tells us that the universe has a very high degree of symmetry.

[3] Cross sections refer to the observable scattering processes which we can actually test experimentally.

point. I expected it would only last 5 years and so a disappointing thing is that nothing better has been found! It's amazing, I mean it's sixty or seventy years ago, and every ten years, you have a new set of experiments proving quantum electrodynamics is correct to two more places of decimals, and now we are up to about twelve figures and it still works! We had all imagined we'd find discrepancies that would tell you what to do next and of course we've never found any discrepancies. It's been a sort of series of disappointments all the way.

PF: Right. Moving into more contentious territory, you hold several controversial views pertaining to climate change. Would you be able to explain what your main thoughts about climate change are, what causes it, how it's going to impact the world, how we manage it, things like that?

FD: Yes. The first thing to say is that there are two effects of carbon dioxide that are really important. There are the effects on climate, and the effects on plant growth, which has nothing to do with climate, and the whole discussion is entirely confined to half of the former. The effects on plant growth of course, are very well documented, and pretty well understood, much better than the effects on climate, and they are totally ignored in the public discussions, so that's part of what makes the situation so confused, so they're only dealing with half of the problem. Next thing to say is the climate is changing. No doubt about that. In the last fifty years, it's got quite a bit warmer, and there's no doubt that that's important, and at the same time, growth of vegetation has grown substantially. The Earth has actually become a lot greener in the last fifty years because of the carbon dioxide, so the effects are happening at the same time.

PF: Right. So, there's this trade-off between the impact on climate which is clearly negative, and the impact on plant growth with seem to be fairly positive.

FD: Yes. I would put it a little bit the other way. I would say the effects on plant growth are clearly positive, and the effects on climate are seemingly negative but we're still not sure. I actually visited Ilulissat, Greenland. It's the place where Al Gore went to make his film and where the global warming is most intense, and it's an amazing place. Huge ice cliffs falling into the ocean, it's all real, the pictures are real and if you talk to the people who live there, they love it. They have had a much better life since the warmth. In the old times, they had to go fishing in horrible weather, and one third of the entire male population died at sea. It was really bad, and now instead of going fishing, they take tourists in boats, and they don't get drowned anymore, and so life is a lot easier for them. So, they hope the warming continues. You see then, even global warming in the most intense places is not regarded as harmful. In other places it is harmful, but anyway

PF: It's something that you can't universally say is bad in all contexts.

FD: Not at all. There are so many different things that are changing. Now we have this problem of wildfires in California, which of course is a big problem, but still, humans on the whole flourish in a warm climate and so it's not at all clear that warming is bad. It's bad for some things, and certainly bad for wildfires, but then again, the cause of wildfires is very complicated too. It's not just climate, there's other things too; they made the mistake of not allowing fires to burn naturally, and so then these severe fires are a result of not having wildfires in the last fifty years. But the main question about climate change is whether it's caused by humans or not, and if it is not caused by humans, it probably means it's the sun. I mean quite a lot of detailed observations have been done in the last fifty years of the sun and the climate together, and there's a very strong correlation, so a big share of the whole thing is probably the sun, and of course the sun is totally out of our control, so until that is settled, you don't really know what the problem is and so it's much too soon to think of fixing it; you don't even know whether it's possible. So, all this talk of saying oh let's fix the climate I think is sort of meaningless. All we know how to do is stop burning coal, and it might or might not have a big effect on climate. We just don't know.

PF: Right. So, we could end up in some sort of catastrophic situation that is totally out of our control.

FD: Yes. For me, the elephant in the room is the ice-age. That is a far worse catastrophe than anything you can think of with warming. The ice-age was really a catastrophe, and it wiped out a whole lot of species, and this whole area was more or less under a kilometer of ice. And that could happen, we don't know; we understand that to just about the same degree as we understand the warming that's going on now. It's really a mystery why we have ice-ages. So, I'm more scared of having another ice-age than I am about having more warming. That's another factor which is not talked about.

PF: That's interesting. Of course, in your past writing, you've written about how, even in the most drastic of circumstances, life generally finds a way to survive, to persevere, and to thrive in the face of difficulties.

FD: Indeed, yes.

PF: So, if we accept this possibility that one day the human race may be entirely wiped out, and that life will somehow keep on going, but we will no longer be part of the story, what would be the collective obituary of the human race? What would be our mark on things if we were to totally disappear?

FD: It's an interesting question. I mean nature will go ahead and in the past when you had big extinctions, it usually took about 10,000,000 years for new species to take over and so it's a slow business, but it happens, and each time there's been a big extinction, there was actually a step forward because you had new species which were much more capable, like mammals replacing reptiles, and it was clearly a step forward, even though it looked bad at the time and I think the same would be true of us. Undoubtedly there's a general movement which in some sense is forward: better brains and better ability to deal with hardships and all kinds of improvements which are likely. If we were wiped out, I think it's very likely that another intelligent species would arise within the order of 10,000,000 years or so. I think an intelligent species is a risky business, causing all kinds of problems, and that's nature. I mean nature loves taking risks and we're part of that and we should recognize that you can't go ahead unless you take risks.

PF: Right. So, it seems like you have an attitude of the optimistic kind of how I learned to stop worrying and love the bomb'.

FD: Yes, exactly, yes.

PF: So, your accolades are impressive, but what do you think your greatest failures have been throughout your career, and do you have any regrets from you career as a scientist?

FD: No! Because that wasn't my priority. I mean I've always said that there are three things in life which are important: there's family, and there's friends, and there's work, and the priority is in that order and I always put the family first and I don't regret that. I mean I'm very happy with my six kids, and my sixteen grandchildren. That's my main contribution, and what I did for science was not that important in a way, and it was also great fun but I've always had the feeling with my science that I didn't care whether it was important; I would be just as happy doing just a piece of pure mathematics if something beautiful came out of it, of no interest to the rest of the world. I've never tried to be important as a scientist and so I don't regret the fact that I don't have a Nobel prize; I never earned it. And, of course, I've had a great time writing books and as a writer. Writing books opened up a whole new world. I mean, I made much more contact with people through the books than I did through the science, and I have a much wider variety of friends as a result, so that was also a choice. The second half of my life was mostly writing books rather than doing science, and I don't regret that.

PF: That's nice. I have one more question for you. You've come out with a new book recently, and you're clearly still very actively engaged with science, so what problems are of interest to you that you're currently playing around with?

FD: Well, this talk about evolution is the main thing at the moment, but most of the time of course I'm just enjoying the community I'm in, especially the astronomers. Having these active astronomers around occupies a lot of my time and just listening to them talk and reading a certain amount, and then of course the email is a great invention. In the old times, people would come and go, and then when they left, you'd lose touch. Now they go off to Australia, but you're still taking emails every day. You don't lose friends anymore. Having all these interesting friends is what it's all about.

PF: That's wonderful. Well that wraps up the questions that I had prepared.

FD: Thanks for coming.

PF: This has been a great pleasure. Thank you very much.

FD: Thank you.

# 5

# *Peter Galison on Photographing a Black Hole*

**Nicole D'Or and Patrick Fraser**

*Senior Editors, University of Toronto*

Dr. Peter Galison is the Joseph Pellegrino University Professor in physics and history of science at Harvard University. Trained both as a physicist and as a historian, Galison has brought a highly original perspective to the historical, philosophical, and social study of physics. His academic historical and philosophical work focuses primarily on the interactions between different physics subcultures, considering how groups of physicists doing experimentation, instrumentation, and theory interact with one another. Together with Lorraine Daston, Galison introduced a novel anthropological approach to studying science in terms of trading zones which has left a substantial imprint on the social study of science, and his book *Image and Logic* is recognized as the most thorough, honest, and historically precise treatments of what it means to 'do' physics in the modern era.

In addition to his academic research, Galison is also a recognized documentary filmmaker, having produced multiple award-winning documentaries, such as *Ultimate Weapon: The H-bomb Dilemma*, which explores the creation of the hydrogen bomb, *Secrecy*, which studies government secrecy, and his latest film *Containment*, which investigates the social and environmental dangers of containing nuclear waste, and the problem of safely storing extremely dangerous material for upwards of 10,000 years. In 2016, Galison helped to establish the Black Hole Initiative which, in 2019, helped to produce the first ever image of a black hole. Galison generously sat down with us during a rare opening in his schedule during his latest lecture tour.

*The following interview has been edited for length and clarity.*

NICOLE D'OR (ND): Thanks for sitting down with us. Could you tell us a little bit about the Black Hole Initiative (BHI), how it came to be, and how you came to be involved in it?

PETER GALISON (PG): My involvement began in 2015 when I collaborated with several colleagues; theoretical physicist Andrew Strominger, theoretical astrophysicists Avi Loeb and Ramesh Narayan, mathematician Shing-Tung Yau, and observer Sheperd Doeleman. We got together in the thought that black holes were like no other objects. They completely fascinate everyone, from mathematicians, to physicists, to observers, to astronomers, to philosophers, to school children, to science fiction writers, to filmmakers. It's just amazing that they have such a broad appeal. There are many objects that are completely fascinating, but they don't have that effect; a neutron star is like the nucleus of an atom, only the size of a city. That's amazing, but it doesn't interest kindergartners or science fiction writers or particularly mathematicians. They're just not in the same class. This is something that really interested all of us and so we put in for a big grant.They loved the idea of the combination of philosophy, physics, mathematics, and observation, so they took a chance on us. So that's how it got started, just with this absolute fascination with an object that grabbed the attention so widely and so far across the disciplines.

ND: It's clear that our understanding of black holes is influenced by many different perspectives, from quantum mechanics, astrophysics, and philosophy. Why is the multidisciplinary nature of black hole research important?

PG: It's a really interesting question. One of the things we didn't know at the beginning was whether this would be parallel play, as they say in psychology, or whether we would really form something together. Each of us had projects that we wanted to pursue. Andy Strominger was then collaborating with Stephen Hawking and with Malcolm Perry and with a young graduate student from Cambridge, Sasha Haco, on trying to solve the so-called information loss paradox, which I can describe if you're interested. Then, there was this attempt to make an image of a black hole – the Event Horizon Telescope (EHT) collaboration – then specific projects; Ramesh Narayan does simulations of the accretion disk around a black hole, this hot plasma that circulates around like the rings of Saturn or planets around the Sun. Shing-Tung Yau, the mathematician, is interested in questions that are raised by instabilities in spacetime and what they tell us about the behaviour of certain classes of equations that he studies. [And] the philosophical questions are great! What happens inside the so-called inner horizon inside a black hole? It looks like you might have solutions that allow backwards time travel or the singularity seems actually to be a point that's not in space and time at all, and what does that mean, and how do we interpret it, and what is the nature of what happens inside a spinning black hole or static black hole for that matter? What is the horizon? There are amazing, interesting questions in all these domains. So, we didn't know at first how much we would be able to create, what I've called in some of my writing, trading zones or interlanguages between different domains. Would we really be able to do something together? More than we could have ever hoped, we have.

At first, we'd have these colloquia where the physicists would [describe] an astronomical unit. I mean, just the basic tools of the trade were completely unfamiliar to these different groups, not just the principal investigators, the six of us, but all the graduate students and post-docs and even some undergraduate who were getting involved. So, we really had to learn a common way of speaking and to learn enough of each other's trade to be able to ask the right questions. So, that has happened. For example, we have a group, a collaboration between philosophers and physicists on the interior of black holes and how to understand that. I've started working on and become a member of the EHT which released this famous picture, which we will come to, I'm sure. The next generation of work may well involve collaboration with theoretical physicists asking questions about how to use the region near the horizon to make predictions of things we might be able to observe that would test aspects of some of the most advanced theories. So, it has really been successful beyond what we even hoped. Our best-case scenario has been exceeded, so that's been great.

ND: That's wonderful!

PATRICK FRASER (PF): With the EHT, there are all these different areas that can lead to further investigation, and there's a vast plethora of areas of inquiry which benefit from these results, but when it was portrayed to the public, it came across in almost every news article that I read about it, as "we've corroborated general relativity again". That's the big claim to fame. They did the same thing with LIGO, I think. Why does the public cares so much about corroborating an old theory and less about the new avenues that can be explored in light of these kinds of projects.

PG: It's a really good question. It's true, nobody has made money betting against Einstein; that's not a good source of income. But I would never describe that as the principal yield of making this image. It's true that in the medium term, we're now able to open up for persistent study, the region right outside the horizon of a black hole and that's amazing. That's what's really important. There are all sorts of things that we're going to be able to understand in that region of the most intense curvature of space and time that we can get at. That's what's important, I think. Then, one of the things that will come out of that will be a probe to see; are black holes the way we expect them from general relativity, does it violate aspects of general relativity, can we differentiate between black holes and other kind of exotic objects that might bend space and time in a somewhat similar way; wormholes and gravastars and boson stars and all these other things. And the answer is we probably will be able to differentiate some of those things.

Then there are other questions. These black holes are at once the darkest objects. Light falls in and doesn't come out. Theoretically, there's Hawking radiation, but that is, for such a big black hole, way below our level of detection. But they're also the brightest objects and the plasma circulates around it and heats itself up to somewhere between a couple of billion and ten billion degrees Kelvin. So, it shines and not only that, the black hole, if it's spinning, can drag space around it in such a way that you have this circulating charge, which creates a magnetic field like a little electromagnet and that can create this helical magnetic fields out of the north and south pole relative to its spin. This plasma gets pulled out into enormous jets that are the size of clusters of galaxies. That's amazing, right? In some ways, this black hole, which is tiny compared to other things, is the source of one of the biggest objects in the universe. It's the darkest object but it's the brightest object. If the beam happens to be pointing toward us, as some of them are, we call them quasars. They're the furthest things we can see, that is to say, the earliest things we can see. Many billions of years ago, the light that we're seeing was sent to us. So they represent kind of lighthouses at the farthest points in the universe from us. We can learn about the structure of the universe from those. Before anyone knew what the power of these quasars was, before anyone knew they were powered by black holes, they were fascinating. Now that we know they're powered by these enormous black holes, these supermassive black holes that seem to be at the centre of every galaxy, and that's amazing. There are about a hundred billion galaxies in the visible universe; that means there are about a hundred billion of these supermassive black holes.

PF: That's pretty cool!

PG: That is pretty cool! For me, the excitement of making the image was being able to see something that even a few years ago, we never thought we'd be able to look and say "that is a black hole" and to then be able to have a persistent stare into it and to be able to look at what goes on – how matter circulates, how these jets are formed, what happens near the horizon – it'll probe theory; it tells us about astrophysics. These jets can prevent stars from being formed in certain regions. They shape the universe in some ways.

ND: The EHT is an array of telescopes across the globe, which is reminiscent of the Magnetic Union and LIGO. What at the challenges and benefits of working on such a large-scale collaborative effort? Is international scientific collaboration the way of the future for physics and astronomy?

PG: Well, I would separate those two. In physics, there's a long tradition of big collaborations, in the Higgs discovery and the big detectors, LINACs and ATLAS and so on that are part of that. Those are billion-dollar detectors. There are, I think, up to 5000 physicists on each of them with thousands of engineers, also, on the projects. They're enormous international collaborations. They're centred at one particular place, although they diffuse the analysis and in the building of these [detectors], there were parts that were assembled elsewhere. In particle physics, that kind of collaboration has been organized. They have constitutions, they have committee structures. It all kind of has been built in place, growing at a steady, maybe accelerating, pace for decades. But certainly, by the 1960's, you had big bubble chamber collaborations. I mean, 'big' by their standards – ten/twenty people on a collaboration. By 1983, the $W$ and the $Z$ particles were discovered. Those had 130/140 people on those experiments. Those were considered enormous at the time. It's been growing ever since.

In astronomy, this is not at all what has happened. Astronomy has had small group collaborations – 2, 3, 10, 20, 50 people. Even the EHT, ten years ago, had 40 people. Now, it's 207. And so, 207 is big by astronomical standards but small by particle physics standards. There's no one place that is the conventionally assumed centre, it's a network; that's different, too, from particle physics. There's nothing like CERN. It's not like there's a billion-dollar detector and it's located in this building at CERN. The other part of your question though, I think leads to an interesting question. Is this the future? Partly what the EHT does is to show that you can make a planet scale telescope. That opens up all sorts of possibilities as well. One of the great challenges of the EHT is that it's a new instrument, a new collaboration with new techniques of making images, looking at an object that has never been seen before. It was like making a camera where you had a new camera back, a new lens, and were testing it on an object that no one had ever seen before; that's not how you test a camera. You would start by taking the lens and put it on an old camera back and look at a test chart to see how well it did in colour, in being able to separate lines of increasing closeness, and ask: did it make distortions at the edges or the centre, all those sorts of questions. Then, when you knew you had a good lens, you try a new camera body attached to it, again on known sources, a known target. Only then would you take your proven lens on a proven camera back that's been tested on known sources and point it at something, you know, trying to take a picture of an insect under a microscope you've never used before. You wouldn't start [that way]. We had to build our ship at sea, so, it was hard.

PF: The EHT image wasn't a standard photograph. You had a highly complex network of radio telescopes that all had various interconnecting parts, whose communication was mediated by things like atomic clocks. With all that sophistication, it wasn't exactly point and shoot. Why do we think that the end result is a photograph of a black hole rather than some sort of a pictorial representation of some more abstract collection of data points? I think this is somewhat related to some stuff you've been looking at before with scientific communities being split between looking at a concrete ontology of the world versus a more interconnected less focused image of how things work. How do you think this project sits in that conversation?

PG: So, let me answer in two parts. First, the question of what kind of image this is. The second is how it relates to the other work that I've done. That is also why I've gotten so involved in it, because it's everything I really care about, at least in my work. So, along with the announcement at 9:07 A.M. on April 10th came 6 published papers. The six published papers, which were released exactly the same time the images went up went up on these screens in Singapore, China, Japan, Santiago, Brussels, and Washington DC, represented a kind of story that goes from beginning to end.

There's an introduction with paper one. Paper two is about the instrument itself, that is to say, this network and how it works and what the technical demands and possibilities for the different radio telescopes are. This was new because it was pushing the edge of what was possible. In order to get the resolution that you need, you essentially have to be able to image the dimple on a golf ball that was located in Mexico City from [Toronto]. It's 1000, 2000 times better resolution than the space telescope has. It's really pushing the edge. It means new electronics, new means of data storage, the means of correlating that data. It means also that all the electronics in each of these telescopes has to be upgraded in order to take that information to be time stamped by an atomic clock. You've got to move this stuff back, correlate it, and line these up to a billionth of a second. Only then you begin to construct your data set.

Paper three is the data, which had to be really understood and sorting out if a telescope was really working or if it's at an angle where it's not reliable, taking into account what's happening with weather. And then [there are] new forms of combining the data that would be less vulnerable, for instance, to a fog bank rolling over your telescope at just the wrong time. Once the data was secured in this way, then there's the image group, in which I was involved. That group really worked for a very long time to develop new methods to analyze this data. Then, to compare the image to theory, you can't just look at Einstein's equations and say, "here's the prediction." You actually have to make an image out of the Einstein equations and this circulating billion-degree plasma that's moving around it. So, there are people that do what's called GRMHD, General-Relativistic Magnetohydrodynamics, which is the fluid flow of hot gas that's turned into a plasma being dragged around this complicated, very curved space that's also being dragged if the black hole is spinning. It's a very complicated system. That's what paper five was about. Paper six was theory. So, you have this whole arc of from the instrument design, through data, imaging, modelling, and theory. In a way, the argument for the validity of the image, that it holds good, is constituted by that whole structure. Then, within, say, the imaging group, there are all sorts of tests that that you have to do.

We first saw an image of the shadow with its ring back in June/July of 2018. The whole of the time since has been just all all-out effort in the image group to make sure that that was a reliable image. We'd test what would happen if you took away one of the radio telescopes. Could it be that one of them was doing something funny? What if you had different image-making computer programs? What if you set the settings, so to speak, on your camera differently? There are all these parameters that have to be put in in order to take this sparse and noisy data and make it into something that you could actually see. Then you can ask the question; in what sense is this like seeing?

My own sense is that seeing is something that isn't just eyes. It's also glasses or contacts or laser surgery or squinting. We do all sorts of things to be able to see better – microscopes and optical telescopes. You could say, "well, the radiation we're looking at is not in the visible spectrum." It's in the infrared at the edge of the microwave because that's the radiation that can get through all the interstellar junk. It does well through the atmosphere. It turns out to be a very good frequency; it's the smallest we can do. What if you wear night vision goggles? If you go out at night and you see a fox walking down your street, you wouldn't say, "that's probably not real," you would still say, "I saw a fox." We are even used to using transformations in frequency to be able to make it visible to us, that's what night vision goggles do. You might say, "there's all this processing that takes place to be able to get it." Well, when you take your phone out, there's a huge amount of processing. It's very complicated. I film sometimes on my iPhone and when I do, I use a program where you set the frame-rate, you set the stabilization, you set the contrast and the focus. You're doing a huge amount of electronic processing and all digital cameras do that. The charge-coupled device that changes the photons into signals has an immense amount of code to function. So, I don't think it's processing that makes it not seeing. If I show you a video of your friend blowing out candles on their birthday, you wouldn't say, "I didn't really see that" or "it wasn't real," which is not to say that we can't be mistaken or there can't be illusions or errors or technical things that produce artifacts. But, it means that we understand seeing to be expansive in this way. You might say, "we're not really seeing the black hole. We're seeing the shadow or the silhouette of a black hole." Well, if you're walking down the beach and you see your uncle running towards you silhouetted against the back, you don't say, "I didn't see my uncle. I saw the silhouette of my uncle." I think that when there are enough distinctive features of something, we do extend the notion of seeing. That's the long answer to that question.

ND: That's very interesting. So, obviously, there are many scientific benefits to imaging a black hole. What do you think is the cultural significance of this achievement?

PG: Black holes have a grip on us. They carry several notions that I think intrigue people. One of them is that there's a kind of portal that's one way. You can check in but not check out. That, I think, is intriguing on many levels, including that it dovetails with the idea that the life cycle of a star, which is born out of gas that's gravitationally attracted. The stars get older and become red giants and, if they're big enough, they collapse into a neutron star. If the neutron star is a couple of times the mass of the Sun, it will collapse into a black hole. It represents, in a sense, the death or end of the life cycle of a star. Astronomers talk that way, not just in popular [discussion], but that's part of what you learn when you learn astrophysics. The idea that the black hole is, in a sense, the end state of the life-cycle of these objects dovetails with already the association with death, these dark stars. Like death, it passes in one direction. I think people feel the metaphorical connection and fascination with these objects. And then, the strangeness of the singularity at the centre, which is where space and time and all of our theories break down, I think that's intriguing to people.

I think even the name [is intriguing]. At one point, they were called collapsars. "Collapsars" are not as riveting as "black holes"; this absence, this strangeness, this thing like the one we observed in M87 is six and a half billion times the mass of the Sun but there's nothing physical in the sense of material there. You could cross over the horizon and you wouldn't feel anything different at the moment you crossed over. You wouldn't be able to get out and things would end badly, but at the moment of passing through the horizon, it's not like you hit something. You can hit the surface of a neutron star (and that would not be good) but that's not true for a black hole. So, this idea of something that's incredibly massive and can break up stars and split them into planet sized pieces or it can accelerate a start to near the speed of light, like in a slingshot. It's incredibly powerful that way but it isn't a material object in the sense that we usually understand. It's just a bending of space. That's amazing to people, I think. And so, people imagine it could be a portal to another part of the universe or another universe. It provokes our imagination in all sorts of ways.

ND: Definitely! What's next for the BHI and how can history and philosophy of science students get involved?

PG: The BHI will continue to make these connections, like between theory and the EHT. I think that will expand. The EHT is going to be extending the network to be able to get more resolution. Eventually, the hope is that we can actually put telescopes in outer space. It's hard to do that but I think we will. Again, that will extend the possibility of making more detailed pictures. We want to make movies of what's happening around the black hole in M87. We want to finish the work that we've started with M87 and look at the Sagittarius A*, which is our very own black hole at the centre of the Milky Way, which is "only" 4.1 million solar masses, instead of 6.5 billion. But it's the same size in the sky, so it's totally within our reach. It's a little more complicated in some ways because it's smaller and things happen faster around it. It's more complicated but it's not a long-term project. I'm very optimistic that we'll be able to approach that. We do have a group in the history and philosophy of black holes. If students are interested in that, we do have graduate students, and as I said, there's an undergraduate, who comes sometimes, from philosophy who participates. There's a graduate student writing her thesis on the EHT in history of science, from Cambridge, England, but she has come to our Cambridge periodically to visit. So, there are different ways that people have gotten involved. If people are interested, they should write me!

# 6

# *Almagest* in Conversation with Ian Hacking

**C. Barker, N. D'Or, G. Foster, P. Fraser, and A. Sarwar**
*Almagest Editorial Team*

Professor Ian Hacking is one of the most prolific philosophers of our time, making many significant contributions to a vast array of topics ranging from initiating the study of philosophy of statistics, to defending and refining our understanding of scientific realism, to clarifying social constructionism. Known for his accessible and jovial writing, Hacking combines wit with a commitment to detail to bring us no less than twelve books on the history and philosophy of science, including *The Taming of Chance*, *Representing and Intervening*, and *The Social Construction of What*.

Born in Vancouver, B.C., Hacking completed his first undergraduate degree at the University of British Columbia before attending the University of Cambridge, where he completed his second B.A., his M.A., and his Ph.D. under the supervision of Casimir Lewy, a former student of Ludwig Wittgenstein. In recognition of his unmatched contributions to the history and philosophy of science, Hacking was awarded the Killam Prize in 2002, the Holberg Prize in 2009, and made a Companion of the Order of Canada in 2004, among many other prestigious accolades.

We met with Hacking at his home in Toronto earlier this year. We'd like to thank Professor Cheryl Misak and Professor Joseph Berkovitz for helping to facilitate the occasion, as well as Professor Ian Hacking for his gracious hospitality and electrifying conversation.

*The following interview has been edited for length and clarity.*

ALMAGEST EDITORS (AE): How did you get involved with the history and philosophy of science?

IAN HACKING (IH): Well, I'm really a philosopher. I'm not a historian. I'm just a philosopher and because of that, I'm interested in everything.

AE: You started in physics, right? How did you become interested in philosophy?

IH: It was entirely fortuitous. That is, I did a math and physics degree at UBC and then applied to Cambridge. There I did what was called a moral science degree, at that time; a philosophy degree. It was a complete switch.

AE: What areas did you study at Cambridge at that time?

IH: Mostly logic. It was a very narrow program. By any American standard, it would be crazy because I did nothing except read Frege and Russell and Austin.

AE: In the early 20th century, physics and philosophy were very interconnected. Since then, we've seen the sciences have pushed themselves away from philosophy, with several prominent physicists expressing negative attitudes towards philosophy of science. How do you think philosophy and science can work together in a constructive manner?

IH: My attitude [towards] philosophy is that it's most importantly the theory of knowledge. Physics is one branch of knowledge, but it's only one branch of knowledge. That's sort of the layout of my life view, I guess, though I didn't start with that view. I don't think that physics is so essential to philosophy; I think there are lots of philosophical problems about physics but it's not a core feature of philosophy. It's true that a number of physicists have been philosophers, but not very many, and certainly not many experimental physicists, whereas I think experimental physics is what's most interesting.

[However,] I think it's most important, if you're going to talk about a science, to read the science, to know the science, practice it, learn it. You can't talk about physics without knowing a lot of physics. You can't talk about biology without knowing a lot of biology. I think that many philosophers just talk off the top of their heads without ever being able to read even a journal like Nature. I do think it is important to be able to read Science and Nature. Even if you don't go beyond those magazines, they still have an enormous amount of information and learning.

AE: Do you think that it would be beneficial for experimental physicists to understand things like intervention and how experimentation is philosophically important to understanding science as a whole?

IH: Well, experiments are essential to [science]; that's how we find out about the world, by doing experiments and having experience. The truth is, I don't think there's any special relation between philosophy and physics except that there are some interesting problems in physics, which are philosophical; problems about space and time for instance.

AE: Right now, the biggest interaction between philosophers and scientists seems to be in bioethics through topics like genetic enhancement. There's a lot of discussion about the ethical implications of scientific investigation into genetically enhancing humans before birth, and a lot of prominent scientists are saying that we should halt investigations into genetic enhancement until we better understand exactly what we're working with. Do you think that we should make sure that we're paying attention to the possible ethical consequences of what we're doing within science, even if that limits how fast science progresses?

IH: That's very important. I think we really don't know what are going to be the consequences of changing the genetic structure of human beings. We just don't know. It may be trivial, it may simply make us healthier, it may make us more smart, it may make us crazy, and we haven't any idea what human beings will do with genetic changes. Very often I'm just left with ignorance and don't have any views at all.

AE: What books are you reading at the moment?

IH: I'm reading rubbish mostly.

AE: Any recommendations?

IH: No, none at all.

AE: Are there any books that you would recommend not reading?

IH: All too many. Most of these [Hacking gestures towards the bookshelf behind him].

AE: Is there a single philosophy book that fundamentally changed your outlook on everything?

IH: No. What changed my outlook was going to Cambridge, and there I was immersed in a very specialized attitude towards philosophy and it was just what I needed. Not to become specialized; most of my fellow students there have become specialists, whereas I've become a generalist. But I think that being a specialist is a necessary precondition to becoming a generalist because if you're going to talk about knowledge you have to know quite a lot.

AE: In Cambridge, did you find something about the pedagogy to be what changed your outlook?

IH: [It was] only good for me; [it was] bad for many people. For me it was perfect, but I had already done a math and physics degree so I had competence in other fields. Not much, but a little bit.

AE: If someone has no knowledge and they want to learn something about everything, where do they start?

IH: Start with something. You may want to learn a lot about life, so learn some biology.

AE: I guess with the internet, we have so much information at our fingertips you become overwhelmed with the amount that you can learn.

IH: Yes, I can feel a bit overwhelmed now. What I try to do is to read science journals; don't try to read pop science. That's much harder, but also much more important, because you get a sense of how people are thinking about something.

AE: What do you think about pop philosophy going on right now, with individuals who were not trained in philosophy being hailed by the general populous as being philosophers, even though a lot of the their work may have been long ago? Do you think that pop philosophy is something we should be wary of?

IH: It depends on the person. Some people like that, others don't. I don't have any strong views about what's right. And I certainly don't have any views about how its right to do philosophy, except it's important to have an open mind. But even there, I think Popper was the most important philosopher of the early 20th century but his mind was totally closed.

AE: Why do you think he was the most important?

IH: Oh, just because he had the best ideas. His ideas of conjecture and refutation have now been absorbed in all the sciences. And that was his idea. The German philosophers of science who were important at the time, the Vienna Circle and so forth, do not seem to me to have made much contribution. Whereas Popper seems to me to have gone to the heart of many questions.

AE: With Popper's falsificationism, in philosophy of inductive logic, there's been a reaction against the logical interpretation of probability that Popper developed. Do you think that the sciences should look towards a more subjectivist account of probability, like Bayesian subjectivism?

IH: The Bayesian subjectivist has a very powerful response to the problem of induction. He tells us how we learn from experience, namely, by Bayesian calculations. And that's an important contribution. I think that Popper was very very narrow minded in his view of what probability is, and with his propensity view. Propensities are one interpretation of probability which suit certain questions, and which don't suit other questions. If we're talking about learning from experience, Bayes' rule is essential.

AE: You say that you think Popper is the most important philosopher of the 20th century, do you think his importance is based on the impact of his work in the real world, rather than just the merit of his ideas?

IH: I would only say philosopher of science. He had an important idea about refutation. The verification principle, which was boasted by Carnap and Schlick and so forth, is a theory of meaning which I don't think is correct. Whereas the falsification principle is not a theory of meaning, it's just a theory of what's the point.

AE: In *Representing and Intervening*, you mention that you don't think a theory of meaning should have the prominent place that it once had within the philosophy of science, and Popper also agreed with that view. Why do you think there is no place for a theory of meaning in the philosophy of science?

IH: Because I don't think it teaches one very much about science. It was an important tool for the Vienna Circle, and we learned from that, but after all, it's now passed. And I don't think we can learn very much more from the verification principle, which is a principle about meaning.

AE: What questions should philosophers of science be focusing on right now?

IH: Well first of all, I should have said this at the beginning: I am out of date. I got sick and I retired, and I do not keep up with the journals, so I'm not qualified to speak about contemporary philosophy. But I am interested in those ethical questions about what to do. You mentioned genetic engineering; I have no very strong clear ideas about that. I think one has to look at the details of what is being done. And I think that genetic engineering is one of the most important discoveries going on right now, and we'll find out more and more about it. And we'll find out more and more about what we can do, and decide, some of us, what we should do.

AE: Do you think the way we interact with technology will shape what kinds of major questions philosophers ask?

IH: Yes it will. But it depends quite a lot on one's person. I continually branched off into new fields, often the first philosopher to do so. And when I did the stuff on statistics, nobody else had worked on statistics, they just wrote about probability and didn't look at what people did with it, namely practice statistics. And I was the person who really started that. Now statistics is a real philosophical problem. My fault.

AE: There is this cliche that correlation doesn't imply causation. It seems that causation must imply correlation. What are your thoughts about how we can learn about causation without "stepping into the world itself."

IH: I don't think all causation is a single thing. I think that there are lots of things that we call causes and they are very different in nature and that some causes which are important in some views are not important in others. I think that human choice is a cause in many things we do, but it's not important in understanding time. And conversely, it's not clear what causal problems arise in connection with time, except trivial ones like 'can causes work backwards' or whatever. If we think that $A$ is a cause of $B$, then $B$ should be changed if we do something different to $A$. A very practical point, but essential. And with correlations, they may remain constant, but only because there's an accident. That is, if there's a correlation between a child's intelligence and how it's brought up, there's so many changes we could make in the child's upbringing that correlate to changes in intelligence, that we can't really say that upbringing is a cause of intelligence, or lack of it, in the child.

AE: In your book *The Social Construction of What*, you indicated that the way in which our languages are structured can have an affect on the way that we interpret and view the world. Do you think that the language of a community can affect how they interpret their world on an ontological level?

IH: Probably yes, but I'd have to know what sort of changing you would make. But I suppose the way we act is highly influenced by our expectations incorporated in language.

AE: Do you think that language is a limiting factor within philosophy. Like late Wittgenstein liked to say, do you think language is sort of the first philosophy?

IH: It's not the first philosophy, but it might be necessary to always come back to think of how language influences our thinking, but that's another case of causation.

AE: Within the realism debate, the criteria of what is real often depends on the correspondence theory of truth. But if we have a different set of criteria that does not rely just on reference, what would the realism debate and its implications look like?

IH: I don't like the idea of a universal theory of truth, as there are different types of truth; for example, truths about emotions that are not founded on reference. So, I wouldn't be able to provide a general answer, since I don't think there is one kind of truth.

AE: So, you would likely believe in a pluralism of truth, which likely depends on the discipline in which one is operating?

IH: Yeah.

AE: Interesting. Do you believe in a pluralism of a similar kind with respect to ethics?

IH: I am an absolutist when it comes to ethics.

AE: We can say that truth may not have much to do with reference in some cases, but it does in other cases, but it may not naturally be the case that one could put together a referential theory of truth with something that is non-referential. How might one reconcile this discrepancy?

IH: I don't feel the need for unity: there are ever so many topics about human emotions that I cannot discuss using concepts from physics, etc., yet I still think I can understand human emotions fairly well without the need to rely on anything else. Perhaps, if we were different kinds of animals, then we'd be telling different stories about our emotions.

AE: With this lack of unity, are you saying that talking about emotions is working in a different logical world than talking about about physics?

It seems that there may be an intersection that unifies the two or perhaps a gradient that allows one to go from physics to emotions and vice versa.

IH: Well, I don't find it necessary to think in terms of reference at all most of the time. If I were describing the things in this room, I would be referring to different individuals, but if I were describing the air currents in this room, I don't think I would be describing objects at all. I'd be describing motions and whatever you detect with anemometers.

AE: In light of this pluralism of truth that you've described, given two extreme ways of doing philosophy of science, one very normative and prescriptive, such as that of Carnap and Hempel, and alternatively that of Feyerabend's 'anything goes' attitude, in this spectrum of beliefs about how scientists ought to proceed, where would you place yourself?

IH: Well, for people like Hempel and Carnap, remember where they grew up. They grew up in Nazi Germany or Nazi influenced countries and it was very important for them to insist that there is such a thing as truth because the Nazis either uttered a lot of falsehoods, or didn't even believe that there was anything 'true' there, and so their philosophy is partly just as opposition to the Nazis; an intellectual opposition. Many people think the Vienna circle were all Jewish, whereas in fact hardly any were, they were just intellectuals, which includes one or two Jews like Otto Neurath, but in general, they were not Jewish. But they were utterly resistant to the evils of the ethics which was being propounded around them.

AE: In a time where fake news and so forth are so common, do you think that we need to adapt our philosophy today to combat this extreme of relativism?

IH: Well I am certainly opposed to the current vogue of Trump's talk. I think that has to be opposed, and fought, and if it can be done with a theory of truth, good for it.

AE: What do you think about the division that definitely still exists today in academia between analytic philosophy and continental philosophy?

IH: Well it's not a distinction I've ever found useful, and I was probably the first English speaking philosopher to take Foucault seriously. That doesn't mean that I am a structuralist or that I am a continental philosopher, just that I think Foucault has a lot of important things to say, and no one else was saying at the time, and I don't have any tension between continental and English philosophy.

AE: If you were to create a syllabus for a philosophy of science class today, what would you find to be central texts?

IH: I would find Kuhn very helpful and essential. He really changed the way in which people talked.

AE: If you could study philosophy for the next hundred years, which questions would you try to answer next?

IH: I think I would start by learning a lot of biology, and start asking questions about life. That's what I would do. But I'm starting fresh. I find biology texts quite difficult. I find that most high school students know more biology than I do, which is right.

AE: Do you have any advice for history and philosophy of science students?

IH: Just read as much as you can! And find out, as long as you're interested, what you can understand. Find people whom you can understand and who teach you something so that when you finish reading a piece, you feel you can say what you learned from it.

AE: In philosophy, there's a tradition of speaking very analytically and anti-historical. Do you think that history has an important place in understanding the philosophy of science

IH: Yes. I mean, I have written a lot about probability, but always in the back of my mind are historical questions about when certain problems became apparent and when they disappeared.

AE: Do you have any other thoughts that you would like to share with the student body?

IH: Learn as much as you can. That's all I've tried to do.

AE: Is there a particular aspect of your career that you found most fulfilling?

IH: Well I think that being fascinated by ideas has always moved me, and I'm really moved by ideas, even if they're bad ideas, they still have a movement to them.

AE: Thank you so much for meeting with us!

IH: Any time.