ACTA PHILOSOPHICA FENNICA

Vol. 98

Beauty, Truth, and Justice

Philosophical Essays on Culture, Science, and Society

ILKKA NIINILUOTO

HELSINKI 2022

ISBN 978–951–9264–96–7
ISSN 0355-1792
Hansaprint Oy
Helsinki 2022

Table of Contents

Preface

In the spring of 1972, precisely fifty years ago, I was preparing my doctoral dissertation on inductive logic as a visiting scholar at Stanford University with Professor Jaakko Hintikka. This experience largely determined by profile and academic career as a logician and philosopher of science – with numerous articles in philosophical journals and main international works *Theoretical Concepts and Hypothetico-Inductive Inference* (with Raimo Tuomela, 1973), *Is Science Progressive?* (1984), *Truthlikeness* (1987), *Critical Scientific Realism* (1999), and *Truth-Seeing by Abduction* (2018).

In my basic studies in mathematics, with specialization in probability theory, I was more fascinated by the *concept* of probability than by the calculus of probabilities. My move from mathematics to philosophy was fueled by the wish to find a field which has substantial content instead of abstract formalism. Here Hintikka was an excellent methodological guide: as a master of formalization he always had in mind some philosophical applications. Thus, my doctoral work on inductive logic was hoped to give tools for defending the realist interpretation of scientific theories, and the later long-time study of truthlikeness has given a basis for a realist theory of scientific progress as increasing verisimilitude.

My other role model in philosophy was Academician Georg Henrik von Wright, a prominent logician and analytic philosopher, who also wrote eloquent essays on human rights and the problems of scientific-technological civilization. In his 1990 article "What Philosophy is for Me?" von Wright confessed that he has not been able to pin down the relations of his literary essays to his "scientific personality", but the efforts to understand one's own times grow "from an inner compulsion". A related division was made by Richard Rorty in his *Philosophy and the Mirror of Nature* (1980) between Kantian *systematic*

philosophy and *edifying philosophy* which substitutes "*Bildung* (education, self-formation)" for knowledge as the goal of thinking and aims at continuing "the conversation of mankind". While von Wright wished to do both, Rorty urged that systematic philosophy should be abandoned and replaced by edifying philosophy.

In my inaugural lecture "How is Philosophy Possible as a Science?" for the chair of Theoretical Philosophy at the University of Helsinki on April 8, 1981 (see *Is Science Progressive?*, 1984, 10-17), I summarized the method of philosophy in three Socratic steps: problematization, explication, and argumentation – and these stages can be practiced both in basic and applied philosophy. Thus, philosophy is not only conceptual analysis but also the study of life stances and world views. Therefore, I argued, there is no need to draw a sharp distinction between systematic and edifying philosophy in Rorty's sense: these two interrelated aspects of philosophical activity at best support each other. This kind of program I have attempted to follow in my Finnish monographs and collections of essays: scientific world view (1984), information society (1989), culture and emergent materialism (1990), reason and values (1994), loving the truth (2003), *Bildung* and university (2011), good life (2015), and technology (2020). Further, in my tasks as the Vice-Rector, Rector and Chancellor of the University of Helsinki in 1998-2013 my starting point was that science policy and university leadership can be understood as applied philosophy of science and ethics.

In this collection I have selected 21 papers which illuminate the interaction between systematic and edifying philosophy. I take up the challenge again from Rorty whose article "Dewey's Metaphysics" (in *New Studies in the Philosophy of John Dewey*, 1977, 45-77) describes Dewey's hope for a culture which "no longer makes the [Kantian] distinctions between Truth, Goodness, and Beauty" and whose "chief enemy" is "the notion of Truth as accuracy of representation". In doing the opposite, I divide the chapters into three parts: *Culture* (with representation and beauty as the key concepts), *Science* (with rationality and truth as the central notions), and *Society* (with justice and quality of life as the main concepts). To study these topics, a sufficient amount of conceptual analysis and conceptual history is needed. In particular, the reader will find in several

chapters applications of Karl Popper's ontology of "three worlds" and G. H. von Wright's concept of a technological norm. It is only from the vantage point of this systematic work that I am able to address edifying themes: 1. temporality of culture, 2. the gap between two cultures, 3. information aesthetics, 4. language as representation, 5. philosophy of film, 6. virtual reality, 7. verisimilitude and fallibilism, 8. the rationality of rationality, 9. critique of creationism and ID-theories, 10. human sciences, 11. futures studies, 12. design science, 13. the aims of a university, 14. the identity of Europe, 15. progress and the Enlightenment, 16. the post-truth era, 17. sustainable development, 18. good life, 19. welfare society, 20. technology policy, and 21. democracy. On the whole, the topics cover important aspects of semiotics, philosophy of culture, philosophy of time, philosophical aesthetics, ontology, epistemology, philosophy of language, philosophy of mind, philosophy of basic and applied science, history of philosophy, philosophy of religion, philosophy of education, environmental ethics, philosophy of war, philosophy of law, philosophy of technology, social philosophy, and political philosophy.

Some of the chapters are based on unpublished manuscripts, articles in Finnish, or updated papers in not very well accessible collections and conference proceedings. They can be read separately, but taken together they inform and support each other. As the articles have been written and presented for different audiences, both academic and popular, there is some variation regarding their level of technicality. But I hope they all exhibit my passion for clarity in philosophy. In collecting and revising them for this issue of *Acta Philosophica Fennica*, during the third spring of the corona pandemic and the third month of Russia's invasion of Ukraine, I have benefited from my position as professor emeritus in the Department of Philosophy, History, and Art Studies at the University of Helsinki, but my thanks extend to the original sources of the chapters during the last decades. As always, special thanks to my family for patience and encouragement.

Helsinki, Finland, May 15, 2022 *Ilkka Niiniluoto*

Culture

Chapter 1: Time and Culture

Following Cicero, culture is literally the "cultivation of spirit" (*cultura animi*), creative activity directed towards producing ideas and philosophies, science and art, pictures and literature, manners and social institutions. Culture is in every respect bound up with time, for it is produced, nurtured, and passed on in the context of succeeding generations. In this chapter, some interesting aspects of this inevitable temporality of culture are charted.

The Dimensions of Time

One major issue in the philosophy of time, on which scholars have consistently failed to agree, is in what respect time is real. Time has a physical dimension which implies a certain order of events and changes in the world regardless of the human beings who perceive these events. This creates a framework of *physical time* (or space-time) in which we can speak objectively of the history of the universe over the 15 billion years.[1] But time has also a subjective dimension, referring to our way of experiencing events and dividing them into past, present, and future. Within our personal *time consciousness*, our present experiences are intertwined with memories of our past history and anticipations of future events. Instead of putting these two views of time against each other, both of these dimensions of time can be regarded as real: physical time as a constituent of the mind-independent nature and experienced time as an element of our mental life.[2]

---

[1] See Hawking (1988).

[2] The reality of time was denied by McTaggart. Objective or physical time has been discussed by Aristotle, Newton, Leibniz, Einstein, Russell, and Hawking. Perceived or experienced time is a topic for Kant, James, Bergson, Husserl, and Heidegger. Cf. the essays in Gale (1978). In Popper's

The threefold division of past, present, and future is accompanied with three types of human capacities and attitudes. The first Finnish university, the Royal Academy, was established in Turku in 1640. Two years later, its first professor of practical philosophy (ethics, politics, and history) Mikael Wexionius published the academic dissertation *De prudentia*, which also was the first book printed in Finland. Following Cicero's *De inventione* (84 BCE), Wexionius (1642) distinguished three components of the virtue of prudence which are related to temporal human attitudes: memory (*memoria*) concerning the past, understanding (*intelligentia*) concerning the present, and providence (*providentia*) concerning the future.

Typical backward-looking emotions, when we remember our past, include pride and remorse, relief and guilt. Our life in the present mood is colored by pain and pleasure, grief and joy, happiness and sorrow, hate and love. Forward-looking emotions include fear and hope.

Memory and the Past

A debated special question in the philosophy of time concerns the extent to which the past is real.[3] Where the future is concerned, we can more or less agree that it is not yet real but opens itself up in front of us as a set of possible alternative courses of events. The doctrine of *presentism* maintains that the only things that are real are those belonging to the present moment of time. This view is a species of actualism which claims that only actually "here and now" existing things are real, so that past events at best *were* real but are not anymore. A counter-argument to presentism is that since the past has already happened it is now objectively defined or fixed: although we can still study and interpret past events, we can no longer influence their course. Hence, they satisfy Charles S. Peirce's "mediaeval definition": those things are real "whose characters are independent of what anybody may think them to be".[4] Also, the past is constantly affecting our own lives in in innumerable ways. Although we may no longer live in the past, or

(1972) terms, physical time belongs to world 1 and experienced time to world 2.

[3] For the reality of the past, against antirealist views, see Niiniluoto (2014).

[4] See Peirce's *Collected Papers* 5.405, 5.430.

even nostalgically wish to return to it, the past remains with us in the form of historical traces it has causally left on us and the memories that we have on it. The recognition of this fact was regarded as one of the marks of civilization by the Finnish philosopher Eino Kaila. This criterion of causation shows that past events are real, even though they are not present here and now. Further, the only way of reaching knowledge about the past is by arguing from present effects to their causes, i.e. by the mode of inference that Peirce called *abduction* or *retroduction*.[5]

Even though the past is immutable as such, historical research can seek answers to many kinds of questions (why? how possible? who? when? where? what?) which afford new interpretations and insights about the development of human culture. In this way, we may gain understanding of our origins and the still operative causal factors, so that our understanding of the present situation is improved.

Cultural Evolution

Culture is a product of human creative practices. Paradigm examples are the design of artefacts (tools, instruments), the construction of works of art (paintings, sculptures, poems, songs, symphonies), and the discovery of new scientific concepts and theories. Typical loci for such activities are technological laboratories, artistic studios, research institutes, and university departments. For everyday life, culture may simply mean the establishment of habits and customs. Sometimes the cultural creations are radically novel, but often they emerge as gradual modifications of earlier products. After their creation, the cycle continues with the preservation and diffusion of cultural entities from one place to another or from one generation to the next. The whole process is usually called *cultural evolution*.

Popper (1972) compared the growth of scientific knowledge to the Darwinian biological evolution, where conjectural theories are tested and refuted by experimental tests. But the intentional creation of hypotheses and theories in science cannot be compared to blind mutations, and the transmission knowledge by learning involves inheritance of acquired characteristics, so

[5] See Niiniluoto (2018).

that cultural evolution is more Lamarckian than Darwinian. Still, the historical development of successive oral or written versions of cultural items like poems and tales can be represented by tree-like "stemmas" which resemble the evolutionary trees for biological species, which means that almost the same kinds of abductive inference schemas can be used in the historical studies in the humanities and biology.[6] Other theories of cultural evolution have compared transmitted cultural units or "memes" to genes in biology.[7]

Against Cultural Erosion

Just as a person's memory may stretch back to the very early years of his or her life, so the collective memory of the human race can be traced back as a chain extending far into our prehistory, as elements of tradition can be handed down from one generation to the next in the form of stories, ballads, and tales preserved in the minds of individuals. The vehicle for this act of cultural transmission is the *spoken language* and its manifestation in the oral traditions of a given community.

The key to success in the preservation of human culture has nevertheless been the acquisition of an ability to record cultural achievements and pass them on in the form of publicly observable signs. Take, for example, the early tools, the art of using which was passed down from parents to children or from craftsmen to apprentices, or the pictures that formed the basis for the early analogical mode of communicating knowledge. One revolutionary step, which is normally taken as marking the beginning of historical time, was the invention of *writing*, as this enabled the products of culture to be recorded and transmitted in a manner that was able to overcome the restrictions of time, place, and human memory. The culture created by human beings and the knowledge generated by human inquiry could be noted down for the posterity as marks on scrolls of papyrus or parchment, reproduced on paper by means of a printing press, and eventually distributed

[6] See Niiniluoto (2018), 66-70.

[7] This term was introduced by Dawkins (1976), who suggests that memes are replicated by imitation. Meme theory has become popular with the development of social media, where messages can be easily transmitted by the share button.

throughout society in the form of printed books. Further stages in the same progression are represented by the digital electronic publications of the late twentieth century, in which the medium employed consists of the circuits and silicon chips of computer memories.

The most famous center of classical learning, which attempted to collect all wisdom of the Antiquity in its library, was the Mouseion established in Alexandria, Egypt, in the 3rd century BCE.[8] The museums, archives, and libraries of our modern society are often referred to as "human memory organizations". This is a propitious way of thinking of them, as their purpose is indeed to document, store, and exhibit cultural achievements from times past. It is true that a museum which places ancient objects on display for people to look at does not usually allow visitors to handle them, and an archive may well make its documents available to readers on microfilm or in digital form rather than handing over the original historical papers, but the purpose remains the same. Similar restrictions also apply to libraries, as the older books in their possession are liable to suffer wear and tear, or simply to degenerate, in the course of time. Thus, we must recognize that while objects of cultural significance serve as traces of the past in our world of the present they are under a constant threat of "erosion" – wear, degeneration, and destruction.

Using the concepts created by Karl Popper, we may say that the products of culture typically belong to three worlds.[9] *World 1* is that of physical objects and processes, such as a book, which has a construction, form, chemical composition, weight, strength, smell, and color of its own. This level of reality also encompasses the letters and pictures printed in a book. For bibliophiles, ancient or old original copies of a book can in themselves be objects of admiration, things to be collected passionately for their own sake. These physical objects are nevertheless subject to erosion by the natural laws of change, so that they can decline in strength in the course of time, fade, turn yellow, develop mold, or even be destroyed by fire etc. Thus, the conservation and restoration of books, films,

---

[8] The buildings of the Mouseion were destroyed in 272 CE. The new Bibliotheca Alexandrina was opened in 2002.

[9] See Popper (1972), Niiniluoto (2006).

audiotapes, and other forms of documentation is one important branch of the services that a library can provide.

World 2 is composed of the subjective contents of human minds. A book as a mentally existing object in world 2 lives in the memories and experiences of its individual readers. Erosion applies here in the sense that its content can be subject to misunderstandings or even be forgotten in time. Libraries nowadays frequently have enormous stocks of material that no one remembers any longer or thinks of looking for. Universities and their libraries can try to counteract this process by providing teaching and demonstrations in the history of learning and ideas, placing the treasures that they have in their possession on display, publishing books and articles about them and developing efficient, comprehensive search systems.

World 3 comprises the public human constructions in culture and society – such as artefacts, works of art, numbers, languages, scientific theories, customs, norms, and social institutions. As a World 3 entity, a book is not merely a piece of the physical world 1 but it has a number of cultural properties and relations (e.g. position in the literary tradition, author and readers, price in the market). Most importantly, the essential property of books as cultural products is the information contained in them, the intellectual property rights over which belong to the author in each case. Just as the printing press marked a major step forward in its time, because it enabled the production of large numbers of copies and successive editions of the same books, so we are able nowadays to store and publish the same work in numerous forms: as a duplicated manuscript, a printed book, a CD-ROM, or an electronic book on www pages. This means that objects of culture can live more independent lives than ever before, as "information packages" in a virtual reality (cf. Chapter 6 below). Scientific knowledge in World 3 is also permanently changing when its tentatively accepted results are updated by new evidence and revised in the light of criticism and new scholarship, i.e. the border between the history of learning and the frontiers of current research is constantly shifting.

These cultural entities in World 3 are also exposed to erosion, bound in time and susceptible to change, in the sense that they can still be eliminated from the sphere of human reality. This can happen if all the records of a work in world 1 are

destroyed and all the thoughts and memories arising in world 2 disappear. Likewise, a cultural object in World 3 can be lost if an electronic document can no longer be opened or cannot be reconstructed by new techniques, or if no one is able to interpret or understand the language in which it was stored.

Some cultural objects are regarded as so valuable that they are granted in their original form a special protected status in the "world cultural heritage". But most achievements of human creativity can be recorded in some form for transmission from one generation to the next, so that it is the task of libraries and other memory organizations to store them, make them available and place them on display, simply in order to ensure that they are not destroyed or forgotten. This will help to avoid the erosion that inevitably detracts our culture. Our libraries of the future will contain cultural documents in a wide variety of forms, both printed works and electronic material, and will thus be better equipped than ever for reserving the past in a living form and maintaining the continuity of our culture on the temporal dimension.

Seize the Moment

The philosophical doctrine of presentism represents a way of thinking that is particularly suited to *postmodernism*, which maintains that we are typically ready to seize upon the transient phenomena of the present without any special yearning for the past or worries about the future. The advice to enjoy the moment comes from the Roman poet Horace, whose recommendation *Carpe diem!* teaches us to take full advantage of life's opportunities whenever they present themselves.

If Horace's dictum is interpreted as a dictum against laziness, it is related to the reification of time as a resource or commodity. Benjamin Franklin expressed this idea in 1748 with the slogan "Time is money". A hundred year later Karl Marx presented his labor theory of value, where the value of an economic good is measured by the number of labor hours required to produce it. The Greek term for the leisure time of free men was *schole*, which is the root of the English "school". If you are not interested in work or education, life may be boring. But the modern age has developed multiple branches of culture industry (e.g. radio, cinema, television, video games),

which help you to entertain yourself and to "spend" and even "kill" time.

But one need not be a presentist or postmodernist to acknowledge the significance of our activities and efforts for contemporary culture. Here we may refer to two other ancient wisdoms. One of the three Delphian maxims, quoted by Socrates in several dialogues of Plato, was *Know thyself!* This principle changed the focus of Greek thinkers from the philosophy of nature to human affairs (such as ethics and politics). In the *New Testament*, Jesus commanded "Love your neighbor as yourself" (*Mark* 12: 31). This means that we should be concerned with the well-being of the living fellow human beings in the same way as we take care of ourselves.

These tasks can be fulfilled in many ways. One of them is the organization of our society by following the principles of justice (cf. Chapter 19 below). The second is the study of the human condition by means of human and social sciences. Such studies not only describe and interpret the present situation, but may also help to design better cultural and social environments (cf. Chapter 10 below). The third is often called the *Diagnosis of Our Time*, following the sociologist Karl Mannheim's wartime essays in 1943. The critical reflection of our present situation, revealing its weaknesses and pitfalls, is an activity well suited to philosophers and artists. In Finland, the diagnosis of our time has been practiced especially by Georg Henrik von Wright (1916-2003).[10] In particular, fiction in literature and cinema is a fine way of posing ethical questions which are immediately relevant to our choices and decisions today.

Provision for the Future

We can provide for the past by interpreting and studying history and by conserving and protecting cultural heritage in archives and museums. But past events have already happened – and we cannot change them anymore. The present is a precious moment which immediately swifts away to the past, so that our choices and decisions today will influence not only ourselves and our neighbors but also the world tomorrow. Therefore, our real responsibility and care – with emotions like

[10] See Österman & Vilkko (2016). See also Chapter 15 below.

hope and fear - concern the future. This insight is expressed by Cicero's and Wexionius's term *providentia*, which was the name of a Roman goddess with the ability to foresee and make provision for the future.[11] In the Christian tradition, providence came to mean God's care and protection of his creations (in Finnish, "sallimus" or "kaitselmus"), but its secular meaning expresses the human care and concern for the future (in Finnish, "huolenpito tulevasta").

Martin Heidegger, speaking about the scarcity or "Being-ahead-of-itself" of human life in his *Sein und Zeit* in 1927, proclaimed that the essence of Being is *Sorge*. This term is often translated as "care", which has an altruistic connotation like *providentia*. Indeed, in Heidegger's system there is no isolated Self or "I", so that *Sorge* does not mean "care for oneself". But still the ontological interpretation of Dasein as care is closely associated with the state-of-mind that Heidegger called *Angst* (anxiety).[12]

Natural science gives us opportunities to predict the future course of regular deterministic systems (like the solar system) on the basis of laws and initial conditions, but it is also known that some deterministic systems (like weather) are unstable or chaotic in the sense that their initial state cannot be known with sufficient precision to make long-term prediction possible. Some natural systems (like radioactive atoms) are indeterministic, as they are governed by probabilistic laws. Social systems (like economy) and socio-technical systems (like cities) can also be highly irregular, even though forecast of some their statistical features may be possible by quantitative methods.

Karl Popper, in his *The Poverty of Historicism* (1957), argued that the future course of history cannot be predicted by us, since it is strongly influenced by the growth of human knowledge – and genuine discoveries of tomorrow are not yet known today. Popper further warned that dogmatic utopian goals have often brought about political totalitarianism (cf.

---

[11] "Providence" is syntactically close to "prevision", which literally means "to foresee".

[12] See Heidegger (1978), 227, 237, 310. In Finnish, the translation of *Sorge* as "huoli" is related to both "huolenpito" (providence) and "huolestuminen" (being worried).

Chapter 21 below). He argued that social sciences should not seek large-scale unconditional prophecies on the basis of mistakenly assumed trends or laws of development, but instead should concentrate on "piecemeal social engineering" which removes "concrete wrongs" and "avoidable suffering".

The paradigm of *futures research* gives an answer to the Popperian worries (see also Chapter 11 below). This approach sees the future as a branching tree with alternative possibilities. The future is still open to some extent, as its features will depend on chance events and human choices. The task of the futurist is to investigate this tree as a whole. Its branches define alternative scenarios, which we may wish to realize or avoid. More precisely, the received view is that the futurist should (i) construct alternative possible futures, (ii) assess the probability of alternative futures, and (iii) evaluate the preferability or desirability of alternative futures.

The tasks (i) and (ii) reformulate the simple predictive model of futurology. Instead of simply announcing what *will* happen, they seriously engage in planning or designing the future by asking what *would* happen *if* we or others make strategic choices in key nodes of the future tree. The task (iii) is value-laden in a delicate way, since the futurist should be able to distinguish between the value commitments of themselves, their employers, and the actual and possible persons occupying positions in the future tree. It is not only a fact of life that such values may differ from each other - assuming the principles of liberal democracy, it is also desirable that people are allowed to make their own morally and politically relevant choices. The community of futurists should promote the humanistic ideals of individual freedom, social responsibility, justice, democracy, and sustainable development by respecting the norms of "critical discursive interactions", as described by Helen Longino in her work *The Fate of Knowledge* (2002). In this way, futures studies is able to serve *providentia*, the goal of Wexionius, in a double sense: to provide *against* the harms and dangers of various kinds of "wicked problems" and to provide *for* the human welfare and good life.

In Finland the Turku School of Economics has been the key location of future studies ever since the visionary professor Pentti Malaska (1934-2012) founded the Finnish Society for Futures Studies in 1980 and the Finland Futures Research Centre

(Tulevaisuuden tutkimuskeskus) in 1992.[13] Through the Finland Futures Academy FFA (Tulevaisuudentutkimuksen verkostoakatemia VVA), founded in 1998, the activities of research and teaching in this field have been spread to ten Finnish universities, and today futures studies has grown into an important academic discipline with a Professor and Master's and Doctor's degrees. Through the connection to the Millennium project the Finnish futurists have been able to make an important impact on the international community – and hopefully, in the footsteps of Mikael Wexionius, to provide for a better world.

Note. This chapter uses some passages from my article "Time, Knowledge, and Culture", in Leena Pärssinen & Esko Rahikainen (eds.), *Printed in the Memory: Library Treasures in the National Library of Finland,* Helsinki: Otava, 2004, 10-12, and the article "How Can Research Provide for the Future?", *Futura* 35 (3) (2016), 81-82.

Literature

Dawkins, Richard (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Gale, Richard (ed.) (1978). *The Philosophy of Time: A Collection of Essays*. New Jersey: Humanities Press.

Hawking, Stephen (1988). *A Brief History of Time*. New York: Bantam Dell.

Heidegger, Martin (1978). *Being and Time*. Oxford: Blackwell.

Longino, Helen (2002). *The Fate of Knowledge*. Princeton: Princeton University Press.

Niiniluoto, Ilkka (2006). "World 3: A Critical Defence". In Ian Jarvie, Karl Milford & David Miller (eds.), *Karl Popper: A Centenary Assessment vol II*, Aldershot: Ashgate, 59-69.

Niiniluoto, Ilkka (2014). "Scientific Realism: Independence, Causation, and Abduction". In Kenneth Westphal (ed.), *Realism, Science, and Pragmatism*. New York: Routledge, 159-172.

Niiniluoto, Ilkka (2018). *Truth-Seeking by Abduction*. Cham: Springer.

Popper, Karl (1957). *The Poverty of Historicism*. London: Routledge and Kegan Paul.

Popper, Karl (1972/1979). *Objective Knowledge: An Evolutionary Approach*. Oxford: Clarendon Press.

[13] See Pouru et al. (2018).

Pouru, Laura, Wilenius, Markku, Holstius, Karin & Heinonen, Sirkka (eds.) (2008). *Pentti Malaska: A Visionary and Forerunner.* Turku: The Finnish Society for Futures Studies.

Wexionius, Mikael (1642). *Discursus politicus de prudentia*. Aboae: Petrus Wald. https://s1.doria.fi/helmi/bk/rv/fem19990060/

Österman, Bernt & Vilkko, Risto (2016). *Georg Henrik von Wright: A Philosopher's Life*. Helsinki: Into.

Chapter 2: The Third Culture: Six Interpretations

C. P. Snow published his much-debated essay on "the two cultures" in 1959. Later it has been claimed that the division between scientific and humanistic cultures has been overcome by "the third culture". John Brockman in his *The Third Culture* (1995), and the Finnish physicist Kari Enqvist in his widely read books on cosmology and quantum mechanics, have argued that the role of intellectuals in our technology-driven society is currently taken over by natural scientists who present their views and opinions directly to the general public in science books and television.[1] Enqvist (1998) adds that the riddles of the universe need no "deep philosophy", but rather will be solved as matters of theoretical and experimental physics.

In this chapter, I shall first recall what Snow originally claimed about the two cultures. Then, from a philosophical perspective, I propose six alternative interpretations about "the third culture".

What C. P. Snow said?

Charles Percy Snow (1905-80) characterized himself as a scientist by his education and as a writer by his calling. He made his master's thesis in chemistry in 1928, and started to work in Lord Rutherford's laboratory at Cambridge. During 1940-70 he published a novel *Strangers and Brothers* in seven volumes. As a free writer Snow also acted in advisory positions concerning industry and technology policy.

Snow's famous lecture was entitled "The Two Cultures and the Scientific Revolution", but the original title "The Rich and the Poor" referred to his concern about the gap between the

[1] Another third-culture celebrity in Finland is the astronomer Esko Valtaoja. In the USA one could mention Richard Feynman (1918-88), a Nobel laureate in physics and a popular science writer, who famously remarked that "philosophy of science is as useful to scientists as ornithology is to birds".

rich and poor people in the world. Like the Marxist historian of science J. D. Bernal, Snow regarded the advance of the scientific revolution as the only and greatest hope of the poor. What Francis Bacon had prophesied already in the early seventeenth century was now becoming a reality: applied natural science is able to prevent unnecessary miseries of billions of people. According to Snow, this fact is understood by scientists who as optimists have "the future in their bones". But, instead, the representatives of the humanistic culture stick themselves to the past and fail to see the great opportunities provided by the industrial applications of science. Therefore, "a gap of misunderstanding" has been created between the literary intellectuals and the natural scientists, and this gap feeds mutual errors between the two cultures. Especially in England literary intellectual are by their nature "luddits", who wish to oppose and destroy technology. As a cure, Snow recommended a reform of education which – instead of too early specialization – guarantees that future politicians and administrators are able to understand science and scientists.

Inspired by heated discussion, Snow published in 1963 his "second look", where he made sure that he is passionately opposed to the gap between the two cultures (Snow 1993). He denied the proposal to separate pure natural science from technology, but admitted that he had overlooked the position of the social sciences. He also warned against trivializing his thesis by suggestions to separate e.g. two thousand and two cultures.

Snow's diagnosis of his own time seems to be based on sweeping and simplified generalizations. Whatever attitudes the humanists may have had, the political leaders in charge of national economies have only rarely been literary intellectuals.[2]

From our current perspective, Snow's pronouncements sound one-sided and narrowly optimist. Today science-based technology is dominating Western economies, just as Snow wished, but the problems of poverty and violence have not disappeared. We agree that the advancement of science and technology is a necessary condition of economic growth and quality of life, but certainly it is not alone a sufficient condition

[2] In the USA, some of the leaders may have a basic education in the liberal arts colleges (cf. Chapter 13 below).

of social progress.[3] We may disagree with the Frankfurt School's rejection of instrumental reason (cf. Chapter 8 below), but – especially since the green movement in the 1970s – we are also aware of the damages that industries have brought to the natural environment. Snow thus appears to be a victim of the "myth of progress" (cf. Chapter 15 below).

Nevertheless, Snow's recommendation to bridge the gap between the two cultures is certainly topical and well-taken. It is important that the decision makers understand the opportunities provided by science and technology – and the factual premises of their solutions are based on research. But it is equally urgent to educate the scientific and technological personnel to understand the environmental, human, and social consequences of their activities. Politicians should have education in science, but likewise engineers and economists should have education in ecology and ethics.

The Third Culture?

How could we promote a new kind of culture which overcomes the opposition between the first (humanistic-literary) and the second (scientific-technological) culture? I present six possible answers.

[1] *The third culture means the victory of the second culture over the first*. This is the view of Brockman and Enqvist. I have serious objections against it. I am ready to grant the natural scientists an important position among the intellectuals of our society. But – even if we don't mention domains like poetry, music, and the fine arts – the scientists are not self-sufficient even in their own endeavors. Reductionism, which would guarantee that physical theories are "theories of everything" and thereby solve all cognitive problems about the world, is a highly problematic position. To be sure, some philosophers too (eliminative materialists, metaphilosophical naturalists) would support this view, but anti-reductionist emergent materialism is definitely against it: psychological, cultural, and social phenomena need their own theories that are not reducible to the physical level. Further, the interpretation and application of physical and biological theories and the development of technologies involve conceptual, normative, and axiological questions that are not scientific but rather philosophical. The

---

[3] Bertrand Russell stated in his *The Scientific Outlook* (1931) that increase in science is "not enough to guarantee any genuine progress", unless it is accompanied by "increase of wisdom" about the ends of life.

second culture needs support from and co-operation with the first culture – with disciplines like philosophy of science, occupational psychology, ethics of medicine, philosophy of technology, philosophy of biology, and environmental ethics.

[2] *The third culture means the victory of the first culture over the second*. This view is militantly defended by postmodernists who wish to deconstruct binary distinctions – among them the dichotomy between nature and culture. Foucault and social constructivists have argued that nature is a category within culture, so that – in Popperian terms – the physical world 1 is reducible to the cultural World 3. This implies that natural science is merely one of forms of "discourse" or literature, whose theories do not have objective contents and truth conditions. The Science Wars, which raged in the late twentieth century, were a result of this clash: the natural scientists raised to protect the integrity of their work and the social status of their profession. However, some of those who advocate the notion of the third culture against philosophers have mistakenly assumed that philosophers are fighting on the side of the relativists and social constructivists. Therefore, it is important to point out that critical scientific realists oppose postmodernism by defending the ontological and temporal primacy of world 1 to the human-made World 3.

[3] *The third culture means interaction between the first and the second*. According to this view, it is possible to build bridges through which the humanists and the natural scientists can interact, communicate, and understand each other – yet maintaining their own identities. Universities with their multidisciplinary research and education programs can serve as such bridges.[4] I think this is a partially correct view, but it need to be complemented by noting some important domains of culture neglected by Snow.

[4] *The third culture means the intersection of the first and the second*. Examples of combinations of humanistic and scientific ingredients include such theoretical and methodological approaches as cognitive science and futures studies. Another example, proposed by Kevin Kelly, the editor of *The Wired*, is "culture industry", arising from the marriage between

---

[4] From an academic perspective, the contrast between science and humanities is sharper in English-speaking countries, where "science" mainly refers to natural science and the humanities are taught in Faculties of Arts. The Finnish term "tiede", like the German *Wissenschaft*, covers both natural and human sciences.

humaniora and information technology. The same marriage has given rise to the research program of "digital humanities", with applications in linguistics, history, and archeology. It is somewhat surprising that Snow does not mention at all industrial design (or applied arts) which links together aesthetic and technological cultures. A notable trend of our current information society is the application of ICT with traditional humanistic areas, especially in the arts and media, in order to promote and commercialize the products of "content industry" (films, videos, computer games, TV, CD-ROM, virtual reality).

[5] *The third culture is constituted by the social sciences*. This view was hinted at by Snow himself in 1963, and has been developed by the sociologist Wolf Lepenies (1985). Some interpretations assume that systematic social sciences function like descriptive natural sciences, while others emphasize their closeness to the hermeneutic or interpretative humanities. A different approach, which might suggest something novel in terms of Snow's dichotomy, is the critical social science aiming at the emancipatory opening of new societal alternatives.[5] This view brings social sciences close to futures studies which actively designs and evaluates alternative future scenarios, and therefore needs help from natural, human, and social sciences as well as from philosophy and value theory (cf. Chapter 11 below).

[6] *The third culture is constituted by analytic philosophy*. This is my favorite answer. One can find philosophical schools which regard philosophy as a form of literature or spirited conversation about eternal and topical themes (in the sense of the first culture),[6] while some schools have attempted to establish a rigorous scientific method for philosophy (in the sense of the second culture).[7] But analytic philosophy at its best is beyond these alternatives. It aims at conceptual clarity in issues that

---

[5] See Habermas (1972).

[6] A prime example is Richard Rorty (1982), whose inspiration comes from American pragmatism (James, Dewey, but not Peirce), later Wittgenstein, Davidson, Heidegger, Foucault, and Derrida. Rorty's attack against systematic philosophy has in turn inspired the project of "post-analytic philosophy", but the collection with this title does not really deliver its promise, since it includes essays by prominent American analytic philosophers (see Rajchman & West, 1985).

[7] Attempts to establish "scientific philosophy" include Husserl's phenomenology, logical positivism, and Marxism-Leninism.

have genuine content but are so far poorly understood. It respects the results of scientific research, among other products of human culture, but puts their interpretation and applications into a critical context, by employing conceptual analysis and ethical evaluation. In my own awakening to philosophy at age of 19 – as a first-year math student who happened to read Eino Kaila's *Inhmillinen tieto* and Wittgenstein's *Tractatus* - I observed that "philosophy can be as exactly systematic as mathematics but as romantic as fictional literature". The history of philosophy in Finland in the twentieth century, as outlined in Haaparanta & Niiniluoto (2003), shows that a strict logician like G. H. von Wright can be at the same the country's most significant public intellectual. This is a good reason for thinking that analytic philosophy can find ways to overcome the limitations of Snow's two cultures.

Note. This chapter has appeared in Kaisa Luoma, Erna Oesch & Risto Vilkko (eds.), *Philosophical Studies in honorem Leila Haaparanta* (Acta Philosophica Tamperensis 4). Tampere: Tampere University Press, 277-284.

Literature

Brockman, John (1995). *The Third Culture: Beyond the Scientific Revolution*. New York: Simon and Schuster.

Enqvist, Kari (1998). *Olemisen porteilla*. Helsinki: WSOY.

Haaparanta, Leila & Niiniluoto, Ilkka (eds.) (2003). *Analytic Philosophy in Finland*. Amsterdam: Rodopi.

Habermas, Jürgen (1972). *Knowledge and Human Interests*. Boston: Beacon Press.

Lepenies, Wolf (1985). *Die drei Kulturen: Soziologie zwischen Literatur und Wissenschaft*. München: Hanser.

Rajchman, John & West, Cornell (eds.) (1985). *Post-Analytic Philosophy*. New York: Columbia University Press.

Rorty, Richard (1982). *Consequences of Pragmatism*. Minneapolis: University of Minnesota.

Russell, Bertrand (1931). *The Scientific Outlook*. London: Allen & Unwin.

Snow, Charles Percy (1993). *The Two Cultures*. Cambridge: Cambridge University Press.

Chapter 3: Beauty and Information

Is beauty in the object or in the eyes of the spectator? Are aesthetic values objective or subjective? These traditional philosophical questions concern ontological and epistemological issues in the theory of beauty, applicable to nature and works of arts. They can also be combined with a methodological approach starting with the practical problem: is it possible to quantify and measure aesthetic properties? New conceptual and mathematical tools for tackling this question have been developed in the so-called information aesthetics. In this chapter, I argue that most recent proposals in this direction are mistaken, since they fail to do justice to insights known already to authors like E. A. Poe. Still, in spite of its limitations in practice, the explication of aesthetic value by means of information value is not a futile exercise. I suggest that it gives some support to qualified objectivism in the theory of beauty.

Thomas Reid on Beauty

In his *Essays on the Intellectual Powers of Man* (1785), the Scottish common sense philosopher Thomas Reid presented an admirably clear analysis on the concept of beauty. This beautiful theory, which appeared five years before Kant's *Kritik der Urteilskraft* (1790), seems to be largely underestimated and neglected by later commentators.

According to Reid, the human mind is able by the "internal taste" to perceive beauty in natural and artefactual objects. However, he argued, all the different things that are called beautiful do not possess any common quality: "the kinds of beauty seem to be as various as the objects to which it is

ascribed".[1] Why, then, objects with nothing in common are called by the same name? The explanation, Reid suggested, is in the fact that these objects have "some common relation to us or to something else".[2]

Beautiful things are similar, in Reid's view, in two respects:

(1) When beautiful objects are perceived, or even imagined, they produce a certain agreeable emotion or feeling in the mind.

(2) This agreeable emotion is accompanied with an opinion or belief of their having some perfection or excellence belonging to them.

Moreover, Reid could not accept the "disrespectful" idea that the Divine Author has provided human beings only with erroneous beliefs. Thus, when the belief mentioned in (2) is true, the object in fact has some *real* perfection - and this is precisely what we call the beauty of that object.

Reid compared beauty to the concept of sweetness, which is associated with our "external senses". In contrast with John Locke, who took sweetness to be a secondary quality, Reid said that sweet and bitter are "attributes of external objects, and not of the person who perceives them". The sensation of sweetness is in the human mind, sweetness itself in sugar. Likewise, the sensation and feeling of beauty is in the mind, but beauty itself in the object. In this sense, beauty is an objective notion, even though different types of objects (e.g. landscape, bird, musical composition, painting, mathematical theorem, human character) have their own types of beauty.

Beyond Subjectivism and Platonism in Aesthetics

Thomas Reid's analysis certainly needs further elaboration. But it has the virtue that it gives, by combining the naturalist

---

[1] Reid's thesis, but applied to the concept of art, was reformulated by Morris Weisz (1956) who was inspired by Wittgenstein's later philosophy. Wittgenstein's notion of family resemblance can be found in Reid's successor Dugald Stewart in 1818 (see Dickie, 1971).

[2] Reid's explanation is essentially the same as the reply of Maurice Mandelbaum (1965) to Weisz.

analysis of beauty with emotions and beliefs, a promising alternative to subjectivism and Platonism in aesthetics.[3]

Aesthetic *subjectivism* in its most radical form asserts that

(3) "x is beautiful" means "I regard x as beautiful".

Here (3) may be an intellectual judgement (about what I assess to be beautiful on some evidence) or an emotional judgement (about what I feel or taste to be beautiful).[4]

Experience of beauty is no doubt a crucially important topic for the psychology of aesthetics, and it can also be investigated by means analytic philosophy (e.g. Wittgenstein's concept of "seeing as") and phenomenology (intentionality of perception).[5] But, as a definition of beauty, (3) is hopeless: the content, intensity, and quality of my experiences of the same or similar objects may vary in a highly context-sensitive way. In this respect, it does not help very much to replace the first-person pronoun "I" in (3) with "every member of a culture A" (*cultural relativism*) or "every qualified observer of type X" (*authority theory*). By what non-arbitrary criterion could we identify the cultural groups or observers who are authorized to act as the judges in matter of beauty?[6]

These variations of subjectivism are incompetent to explain why some objects are valued as beautiful, and others are not. In this respect, they resemble the *emotivist* thesis that

(4) "x is beautiful" means the same as the exclamation "Oh x!".

Perhaps the most interesting form of subjectivism has been developed by Immanuel Kant. However, it is not at all evident

---

[3] The structural similarity of positions in ethics and aesthetics can be clearly seen in Dickie (1971). This does not mean that the counterpart theories are equally plausible. Thus, I am inclined to defend objectivism about beauty, but moderate relativism concerning ethical values. See Niiniluoto (1991).

[4] "Each mind perceives a different beauty", wrote David Hume in 1757 (see Hume, 1894, 136).

[5] See, e.g., Scruton (1974).

[6] If we need to apply aesthetics in practice, I am ready to admit that the best method is to rely on expert judgement by well-educated persons. To find such persons, I might consult the International Institute of Applied Aesthetics in Lahti. But, in my view, this is not adequate as a *definition* of beauty.

that judgments of taste, when they don't involve "interest and concept", are really capable of reaching a "universal" and "unconditional" consensus among human beings, as Kant required.[7] In this respect, Kant's account of beauty is equally problematic as the pragmatist *consensus theories* of truth.[8]

The value subjectivists start from mental phenomena (experience, taste, emotions in individuals and groups), but fail to link them with the nature of the objects of aesthetic judgements. The value *objectivists* start from the other end by counting beauty as a property of the object, and then argue that the existence of this objective aspect explains our typical experiences with the object.

The oldest form of objectivism is *Platonism*, which takes beauty to be an ideal transcendent form. Its main problem - apart from ontological extravagance - is the difficulty in explaining how particular objects in our world could have their share of the ideal beauty.

Another objectivist theory takes beauty to be an *independent quality* of some objects.[9] For example, this theory could assert that my cat Blackie has the properties of having four legs, being black, beautiful, five years old, etc. Most versions of this account presuppose that we have some special capacity of experiencing or directly knowing this quality of beauty. But it is highly implausible to assume that the beauty of my cat could be somehow independent of her other properties.

Reid's theory is a suggestion towards aesthetic *naturalism*, which takes beauty to be a "real excellence" in the object, with the capacity of producing agreeable emotions and beliefs within our minds. But there is no common quality of beauty among all beautiful things, but each type of object has its own beauty depending on its natural qualities or properties:

(5) For an object x of type A, "x is beautiful" means "x is G", where G is definable by the natural properties of x.

---

[7] See Kant (1952).

[8] See, e.g., Niiniluoto (1999), 101, 236.

[9] Dickie (1971) calls this view "Platonism 2". Its counterpart in ethics, advocated by G. E. Moore, is known as non-naturalism. As this view does not presuppose the independent existence of the universal "beauty", it need not involve a Platonist ontology.

To use a fashionable term, the concept of beauty supervenes on the natural properties.[10]

In most cases, the beauty of an object depends on the "sum" of many of its properties. For example, the shape, fur, and color of my cat all contribute to her beauty. These combinations of relevant properties may be specific to the types of objects: woman, man, cat, car, forest, poem, and theory are evaluated on different standards. Further, beauty may be a "holistic" or "emergent" property of a whole.[11] Human face is a good example: attempts to combine together the most beautiful eyes, nose, mouth, and hair in the world have not led to any appealing results.[12]

Sircello's Theory of Beauty

What we have said so far can be compared to Guy Sircello's *A New Theory of Beauty* (1975), one of the best analytic attempts to present a theory of beauty.

Sircello's promising starting point is that beauty is primarily a property of properties of objects - hence, a kind of second-order property. Instead of Platonic transcendent forms or Aristotelian immanent universals, Sircello means by properties what other philosophers have called property-instances, tropes, and abstract particulars: the greenness of those particular hills, the clearness of Helen's skin (*ibid*., 14).[13] Thus, beauty is not attributed to clearness-of-skin in general, but to Helen's skin as it is now.

According to Sircello, a property F is a PQD (a Property of Qualitative Degree) if it is possible for one object to be "more or less" F than another object, but there is no "single scale" for measuring this degree (*ibid*., 39). Thus, vividness, delicacy, and depth of colors are PQD's, but being square, being pregnant,

---

[10] This term was introduced in ethics by Richard Hare (1952).

[11] Cf. the "regional properties" of Beardsley (1958).

[12] It is well-known that the ideals of feminine beauty vary in different cultures. This has not prevented the organization of international beauty contests (Miss Universe) or disturbed the global market for fashion products, cosmetic companies and surgery. For a person, a beautiful appearance may have not only intrinsic but also instrumental value (with economic, erotic, hedonistic, vital, and social dimensions).

[13] For tropes, see Campbell (1990).

being tall, and being hot are not. Then Sircello's Theory of Beauty says that a property F, as a PQD of object x, is beautiful if and only if (i) F is not a property of deficiency, lack, or defect (e.g. "being deformed"), (ii) F is not a property of the appearance of deficiency, lack, or defect (e.g. "being rotten-smelling"), and (iii) F is present in x to a very high degree (*ibid*., 43).

It is easy to understand why Sircello wants to exclude genuine quantitative properties from his class of PQD's: given his condition (iii), all tall basketball players would be "beautiful" with respect to their tallness. But this seems to be an *ad hoc* requirement: why could not numerically quantifiable properties (e.g. scales and proportions of landscapes and buildings) be potentially beautiful? Of course Sircello could argue that his treatment is restricted to aesthetically relevant properties, but that would make his theory circular.

The crucial condition (iii) appears to be problematic. A dim painting of morning mist need not contain any positive comparative quality "in a very high degree".

Sircello's next step is to say that an object (or concrete particular) is beautiful only if it has a beautiful property. In other words,

> (6) "x is beautiful" means that there is a trope F such that x has F and F is beautiful.

Sircello presents this only as a necessary condition for an object to be beautiful. But he asserts that the beauty of objects is always reduced to the beauty of their properties.

Sircello's treatment of properties is curiously atomistic, however. To link (6) with the naturalist approach (5), single tropes should be replaced by combinations or bundles of tropes. He notes the phenomenon of disharmony of an object's properties (*ibid*., 101), but does not otherwise recognize the problem that the beauty of PQD's may be context-sensitive. Two fine colors put together may give an awful result. Thus, Sircello fails to pay enough attention to the object as a whole and to the interplay of its properties as a basis of its beauty.

Holistic Beauty and Numerical Aesthetics

Among the holistic definitions of beauty, one may mention harmony and symmetry (Plato), proportion and symmetry

(Vitruvius), perfect harmony of parts (Leon Battista Alberti), uniformity in variety (Francis Hutcheson), and balance of form expressing life (see Eliel Saarinen, 1985). All these concepts strive to express something like the *compositional* or *structural beauty* of a complex object or system. The attempt to define such notions by means of geometrical and combinatorial concepts is called *numerical* or *mathematical aesthetics*, and it is continued in *information aesthetics* and *cybernetic aesthetics*.[14] The most important centers of this program in the 1960s and 1970s were Stuttgart (Max Bense) and Strasbourg (Abraham Moles).

Assume that M(x) is a measure of the aesthetic value of an object. Then we could say that object x is "structurally beautiful" if M(x) is sufficiently high. Object x is structurally more beautiful than object y if M(x) > M(y). Definitions of this form do not give any universal conception of beauty, i.e. necessary and sufficient conditions for being beautiful, but only a special notion applicable to certain types of objects. On the other hand, it is important that this class of objects includes both natural things (e.g. landscapes) and artifacts (e.g. works of art).

This approach is also in another way in conformity with Reid's. Complex objects can be looked at in different ways, as Gestalt psychology has taught us.[15] These ways or modes of perception have to be learned. When the objects produce pleasant feelings in our mind, we believe that this emotion is not merely internally generated, but there is also some real excellence in the composition and structure of the object. It is this real perfection that is under study within numerical aesthetics.

A pioneering work in mathematical aesthetics is G. D. Birkhoff's *Aesthetic Measure* (1932), which proposes a measure M for the aesthetic value of geometrical polygons - extending it also to ornaments, simple musical pieces, and poems.[16] According to Birkhoff's formula,

(7) $M = O/C$,

where C is the *complexity* of a given polygon (i.e. the number of straight lines needed to cover its sides), and O is its *order*.

---

[14] See Birkhoff (1958), Moles (1968), Bense (1969), Maser (1970), Arnheim (1971), Gunzenhäuser (1975), Stiny & Gips (1978), Frank & Franke (1997).

[15] See Arnheim (1967).

[16] See Birkhoff (1952).

The order is defined by O = V + E + R + HV - F, where V is vertical symmetry, E is equilibrium, R is rotational symmetry, HV is relation to a horizontal-vertical network, and F is a penalty for too small angles or too close vertices.

Since the 1950s it became customary to interpret aesthetic measures in terms of the psychology of perception. Birkhoff's formula is then related to the way in which perceptual information is processed in the human mind. Aesthetic experience is thereby psychologically explained as the enjoyment or pleasure due to the effort of conceiving the perceived object, when the mind is functioning in its normal way.[17] If the object is too complex, the conceptual organization or structuring does not succeed and aesthetic pleasure is not produced.[18]

It is historically interesting - though almost forgotten - that the American author Edgar Allan Poe pondered about problems in information aesthetics already in the 1840s. According to Poe, poetry is defined as "the rhythmical creation of beauty". In the essay "The Rationale of Verse", he noted that "man derives enjoyment from his perception of equality", where equality embraces the ideas of similarity, proportion, identity, repetition, adaptation, and fitness. Poe goes on to suggest that such enjoyment from perceived equalities would increase in the form of a series of potencies:

> I have no doubt, indeed, that the delight experienced, if measurable, would be found to have exact mathematical relations such as I suggest; that is to say, as far as a certain point, beyond which there would be a decrease in similar relations.[19]

The last sentence is explained in a short note "Rhyme", where Poe reminds of Francis Bacon's just aphorism about human faces:

---

[17] According to Moles, this efficiency is about 100 bits per second. See Moles (1968) and Arnheim (1971). The relevance of scale to beauty was discussed already in Aristotle's *Poetics*.

[18] Stiny and Gips (1978), 119, quote F. Hemsterhuis's 1769 definition of the beautiful as "that which gives the greatest number of ideas in the shortest space of time". One of their definitions of the "evocative" aesthetic value is the ratio of the measure of the evoked emotions to the measure of the evoking object (*ibid.*, 115). This measure allows that a sentimental piece of popular music is more beautiful than an abstruse symphony.

[19] See Poe (1977), 122, 310.

> There is no exquisite beauty without some strangeness in the proportions.

Regularity and expectedness have to be balanced by novelty, originality, and unexpectedness:

> Perfection of rhyme is attainable only in the combination of the two elements, Equality and Unexpectedness. But as evil cannot exist without good, so unexpectedness must arise from expectedness.[20]

Syntactical Information and Redundancy

Claude Shannon's 1948 theory of communication is based on a *syntactical* concept of information. Assume that the signs a_1, ..., a_n occur in a communication channel with regular relative frequencies p_1, ..., p_n. Then the *entropy* associated with a message containing one sign is defined by

$$(8)\quad H = - \Sigma\, p_i \log p_i,$$

where log is the logarithm function with base 2. Entropy H is a measure of the "uncertainty" or "unexpectedness" of a message. It receives its maximum value H_{max} (= logn), when all signs a_i are equally frequent or probable, i.e. $p_i = 1/n$ for all i = 1, ..., n. Entropy H receives its minimum value 0, when one of the signs occurs with probability 1, i.e. can be expected with certainty. The *information* or "surprise value" provided by a message can then be defined by change in entropy, i.e. as elimination of uncertainty.

The transmission of information in a channel is most efficient, when the source has the maximum entropy H_{max}, since then there is much uncertainty to be eliminated. But in natural language there is always some redundancy, due to differences in the probabilities of different signs and statistical correlations between successive signs.[21] In other words, not all combinations of signs are equally probable. The *redundancy* R of an information source with entropy H can be defined by

$$(9)\quad R = 1 - H/H_{max}.$$

---

[20] See Poe (1977), 311.

[21] Poe was aware that such statistical regularities can be used as a method of deciphering secret writing. See his story "The Gold-Bug" (1843).

Hence, R has the value 0, if $H = H_{max}$, and the value 1, if $H = 0$. The redundancy of ordinary English text is about between .5 and .8.

The function of redundancy in normal communication is to increase the possibility of understanding: even if all letters or words are not heard, the missing parts can easily be guessed. But it has also another function: to produce aesthetic pleasure. This observation was made in the 1950s, when Shannon's theory was applied to art, especially music: if art is a semiotic process, where certain combinations of sign occur in relatively regular frequencies (at least within styles and traditions), then the corresponding subjective probabilities give structure to the anticipations and expectations of the spectator or the listener.[22] Without some redundancy such probabilistic expectations would not be possible.

An excellent formulation of the relation of redundancy and aesthetic value was given by the Finnish mathematician Gustav Elfving (1965).[23] Compare the sentences

(a) KÄKI KUKKUU KAUNIHISTI

(b) JKZAO AWQIA RCEXR LUMSZ.

Sentence (a) belongs to a language (Finnish, in fact) with low entropy, while sentence (b) belongs to a language with high entropy and with highly effective transmission of information. But it is clear that in (a) part of the efficiency is sacrificed in favor of aesthetic qualities: (a) is more beautiful than (b).[24] But redundancy should not be too large, either: the sentence

(c) KKKKKKKKKKKKKKKKKKKKKK

is not aesthetically pleasant, since it contains too much repetition. The aesthetic value of a sentence presupposes that it is not too chaotic or irregular (like (b)) or too repetitive or regular (like (c)), but has a proper degree of redundancy.

Elfving generalizes this idea to two-dimensional pictures with white and black squares. Aesthetically interesting results,

---

[22] See, e.g., Attneave (1959), Meyer (1967).

[23] Gustav Elfving (1908-84) was my teacher at the University of Helsinki, when I wrote in 1968 my Master's thesis on Bayesian statistics.

[24] Another example could be the Finnish sentence AJAKAA HILJAA SILLALLA, which once won an international linguistic beauty competition.

"stochastic" or "aleatory" works of art, are obtained, when there are some, but not too high, statistical correlations between the colors of adjacent squares. In principle, this idea is very generally applicable, since almost all pictures can be coded in a digital form (raster photographs in newspapers, photo processing by computers). This is the basic idea of later approaches to "generative aesthetics" or *computer art*.

Quantitative measures of beauty have been proposed in information theory by replacing, in Birkhoff's formula M = O/C, the order O by redundancy R and complexity C by entropy H. For example, Rul Gunzenhäuser (1975) defines aesthetic value by

$$(10)\quad M = R/H = 1/H - 1/H_{max}.$$

Then M = 0, if R = 0, i.e. the object is chaotic (cf. Elfving's (b)). But M receives the value infinity, if R = 1 and H = 0, i.e. the object is completely regular (cf. Elfving's (c)). Thus, definition (10) does not take into account the insight of Bacon, Poe, and Elfving: too regular objects are not excellent in their beauty.

For this reason, definition (10) should be replaced by a function, which increases with the difference $R - R_{opt}$ between actual redundancy R and the optimum redundancy R_{opt} from the aesthetic viewpoint. If optimum entropy H_{opt} is defined by $R_{opt} = 1 - H_{opt}/H_{max}$ (cf. formula (9)), then a simple piecewise linear function of the required kind is

$$(11)\quad M = 1 - |H - H_{opt}|/[H_{max} - H_{opt}],$$

where |a| is the absolute value of a. This function receives its maximum 1, if $H = H_{opt}$, and its minimum 0, if $H = H_{max}$. For

H = 0, M receives the value $(H_{max} - 2H_{opt})/(H_{max} - H_{opt})$ (see Fig. 1).

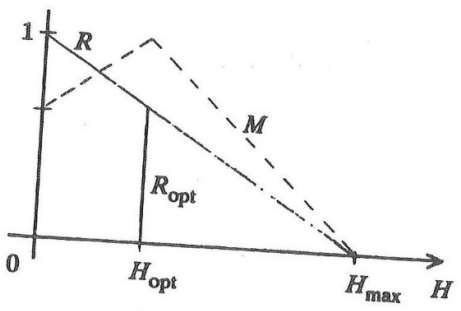

Fig. 1. Beauty and Entropy

Art and Semantic Information

The process of generating works of art can, at least in some cases, be simulated by stochastic processes. However, as Stiny and Gips (1978) point out, the assumption underlying Shannon's information theory, viz. the existence of regular relative frequencies, is very artificial in many situations involving words, pictures, and musical tones.

Tarkko Oksala (1981) has proposed that the products of architecture are evaluated by the semantic information measures of Rudolf Carnap, Yehoshua Bar-Hillel, and Jaakko Hintikka.[25] The basic idea is to conceive a building as a composition from simple standard elements, so that its information value depends on its improbability with respect to the possibilities of making different combinations of elements.

More precisely, the semantic *information content* of a sentence h in a language L is the class of states of affairs excluded by h, where the relevant states of affairs are describable in language L. If these states w are assigned with logical or epistemic probabilities P(w) (expressing degrees of belief in w), or by

[25] For semantic information, see Hintikka and Suppes (1970).

probabilistic weights of significance, the amount of information in h can be defined by

$$(12)\quad \mathrm{cont}(h) = 1 - P(h) = P(\sim h) = \sum_{w \vdash \sim h} P(w)$$

or

$$(13)\quad \mathrm{inf}(h) = -\log P(h).$$

A practical limitation of this approach arises from the fact that a reasonable account of the required probability measures has been developed only for monadic first-order languages which contain one-place predicates but not relations.

Beauty and Kolmogorov Complexity

A.N. Kolmogorov defined in 1963-64 a non-probabilistic measure of information, based on the notion of computational complexity. Let x be a finite binary sequence (consisting of zeros and ones), and let L(x) be its length. Then the *Kolmogorov complexity* K(x) of x is defined as the length of the shortest sequence (or program) by means of which a universal Turing machine is able to generate x. If x is very regular, a short description is sufficient to produce it (i.e. K(x) is small), but a highly complex or "random" sequence without repetitions and invariance needs x itself as input (so that K(x) is of the same size as L(x)).

As any finite text, consisting of signs, can be coded by binary sequences, the definition of K(x) can immediately be generalized to compositional objects and systems. Stiny and Gips (1978) apply it to texts, geometrical figures, pictures, and musical pieces. They define in effect the aesthetic value of an object x by the ratio of its length L(x) and complexity K(x):

$$(14)\quad E(x) = L(x)/K(x).$$

They note that this corresponds to Francis Hutcheson's (1725) formula: beauty is "a compound ratio of uniformity and variety".[26] While the length L(x) of x is a measure of the possible

[26] See Stiny and Gips (1978), 108. A counterpart to measure (14) in the philosophy of science is Eino Kaila's concept of relative simplicity, defined as the ratio between the explanatory power of a scientific theory and its complexity. See Niiniluoto (1994).

variety of x, the inverse 1/K(x) is a measure of the unity, organization, coherence, or regularity of x.[27]

The Stiny - Gips measure E is obviously a non-probabilistic counterpart to aesthetic measures defined by entropy H. In Elfving's examples, K and H correspond to each other: K(x) is small for the regular (c), K(x) is large for the irregular (b), and K(x) is between these extremes for (a). Hence, the counterpart of (14) in statistical information theory would be

(15) H_{max}/H.

But this measure is again unsatisfactory - and for the same reason as R/H in (10): when the regularity of an object exceeds its optimum value, its beauty should start to decrease. Thus, (14) should be replaced by a function like

(16) $1 - |K(x) - K_{opt}|/[L(x) - K_{opt}]$

which decreases when the complexity K(x) differs more from its optimum value K_{opt}.

Concluding Remark

The attempt to define aesthetic measures by means of information theory is an unfinished program. At best it can give a partial analysis of a special notion of beauty.[28]

Even if we could reduce the notion of beauty, via information, to probability, this would not solve the philosophical dispute whether beauty is subjective or objective, since probability itself has subjective and objective interpretations.

If the probabilistic measures are based upon subjective or personal probabilities - i.e. upon beliefs, expectations, surprise

---

[27] See Stiny and Gips (1978), 103, 109.

[28] In 1995 I stated that I don't wish to make any predictions about the future prospects of information aesthetics. This caution was warranted, as this research program seems to be degenerating. In a recent survey, Nake (2012) admits that information aesthetics *was* an "heroic experiment", but its main achievement is the emerging new technology of computer art. In another survey, Giannetti (undated) claims that the rationalistic syntactical framework of cybernetic aesthetics fails to take into the account the subject, context, and semantic content of communication. But this seems to be a limited perspective of the richer possibilities of analyzing beauty by different concepts of information.

values, and weights of significance relativized to persons -they can be interpreted as measures of experienced beauty.

If the measures are based on physical probabilities (associated with the stochastic generation of the object) or on complexity (associated with a linguistic description of the structure of the object), aesthetic value will be a function of the real properties and relations in the object. In this sense, structural beauty gives grounds for a naturalist and objectivist theory.

Note. This chapter is based on a lecture in a symposium on Beauty, organized by the Philosophical Society of Finland in Tampere in January 1987. Its earlier Finnish version, "Kauneus ja informaatio", appeared in Markus Lammenranta & Veikko Rantala (eds.), *Kauneus*, Filosofian laitoksen julkaisuja, Tampereen yliopisto, Tampere, 1990, 200 - 222. The unpublished English version was presented Aesthetics in Practice, XIIIth International Congress of Aesthetics, in Lahti, August 3, 1995.

Literature

Arnheim, Rudolf (1967). *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley: University of California Press.

Arnheim, Rudolf (1971). *Entropy and Art*. Berkeley: University of California Press.

Attneave, Fred (1959). "Stochastic Composition Processes". *Journal of Aesthetics and Art Criticism* 17.

Beardsley, Monroe (1958). *Aesthetics*. New York: Harcourt, Brace & World.

Bense, Max (1969). *Einführung in die informationstheoretische Ästhetik*. Reinbeck: Rowohlt.

Birkhoff, George D. (1932). "A Mathematical Theory of Aesthetics and its Application to Poetry and Music". In *Collected Mathematical Papers*, vol. 3, New York: Dover, 1958.

Campbell, Keith (1990). *Abstract Particulars*. Oxford: Blackwell.

Dickie, George (1971). *Aesthetics: An Introduction*. Indianapolis: Bobbs-Merrill.

Elfving, Gustav (1965). "Information and Esthetic Evaluation". In *Studia Logico-Mathematica* (Acta Philosophica Fennica 18), Helsinki: Philosophical Society of Finland, 7-11.

Frank, Helmar & Franke, Herbert W (1997). *Ästhetische Information: Eine Einführung in die kybernetische Ästhetik*. München: KoPad.

Giannetti, Claudia (no date), "Cybernetic Aesthetics and Communication". www.medienkunstnetze.de/themes/aesthetics_of_the_digital/cybernetic_aesthetics/16/

Gunzenhäuser, Rul (1975). *Mass und Information als ästhetische Kategorien*. Baden-Baden: Agis-Verlag,

Hare, Richard (1952). *The Language of Morals.* Oxford: Oxford University Press.

Hintikka, Jaakko & Suppes, Patrick (eds.), *Information and Inference*. D. Reidel, Dordrecht, 1970.

Hume, David (1757/1894). "On the Standard of Taste". In *Essays on Moral and Political*. London: Routledge and Sons.

Kant, Immanuel (1790/1952). *Critique of Judgment*. Oxford: Oxford University Press.

Mandelbaum, Maurice (1965). "Family Resemblances and Generalizations Concerning the Arts", *American Philosophical Quarterly* 2, 219-228.

Maser, Siegfrid (1970). *Numerische Ästhetik*. Stuttgart: Karl Krämer Verlag.

Meyer, Leonard (1967). *Music, the Arts, and Ideas.* Chicago: University of Chicago Press.

Moles, Abraham (1968). *Information Theory and Esthetic Perception*. Urbana: University of Illinois Press.

Nake, Frieder (2012). "Information Aesthetics: A Heroic Experiment", *Journal of Mathematics and Arts* 6 (2-3), 65-75.

Niiniluoto, Ilkka (1991). "What's Wrong With Relativism?", *Science Studies* 4 (2), 17-24.

Niiniluoto, Ilkka (1994). "Descriptive and Inductive Simplicity". In Wesley Salmon & Gereon Wolters (eds.), *Logic, Language, and the Structure of Theories*, Pittsburgh: University of Pittsburgh Press, 147-170.

Niiniluoto, Ilkka (1999). *Critical Scientific Realism*. Oxford: Oxford University Press.

Oksala, Tarkko (1981). *Logical Aspects of Architectural Experience and Planning*. Research Papers 66, Otaniemi: Helsinki University of Technology.

Poe, Edgar Allan (1927/1977). *Poems and Essays*. London: J. M. Dent and Sons Ltd.

Reid, Thomas (1969). *Essays on the Intellectual Powers of Man.* Cambridge, Mass. The MIT Press.

Saarinen, Eliel (1985). *The Search for Form in Art and Architecture*. New York: Dover.

Scruton, Roger (1974). *Art and Imagination*. London: Methuen.

Sircello, Guy (1975). *A New Theory of Beauty*. Princeton: Princeton University Press.

Stiny, George & Gips, James (1978). *Algorithmic Aesthetics: Computer Models for Criticism and Design in the Arts*. Berkeley: University of California Press.

Weisz, Morris (1956). "The Role of Theory in Aesthetics", *The Journal of Aesthetics and Art Criticism* 15, 27-35.

Chapter 4: Representation, Caricatures, and Truthlikeness

Representation is a key concept of the philosophy of culture, since representing is one of the functions of practically speaking all human-made cultural artefacts. Indeed, representation is a common aspect of our everyday life, language, art, and science. But this notion cannot be taken for granted, and in the postmodern era it has been subjected to criticism and even attack by many philosophers. Even though some of these critics call themselves "pragmatists", I argue in this chapter that a satisfactory account of representation is provided Charles S. Peirce's semiotics. This semiotic approach is also unified in the sense that it illuminates variations of this theme within cognitive science, philosophy of language, philosophy of science, and aesthetics. The related problem of misrepresentation, especially the ability of caricatures to refer to their objects, can be clarified by the account of truthlikeness and idealization developed by philosophers of science in the spirit of critical scientific realism.

Representation Embattled

Examples of representation can be found in the use of language in everyday life, works of art in various fields, and theories and models in different scientific disciplines. We use several kinds of signs to refer to objects around us and to represent even fictional objects in art and unobservable theoretical entities in science. As Woosuk Park (2014) suggests, we need a unified theory of representation which seeks connections between linguistics, cognitive science, aesthetics, and philosophy of science. But such an account should also give a reply to the postmodern critics who have challenged the need or the viability of the notion of representation altogether. This issue is

connected to the philosophical problem whether it is possible to combine realism (i.e. mind-independent reality as the object of knowledge) and fallibilism (i.e. uncertainty and corrigibility of all factual knowledge claims). Therefore, the account should also deal with the problem of misrepresentation which philosophers of science have discussed in their treatments of truthlikeness and idealization.

In this chapter, I argue that the ingredients of a rich account of representation are provided by the theory of semiotics that Charles S. Peirce (1839-1914) developed already a century ago.[1] Peirce was the American founder of pragmatism whose system of ideas was based on his theory of signs or "semeiotic".[2] In the mid-twentieth century in Europe, Peirce's semiotics had to compete with Saussure's alternative approach. But even though semiotics has become a highly fashionable international research field, it is curious how often authors who are discussing and criticizing the notion of representation ignore Peirce's seminal contributions in this area.

Perhaps the strongest attack on representation comes from Richard Rorty's (1980) attempt to rewrite the history of philosophy without assuming that the human mind is "a mirror of nature". Rejecting "privileged representations" like Immanuel Kant's intuitions and concepts, and appealing to "heretical followers" of Edmund Husserl and Bertrand Russell, Rorty takes his clues from John Dewey's rejection of the "spectator theory of knowledge", Donald Davidson's philosophy of language which admits that we and our languages have only causal relations to the world, Ludwig Wittgenstein's conception of language as a tool rather than a mirror, W. V. O. Quine's rejection of the analytic-synthetic distinction, and Wilfrid Sellars' account of truth in terms of correct assertability by means of inferential steps within language. Rorty thinks that Peirce is "overpraised", as "his contribution to pragmatism was merely to give it a name, and to have stimulated James".[3] His version of pragmatism without the correspondence notion of truth rejects the idea of a world which is independent of our conceptual schemes and our knowledge – for him this is just a

---

[1] See Niiniluoto (2014).

[2] See Peirce's *Collected Papers* (*CP*) (1931-35), (1958); Short (2007).

[3] See Rorty (1982), 160-161.

meaningless "world well lost".[4] Rorty (1991) concludes with "antirepresentationalism" which "does not view knowledge as a matter of getting reality right, but rather as a matter of acquiring habits of action for coping with reality". This framework eschews discussion of realism-vs.-antirealism by denying that the notion of "representation" has any useful role in philosophy.

Huw Price (2010) calls his position "non-representationalism" or "anthropological pragmatism". With influences from Robert Brandom's Sellarsian view of statements as inference tickets instead of word-world correspondences, and from "expressivism", where moral statements are interpreted as expressions of evaluative attitudes rather than assertions of matters of facts, Price's account of the functions of our linguistic interactions is such that the question about more or less correct representation of reality does not even arise.

For a Peirce-scholar any position of "pragmatism" without semiotic relations seems utterly strange. One may recall that Peirce in 1905 renamed his doctrine as "pragmaticism", which is "ugly enough to be safe from kidnappers" (*CP* 5.414). But the real issue here is not only nomenclature about a philosophical school, but rather what Jaakko Hintikka (1997) has called "the ultimate presupposition of twentieth-century philosophy". Hintikka draws a contrast between two positions: language as a *universal medium*, which claims that we cannot step outside language so that semantic relations between language and the world are inexpressible or ineffable, and language as a *calculus*, which allows the systematic variation and study of the interpretation of linguistic expressions and their referents in the world. In the former camp, he includes Frege, Russell, Wittgenstein, Quine, and Heidegger, in the latter Peirce, Hilbert, Husserl, later Carnap, Tarski, and Hintikka himself.[5] Hintikka defends his calculus view by interpreting Wittgenstein's language-games as a method of establishing language-world connections, and develops this idea in detail with his game-theoretical semantics for natural and formal languages. On the

---

[4] *Ibid*., 3-18.

[5] This distinction, which van Heijenoort originally formulated in the context of logic, is a useful heuristic, but its application to specific authors may be debated.

basis of this distinction, one may find that the presuppositions of contemporary antirepresentationalists – from Davidson and Sellars to Rorty and Price – belong to the universal medium view.

In aesthetics, the trend of postmodernism suggested that the notion of representation needs rethinking, as the artists have learnt to play with many kinds of representation modes,[6] but soon it started to refer to new decorative styles with emphasis on play and parody. Nelson Goodman's *Languages of Art* (1976) criticized heavily resemblance theories of representation. Many French post-structuralists and deconstructionists concluded that semiotic systems are closed in the sense that they cannot refer to anything extra-linguistic outside the endless relations between signs. According to Jacques Derrida, "there is nothing outside language", and Jean Baudrillard (1983) claimed that signs not only distort and "pervert" their referents but often function as *simulacra* which hide the disappearance or non-existence of their apparent objects. Among philosophers of science, Baudrillard's account has been used as an argument for the thesis that "realism is dead".[7]

Representation is a widely used concept in cognitive science, as one can see in Jerry Fodor's *Psychosemantics* (1987). Here this notion as such is not under attack, since the need of mental representations is presumed, but the problem lies in its use in specific senses which – without arguments – exclude others.

Philosophers of science, who are interested in the representative capacity of scientific theories and models, have proposed "minimalist" approaches which are not committed to any specific type of representation. For example, Mauricio Suárez (2004) argues that representation should not be tied with the traditional notions of isomorphism and similarity: for a minimalist "inferential" account of representation it is sufficient that the "representational force" of model M points toward target R, and that M allows competent and informed agents to draw specific inferences regarding R.[8]

---

[6] See Wallis (1984).

[7] See Wise (2005).

[8] Dyadic notions of representation between models and targets have been challenged by Tarja Knuuttila (2009).

Peirce's Semiotic Account of Representation

Many of the worries, which have motivated attacks against representation, are answered by the theory of semiotics that Peirce developed around the year 1900. According to Peirce, "a sign, or *representamen*, is something which stands to somebody for something in some respect or capacity" (*CP* 2.228). This notion of representation is extremely broad, as it allows anything to be a sign as long as some interpreter finds a ground for taking it to denote or stand for something.[9] It is irreducibly triadic, as it always involves a sign (or sign-vehicle), object, and interpretant. Even though the "interpretant" for Peirce is not the same as "interpreter" in person, but rather another sign or "an idea to which it gives rise" (*CP* 1.139), which suggests the process of "unlimited semiosis",[10] the "final interpretant" gives the whole truth about the sign's object.[11] The "immediate object" is the object "as cognized in the sign", and the "dynamic object" is the object itself "in relations as unlimited and final study would show it to be" (*CP* 8.183).

Instead of Peirce's triadic approach, the European trends followed Ferdinand Saussure's semiology, where a sign is a dyadic relation between a "significant" and "signifié".[12] Typically this is a coded conventional relation between a spoken word and the expressed mental idea or meaning, so that semiology does not include the realistic sign – object relations of Peirce's triad. This led to the post-structuralist and postmodernist images of us in the midst of an unlimited semiosis without links to extra-linguistic reality.[13]

John Deely, a specialist in semiotics and Thomist philosophy, has argued that Peirce with his pragmaticism "stands at

---

[9] Smoke is an indexical sign of fire, as fire causes smoke, but to treat the natural event of smoke as a representation it has to interpreted as a sign. The debate whether names are natural or conventional was started with Plato's dialogue *Cratylus* (see Niiniluoto, 2008).

[10] See Eco (1976), 68.

[11] See Short (2007), 190.

[12] See Culler (1976).

[13] However, perhaps the main point of Derrida's deconstructionism was to oppose Platonist universals, or the assumption of "a transcendental signified", instead of denying the objective material world.

the interface of modernity and postmodernity, the last of the moderns and the first of postmoderns".[14] He argues that Descartes ignored the semiotic developments in the Iberian Late Latin milieu of his day. The history of philosophy would have taken an alternative course, if the work culminating in John Poinsot's *Tractatus de Signis* (1632) with its triadic notion of sign had been taken seriously. In this very special sense, "the postmodern era in philosophy" started in 1867, when Peirce published his new list of categories. One may agree with Deely that Peirce, with his semiotic consciousness, was able to break with the shortcomings of the modern paradigm. Still, it is risky and potentially misleading to characterize Peirce as "postmodern",[15] since this term is standardly – but in Deely's view "falsely" – applied in France and the United States to new radical forms of linguistic idealism and relativism.[16] Peirce's triadic notion of sign - which links language, reality, thought, action, and community - is in fact a tool for fighting against such vulgar anti-realist doctrines. As argued in Niiniluoto (1999), Peirce has an important position in the history of scientific realism. Deely's other key figure of "postmodernism", Martin Heidegger, is certainly anti-modern in several senses. Heidegger's tendency of going back to the archaic roots of philosophical and theological thinking, his anti-humanist outlook, and his scepticism about the project of Enlightenment and scientific-technological progress suggest that he should be placed within the camp of *pre*moderns rather than *post*moderns. This is reconfirmed by the observation that his account of language and truth (i.e. *aletheia* as "disclosure of being") is not based upon the semiotic notion of representation. Further, given that Heidegger's thought is committed to the universal medium view of language, Heidegger and Peirce appear to be in opposite camps.

Peirce introduced several threefold divisions of signs, so that on the whole his system is quite complex. But his semiotics is "minimalist" in the sense that its taxonomy of signs includes as special cases different grounds for denotation or the sign – object relation (*CP* 2.247-249): an *index* refers to its object by

---

[14] See Deely (1994), 20.
[15] See Niiniluoto (2008).
[16] See Cobley (2009), 4.

virtue of a causal connection (e.g. smoke is a sign of fire), an *icon* by its likeness or similarity with the object (e.g. a picture of a cat is an icon of a cat), and a *symbol* by a rule or convention (e.g. the word "cat" is a symbol of a cat).

Note that this classification of signs is not exclusive, since it allows mixed cases. Photographs are indexical, as they are causally produced by their objects, and at the same time iconic, as they are similar to their objects. Theoretical terms in science may have an indexical element, as their descriptions sometimes employ causal terminology (e.g. electrons leave a trace in Wilson chambers), but still such terms are conventional symbols.

The flexibility of Peirce's approach is seen in the possibility of applying it to a large class of cultural entities which are not signs in any ordinary sense (e.g. buildings, furniture, clothes, and household tools). Such *product semantics* takes its starting point in Eco's suggestion that a sign has its own object as its "primary denotation" and the object of its interpretant sign as its "secondary denotation" or "connotation". Following the idea of functionalism ("the form follows function"), a product of industrial or architectural design may denote its own function. For example, a knife has to be sharp but also look like a cutting tool, the handle of a motor saw should show how one grips it, and a Gothic cathedral expresses the relations of humble believers to their mighty god.[17] A screwdriver shaped as the Eiffel tower serves as iconic connotation to a tourist trip to Paris.

When Goodman (1976) claimed that resemblance is not sufficient for representation, for Peirce's followers he is right in the trivial sense that a two-place relation between two entities cannot establish a representative relation without the activity of the community of interpreters. For example, as Short (2007, 215) remarks, identical tokens of a letter do not represent each other, unless they are so interpreted. In the Peircean approach, a painted portrait is an icon of an actual person (whom it may fail to resemble) only when accompanied by an indexical sign, such as an attached label, or in respect to the history of its

[17] Vihma (1995) applies Peirce's semiotics to steam irons and bicycle helmets.

production.[18] Further, it does not even make sense to ask whether the relation of representation is reflective and symmetric, since it is not dyadic. The role of interpretative cultural codes is highlighted in Umberto Eco's (1976) sophisticated discussion of pictorial representation. For similar reasons, it is misleading to treat isomorphism simply as a two-place relation between two structures M and R, since an isomorphism always presupposes as the third component a key or mapping between the corresponding elements of M and R. This was clearly recognized by Erik Stenius (1964) in his careful explication of the picture theory of language in Wittgenstein's *Tractatus*.

Peirce's semiotic lesson can be seen in the recent "pragmatic" accounts by philosophers of science. Using the slogan "no representation without representers", Ron Giere (2006) replaces the dyadic relation "X represents W" by "S uses X to represent W with purpose P".[19] Bas van Frassen (2008), whose "empiricist structuralism" requires that empirical substructures of theoretical models should be isomorphic to "data models" and "surface models" of observable phenomena, uses a triadic account including a representational structure X, a target structure W, and a user, and it is up to the user's decision that X represents W. Valeriano Iranzo (2014) protests that appeal to user's decision is not sufficient and should be constrained by an additional condition that a representational model in science should be able to convey some truthful knowledge about the purported target.

The distinction between indices, icons, and symbols shows that all representation need not be iconic or based upon similarity. When Goodman (1976) and others argue that similarity is not necessary for representation, Peirce's semiotics agrees: smoke does not resemble fire. Goodman's (1972) more specific objections to the notion of similarity are answered in Niiniluoto (1987), 35-38. Goodman claimed that "similarity cannot be equated with, or measured in terms of, possession of common characteristics". But his thesis that "any two things have exactly as many properties in common as any other two" is based on a nominalist account of properties in terms of all

[18] See Short (2007), 216. Cf. Blumson (2014).
[19] See also Mäki (2009).

Boolean combinations of predicates. The problem disappears when degrees of similarity are defined relative to a selection of respects which are taken to be relevant to the comparison.[20] As representation without resemblance or "mirroring nature" is possible, Rorty's anti-representationalism is based on a too narrow notion of representation. Still, as Park (2014) argues, resemblance relations may have interesting roles in semiotic systems.

All representation need not be based on causal relations, either. Peirce's symbols, among them words in natural language, are conventions by the linguistic community: the words "cat" in English and "kissa" in Finnish do not resemble cats as animals, and they need not have causal relations to cats. Still these words have a conventional meaning which allows them to refer to cats. Fodor (1987), who thinks – with many physicalist cognitive scientists - that all signs should be icons (the resemblance theory) or indices (the causal theory), tries to analyze symbols by a relation between tokens and their causes: the tokenings of the symbol "A" are nomologically dependent of the instantiations of the property A. For example, the mental use of the token "cat" is caused by the presence of cat. This idea is related to causal theories of perception (even though the choice of the token "cat" in English, instead of "kissa" in Finnish, depends on the cultural context). But this account is not applicable even to typical terms in natural language, which can be thought and used in many contexts without the presence of their referents – the power of symbolic languages is that they allow us to go beyond "here and now" and to refer to past, present, future, and possible objects. Nor is it promising in the context of science and art, where symbols are used to refer to unobservable and fictional objects. Among the problems of this account is its rejection of the possibility of misrepresentation.[21]

Misrepresentation, Truthlikeness, and Caricatures

While "reference" and "denotation" are success terms in ordinary discourse, an adequacy condition for a theory of representation is that it should be able to give an account of various

[20] See Niiniluoto (2013).
[21] See Park (2014).

kinds of misrepresentation. Pero and Suárez (2016) distinguish mistargetting (application of a representation to a wrong object or target) and inaccuracy, where the latter may involve abstraction (neglecting some features of the target), pretence (ascribing features not possessed by the target), and simulation (both abstraction and pretence). They carefully argue that the notion of homomorphism does not allow weakenings which would help to understand the phenomenon of misrepresentation. This argument is incomplete, as it does not consider partial isomorphisms which structural realists have used to explicate the notion of partial truth,[22] and thereby to cover at least abstraction as a form of inaccuracy. But while this specific argument about morphisms may be otherwise right, it does not show that all similarity-based treatments of representation are mistaken.

Park (2014) observes that problems of misrepresentation in art and science are similar in interesting ways. He refers to Hopkins (1998) and Blumson (2009) who have discussed examples of pictorial misrepresentation by examples of caricatures and police depictions of dangerous criminals. It is a remarkable coincidence that I have used precisely the same examples in my account of reference by truthlike scientific theories in Niiniluoto (1997). Indeed, the notion of truthlikeness, as applied to idealized scientific theories, gives a useful and viable approach to misrepresentation, and thereby helps to defend critical scientific realism.[23]

Peirce did not speak of misrepresentation in his semiotics, but as a fallibilist he was aware that many scientific theories are erroneous. For him, science is a "self-corrective process" (*CP* 5.575) which leads different minds from antagonistic views to "one and the same conclusion" (*CP* 5.407). This ideal limit of the scientific opinion is what is meant by truth (*CP* 5.565). In particular, induction pursues "a method which, if duly persisted in, must, in the very nature of things, lead to a result indefinitely approximating to the truth in the long run" (*CP* 2.781). W. V. O. Quine (1960) objected that Peirce's characterization of truth as the limit of inquiry is mistaken, since the notion "nearer than" is defined only for numbers but not for

[22] See Da Costa and French (2003).
[23] See Niiniluoto (1987, 1999).

theories. In the same year, probably inspired by Quine's critique of Peirce, Karl Popper started to develop his comparative notion of verisimilitude for scientific theories.[24]

Popper thus suggested that it makes sense to say that one scientific theory is "closer to the truth" than another. When his explication of this notion of *truthlikeness* or *verisimilitude* was refuted by David Miller and Pavel Tichý in 1974, a new "similarity approach" was started. It defines the degree of truthlikeness of a theory T by the requirement that the complete states of affairs ("possible worlds") allowed by T are similar to the true state. More precisely, a theory T in language L is truthlike if it is similar to the complete truth C* (in so far as it is expressible in L) about a real system W, or if the models M of T are similar to W (see Fig. 2). (The latter condition is equivalent to Giere's (1986) condition, even though he does not want to use the "hybrid" concept of truthlikeness.) This notion does not reject the absolute notion of truth (explicated by Tarski's model theory) but presupposes it: a theory T in language L is maximally truthlike if and only if it is equivalent to the complete truth in L.

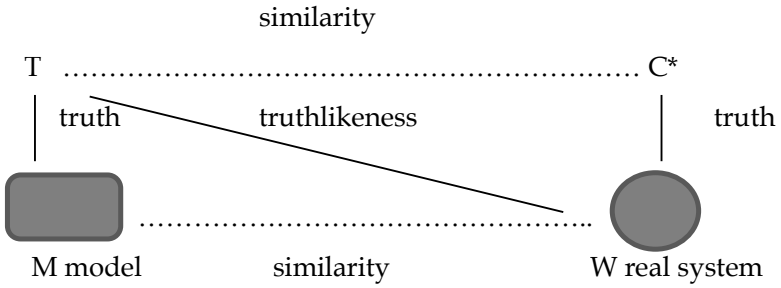

Fig. 2. Truth + Similarity = Truthlikeness

This similarity account allows us to say that one theory is a better representation of reality than another. In particular, it can be applied to *idealized* theories and models which either ignore or distort some relevant aspects of the target: for example, Boyle's law is true about ideal gas but only truthlike about real

[24] See Popper (1972), 231.

gas.[25] So, a theory or a model as a whole is compared to the real system by analogy or resemblance,[26] sometimes using mathematical equations and diagrams, even if the terms and functions used by such theories are symbols rather than icons in Peirce's semiotics. In fact, Pero and Suárez's (2016) talk about positive and negative analogy can be directly associated with Amos Tversky's treatment of *degrees of similarity* by "feature matching", which has recently been used to measure the similarity between scientific models and targets and the truthlikeness of scientific theories.[27] The roots of this idea go back to John Stuart Mill's 1843 treatment of analogy, which Peirce elaborated in 1883 in his treatment of hypothetical analogical inference (*CP* 2.706). According to Peirce, the probability of such inference depends on the "r-likeness" of two objects a and b, defined as the ratio of the number of shared properties of a and b to the number of all properties of b.[28]

Scientific theories, which describe domains of entities with their properties and lawful relations, are often compared to maps which more or less successfully represent some selected structural relations of landscapes or territories.[29] For example, a map of London subway describes correctly the available routes of trains, but the locations and distances between stations are only approximate. The notion of truthlikeness allows us to treat maps in the realist way as truthlike descriptions of reality. In the same way, theories can be compared to portraits and drawings as depictions of human subjects. No two-dimensional portrait or three-dimensional sculpture is identical with its object, nor a completely accurate representation of its object, since it always involves a selection of features. In addition to mere likeness of facial structure and expression, a good portrait (even a photograph) uses perspective, colors, lights, and other contextual aspects to illuminate the mood and personality of its target. But still portraits are similar to their objects, so that even computers can run algorithms of face recognition

[25] See Niiniluoto (1986).
[26] See Niiniluoto (2012, 2013).
[27] See Niiniluoto (1987), 33-34; Weisberg (2013); Cevolani, Festa, and Kuipers (2013).
[28] See Niiniluoto (1987), 24.
[29] See Giere (1986).

which identity persons by comparing digital pictures with pre-existing patterns.

Scientific idealizations can be compared to *caricatures* which playfully and amusingly select and exaggerate some features of their targets.[30] As caricatures to some extent misrepresent their targets, their ability to refer to their targets is denied by Fregean descriptive theories of reference, which require that a theory can refer only to those entities which it correctly describes. This theory of reference led Thomas Kuhn to his dramatic antirealist conclusions: false theories do not refer to real entities, and theory-changes involve world-changes. However, if we adopt a principle of charity to the effect that a theory refers to those objects which it describes in the most truthlike manner, then such caricatures can refer to their targets.[31] More precisely, a term t in a scientific theory T refers to the object b which maximizes the degree of truthlikeness of T as applied to b. This modified account of descriptive reference includes both singular reference to particulars and general reference to kinds. The possibility of reference failure or mistargetting is explained by choosing a threshold or a lower value for the required degree of truthlikeness: reference is successful only if the fit is good enough.

Park (2015), who endorses and elaborates this idea, notes that Ernst Gombrich devoted an entire chapter of his *Art and Illusion* (1960) to caricatures. It is well known that Gombrich was influenced by Popper's philosophy of science, but his classic work was published just before Popper started to develop his ideas about verisimilitude. Moreover, Popper's attempted explication of this notion was based upon the idea of overlap of sets of sentences – without invoking the concept of similarity. So it is not clear whether Popper and Gombrich influenced each other in this matter.

Goodman (1976) does not mention caricatures in his attack against "the copy theory of representation", even though caricatures avoid the objection that resemblance is reflexive and symmetric. With some possible exceptions (like Donald Trump), no man is a caricature of himself, and Barack Obama is not a caricature of any of his caricatures. But Goodman cites

---

[30] See Niiniluoto (1999), 128.

[31] See Niiniluoto (1997).

Gombrich's (1960) observation that there is no "innocent eye". So there is a problem of specifying the target of pictorial representation, since it is "not an object the way it is, nor all the ways it is, nor the way it looks to the mindless eye".[32] But it is a virtue of caricatures as an example of iconic signs that they make the issue of resemblance vividly visible. Caricatures are not copies like fingerprints and DNA samples used as forensic evidence, since they are intentionally selected and exaggerated – involving abstraction, pretence, and simulation in the sense of Pero and Suárez (2016). But for many caricatures we can easily identify their targets, who typically are well-known politicians or other public figures. In spite of the boosted nose, eyes, ears, mouth, or jaw, we know well enough how they "really" look like – either by acquaintance or by photos. Even popular caricatures of Charles Darwin as an ape sufficiently preserve his facial features so that it is easy to recognize this great scholar. But in the case of theoretical terms in science (e.g. electron, gene, quark, Higgs field), the target is not yet known, except some partial, uncertain, and indirect evidence. Therefore, as scientific theories are fallible attempts to describe some so far unknown theoretical entities on the basis of incomplete information, they can be compared to the wanted-for posters of unknown criminals sometimes published by police on the basis of partial information by eyewitnesses.[33] Sometimes such pictures are successful, so that the criminal is correctly identified and captured, but they can also be misleading and direct the search to a wrong person.

The caricature theory of reference shows that there is an important analogy between scientific idealizations and pictures in art. At the same time this theory reaffirms the point that representation and misrepresentation by similarity relations has an important role both in science and art.

Note. This chapter is based on my article "Remarks on Representation and Misrepresentation", *Estudios Filosóficos* 69 (2020), 253-264.

[32] Goodman (1976), 9.
[33] See Niiniluoto (1999), 132; Blumson (2009).

Literature

Baudrillard, Jean (1983). *Simulations*, New York: Semiotext.

Blumson, Ben (2009). "Images, Intentionality and Inexistence", *Philosophy and Phenomenological Research* 79, 522-538.

Blumson, Ben (2014). *Resemblance and Representation: An Essay on the Philosophy of Pictures*. Cambridge: Open Book Publishers.

Brzezinski, J., Coniglione, F., Kuipers, T. A. F. & Nowak, L. (eds.) (1990). *Idealization I: General Problems.* Rodopi: Amsterdam.

Cevolani, Gustavo, Festa, Roberto & Kuipers, Theo (2013). "Verisimilitude and Belief Change for Nomic Conjunctive Theories", *Synthese* 190, 3307-3324.

Cobley, Paul (2009). *Realism for the 21st Century: A John Deely Reader*. Scranton: Scranton University Press.

Culler, Jonathan (1976). *Saussure.* Glasgow: Fontana/Collins.

Da Costa, Newton C. A. & French, Steven (2003). *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning*. Oxford: Oxford University Press.

Deely, John (1994). *New Beginnings: Early Modern Philosophy and Postmodern Thought*. Toronto: University of Toronto Press.

Eco, Umberto (1976). *A Theory of Semiotics*. Bloomington: Indiana University Press.

Fodor, Jerry (1987). *Psychosemantics*. Cambridge, MA: The MIT Press.

Giere, Ron (1986). *Explaining Science: A Cognitive Approach*. Chicago: The University of Chicago Press.

Giere, Ron (2006). *Scientific Perspectivism*. Chicago: The University of Chicago Press.

Gombrich, Ernst (1960). *Art and Illusion: A Study of Psychology of Pictorial Representation*. London: Phaidon Press.

Goodman, Nelson (1972). *Problems and Projects.* Indianapolis: The Bobbs-Merrill Company.

Goodman, Nelson (1976). *Languages of Art*. Indianapolis: Hackett.

Hintikka, Jaakko (1997). *Lingua Universalis vs. Calculus Ratiocinator: An Ultimate Presupposition of Twentieth-Century Philosophy.* Dordrecht: Kluwer.

Hopkins, Robert (1998). *Picture, Image and Experience.* Cambridge: Cambridge University Press.

Iranzo, Valeriano (2014). "Models and Phenomena: Bas van Fraassen's Empiricist Structuralism". In Wenceslao Gonzalez (ed.), *Bas van Fraassen's Approach to Representation and Models in Science*, Dordrecht: Springer, 63-76.

Knuuttila, Tarja (2009). "Isolating Representations Versus Credible Constructions? Economic Modelling in Theory and Practice", *Erkenntnis* 70, 59-80.

Mäki, Uskali (2009). "MISSing the World: Models as Isolations and Credible Surrogate Systems", *Erkenntnis* 70, 29-43.

Niiniluoto, Ilkka (1986). "Theories, Approximation, Idealizations". In Ruth Barcan Marcus, G. J. W. Dorn & Paul Weingartner (eds.), *Logic, Methodology, and Philosophy of Science*, vol. vii, Amsterdam: North-Holland, 255-289. (Reprinted in Brzezinski et al. (1990), 9-57.)

Niiniluoto, Ilkka (1987). *Truthlikeness*. Dordrecht: D. Reidel.

Niiniluoto, Ilkka (1997). "Reference Invariance and Truthlikeness", *Philosophy of Science* 64, 546-554.

Niiniluoto, Ilkka (1999). *Critical Scientific Realism*. Oxford: Oxford University Press.

Niiniluoto, Ilkka (2008). "Semiotics and Philosophy: Comments on Deely". In Robert S. Hatten, Pirjo Kukkonen, Richard Littlefield, Harri Veivo & Irma Vierimaa (eds.), *A Sounding of Signs: Modalities and Moments in Music, Culture, and Philosophy - Essays in Honor of Eero Tarasti on his 60th Anniversary* (Acta Semiotica Fennica 30). Helsinki: International Semiotics Institute at Imatra, 338-346.

Niiniluoto, Ilkka (2012). "The Verisimilitude of Economic Models". In Aki Lehtinen, Jaakko Kuorikoski & Petri Ylikoski (eds.), *Economics for Real*, London: Routledge, 65-80.

Niiniluoto, Ilkka (2013). "Models, Simulations, and Analogical Inference". In Vassilios Karakostas & Dennis Dieks (eds.), *EPSA 11: Perspectives and Foundational Problems in Philosophy of Science*, Cham: Springer, 19-27.

Niiniluoto, Ilkka (2014). "Representation and Truthlikeness", *Foundations of Science* 19, 375-379.

Park, Woosuk (2014). "Misrepresentation in Context", *Foundations of Science* 19, 363-374.

Park, Woosuk (2015). "Towards a Caricature Model of Science". In Lorenzo Magnani, Ping Li & Woosuk Park (eds.), *Philosophy and Cognitive Science II: Western & Eastern Studies*, Cham: Springer, 77-93.

Peirce, Charles S. (1931-35). *Collected Papers*, vols. 1-6, ed. by Charles Hartshorne & Paul Weiss. Cambridge, MA: Harvard University Press.

Peirce, Charles S. (1958). *Collected Papers*, vols. 7-8, ed. by Arthur Burks. Cambridge, MA: Harvard University Press.

Pero, Francesca & Suárez, Mauricio (2016). "Varieties of Misrepresentation and Homomorphism", *The European Journal for Philosophy of Science* 6, 71-90.

Price, Huw (2010). *Naturalism Without Mirrors*. Oxford: Oxford University Press.
Popper, Karl (1972). *Objective Knowledge*. Oxford: Clarendon Press.
Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: The MIT Press.
Rorty, Richard (1980). *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
Rorty, Richard (1982). *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press.
Rorty, Richard (1991). *Objectivity, Relativism, and Truth.* Cambridge: Cambridge University Press.
Short, T. L. (2007). *Peirce's Theory of Signs.* Cambridge: Cambridge University Press.
Stenius, Erik (1964). *Wittgenstein's Tractatus: A Critical Exposition of Its Main Lines of Thought*. Oxford: Blackwell.
Suárez, Mauricio (2004). "An Inferential Conception of Scientific Representation", *Philosophy of Science (Symposia)* 71, 767-779.
van Fraassen, Bas (2008). *Scientific Representation: Paradoxes of Perspective.* Oxford: Oxford University Press.
Vihma, Susann (1995). *Products as Representations: A Semiotic and Aesthetic Study of Design Products*. Helsinki: University of Art and Design.
Wallis, Brian (ed.) (1984). *Art After Modernism: Rethinking Representation.* New York: The Museum of Contemporary Art.
Weisberg, Michael (2013). "Getting Serious about Similarity", *Philosophy of Science* 79, 785-794.
Wise, M. Norton (2005). "Realism is Dead". In Martin R. Jones & Nancy Cartwright (eds.), *Idealization XII: Correcting the Model. Idealization and Abstraction in the Sciences*, Amsterdam: Rodopi, 269- 285.

Chapter 5: Film and Reality

On his first trip to Leningrad, a Finnish film critic arrives at the square of the Winter Palace, the famous scene of the October 1917 revolution, stops and enthusiastically sighs: "This is the celebrated square where Sergei Eisenstein in 1927 filmed his ingenious *October*."

This true story, for its part, throws light on the confusing relation between film and reality. Film has become an inseparable part of the modern world, and it is — perhaps more effectively than any other form of art — able to record and capture reality, on the one hand, and to create an illusionary and fictive reality, on the other. The narrative language of film, based on picture and sound, enables one to make denser representations of the world - variously characterized as "dream-like", "greater than life", or even "more real" than the world itself.

The diverse relation of film and reality is, indeed, a challenging and controversial basic question of film studies. Philosophy can provide conceptual tools for clarifying it. Simultaneously, philosophy gives us an excellent opportunity to more generally illuminate the questions of what kind of entities works of art are and how they can function as signs referring to reality.

Film Theory

My subject in this chapter is so wide that it is, in a way or another, related to all problem areas of cinema studies. Therefore, before beginning, I must make some limitations.

Firstly, I am not going to deal with those scientific regularities and technological inventions through which the possibility of photography and motion pictures has become actualized. Secondly, I am not going to consider those regularities of perceptual psychology which enable human beings to see a film of 24 pictures per second projected onto a screen as a uniform

movement or sequence of events. Thirdly, I am not going to talk about film in a general way as a part of social reality — about cinema culture belonging to the object of study of history and social sciences (film production, film-makers, distribution, cinemas, film shows, cinemagoers, film magazines, film critics, social conditions and effects, etc.). Fourthly, I am not going to discuss in detail the specific ways of expression used in the language of film, the interpretation of the content of films, or their aesthetic value.

What, then, remains of film study, if technology, psychology, history, sociology, grammar, aesthetics, and criticism are skimmed off? One could express the answer by means of the words *the philosophy of film*, but it is more usual to talk about *film theory* in this context — like, for example, Dudley Andrew in his excellent reviews *The Major Film Theories* (1976) and *Concepts in Film Theory* (1984).

The task of film theory is to present an analyzed and argued conception of what kinds of entities films are. Thus, it attempts to answer André Bazin's (1973) question: what is a film? Since films as works of art are more like processes than objects, we can also follow Nelson Goodman (1976) in presenting the question in another form: how is a film? These problems concerning the "essence" or "way of being" of films belong to *the ontology of film*.

With regard to technology and expression, film is a constantly changing and developing historical phenomenon. Therefore, an attempt to find a definition of the essence of film has been regarded as a fallacious "essentialism" — as "chasing wild geese".[1] In the ontology of film, instead of searching for timeless essence, one should ask what individual films are: what kind of entities are, for example, Murnau's *Tabu*, Chaplin's *Modern Times*, and Hitchcock's *Vertigo*?

The representative and communicative function is essentially related to films as cultural objects: as works of art in general, films function as signs,[2] which refer to a reality "external" to themselves (or intertextually to other films). Thus, in *the*

---

[1] Jarvie (1987), 157.

[2] Goodman (1976) states that works of art function as "symbols", but I myself prefer to use this term in a stricter Peircean sense.

semiotics of film, which is a part of film theory, films are conceived as significant "texts" constructed in the language of film.

The problems of the ontology and semiotics of film are already apparent in the usual distinction between animations, documents, and fictions. As a cartoon, animation differs from the forms of film based on photography: the producing of the film material itself is different. Document and fiction, again, are distinguished from each other on the basis of their relations to reality.

The "issue of realism" connected with the latter distinction is, however, problematic in many ways. It is customary to add to dramatized films, already at the production stage, a text according to which "the events and characters of this film are fictive, and their possible similarities to real people, past or present, are accidental, unintended". The relation of film and reality is even more complex than in the case of novel: even in a purely fictive film, the camera in a sense "documents" the "reality" of actors and stage setting in front of it.

A further complication about the realism problem arises, when we relate the content of films to the temporal distinction between the past, the present, and the future (cf. Chapter 1 above). Documentary films contribute to the task of remembering and preserving culture, but also fictive historical films are often realistic enough to help in interpreting and understanding the past. Edwin Laine's 1956 film on "the unknown soldier" based on Väinö Linna's novel has deeply shaped the national conception of the war between Finland and Soviet Union in 1941-44. On the other hand, D. W. Griffith's *The Birth of a Nation* (1915), in spite of being an artistically innovative and groundbreaking historical drama, gives a misleading and racist description of the American Civil War. Fictive films can also contribute to our expectations about the future. Fritz Lang's *Metropolis* (1926) is a visionary story about a possible future society. Its main philosophical theme – the relation between mechanical machines, human persons, and big cities – is closely linked with the modernist cultural and ideological trends of the 1920s, but it also anticipates the iconography of androids, cyborgs, and dark cities in recent movies and rock videos. Lang's themes are strikingly up-to-date, since new

information technologies and robotics are transforming what used to be fictional machines into reality.

The Philosophy of Film

The ontological and semiotic questions in film theory are not only topics of academic discussion. They also have an important influence on producing and assessing films. This is demonstrated by the various programmatic statements that great directors (among others, Dziga Vertov, Sergei Eisenstein, Robert Flaherty, Roberto Rossellini, Eric Rohmer, Pier Paolo Pasolini, Andrej Tarkovsky, and Jean-Luc Godard) and critics (André Bazin) have made on the realism debate.[3]

As a rule, the discussion on the relations of film and reality is reducible to principal philosophical problems.[4] This is especially clearly visible in Andrew's works, whose central theoreticians base their views on a philosophical doctrine or a philosophically interpreted research program of cultural study: Hugo Munsterberg and neo-Kantianism, Rudolf Arnheim and Gestalt psychology, Sergei Eisenstein and dialectical materialism, Béla Balász and formalism, André Bazin and catholic existentialism, Siegfrid Kracauer and naturalism, Christian Metz and semiotics, Henri Agel and phenomenology. Since France has been the center of film theory, the discussion of the past few decades has mainly relied on the traditions and new trends of Continental philosophy — Marxism, psychoanalysis, hermeneutics, phenomenology, existentialism, structuralism, semiology, post-structuralism, deconstruction.

On the contrary, my own field, the so-called analytic philosophy, has achieved almost nothing interesting with regard to film. If the subject has been touched instead of the usual silence, the tone has been negative: Gilbert Harman condemns the semiotics of film as fruitless;[5] Roger Scruton (1983) considers photography and film as "pornographic" media, which

---

[3] Useful anthologies on this issue are Mast and Cohen (1979) and Williams (1980). See also Armes (1974).

[4] As an old cinema enthusiast from my student days, who has since then become a professional philosopher, I have been surprised to see how purely film theory is *included* within a philosophical debate.

[5] See Mast and Cohen (1979).

cannot create representative art — except for the plot of a film. There are, however, positive exceptions: a couple of articles by Arthur Danto, Stanley Cavell's *The World Viewed* (1979), and, especially, I. C. Jarvie's versatile *Philosophy of the Film* (1987).[6]

The disinclination and incapability of traditional analytic philosophy to approach the problems of film can perhaps be understood: there are other reasons in addition to personal lack of interest or the small valuing of film among forms of art. Before Charles W. Morris in the United States in 1930s revived semiotics exposed by Charles S. Peirce at the end of the 19th century and Nelson Goodman published his influential work *The Languages of Art* (1968), analytic philosophers primarily concerned with the philosophy of language did not have any such notion of *language* available which could be useful in dealing with film.

Film as a language differs in a significant way from natural language, as has been emphasized from the 1960s by Christian Metz, who has applied de Saussure's semiology into film.[7] An individual shot does not correspond to a word of a language but to a sentence: a picture of a revolver must be understood as the sentence "There is a revolver here". Film is not a *langue* in Saussure's sense — i.e. a linguistic system which would have its smallest significant elements corresponding to the phonemes and morphemes of natural language.

On the other hand, film does not structurally correspond to the formal languages of logic, in which the analytic tradition of Frege — Russell — Wittgenstein — Tarski — Carnap has sought its models of language. These languages have been assumed to possess a well-defined syntax (grammar) and a semantics following "Frege's principle": the meaning of a sentence is a function of the meanings of the words occurring as its parts. This kind of simple principle of compositionality is not applicable to the language of film, which utilizes film editing: the central insight of the montage theory proposed by Kulesov and Eisenstein in the 1920s is precisely the context-dependency of pictures, the dependency of the meaning of a

---

[6] The collection Freeland & Wartenberg (1995) is a mixture of different philosophical approaches.

[7] See Mast and Cohen (1979).

picture on other pictures in relation to which it is presented.[8] The same remark is strengthened by the multifarious nature of the language of film: simultaneous effects of sound and music influence the meaning of pictures. For these reasons, the seemingly interesting attempt to find a parallel between the pictorial language of film and the "picture theory of language" of Wittgenstein's *Tractatus* (1921) does not work.

However, in my view, Metz's thesis that the language of film has no negation is far less significant. This thesis itself is questionable: a film as interpreted by a spectator can by different means tell and show that, for example, the man is *not* young anymore, the wife is *not* at home, the relation of father and son is *not* close. As a part of a sound picture, there may also be a narrator, who possesses the whole variety of natural language expressions. However, if film is in a narrower sense understood as a kind of visual perceptual language, it can be mentioned as an interesting parallel phenomenon that, by the inspiration of logical positivism, formal languages without negation have been studied in logic.

The Realism Debate

The question of the existence of an "external world" — a reality independent of human cognition — and of human possibility of obtaining true knowledge of it has often been considered the "basic question" of philosophy. Ontological and epistemological *realists* give positive answers to these questions, while the answer of *antirealists* (among others, subjective idealists and skeptics) is a negative one.

The leading idea of the Vienna Circle of the 1920s was that the entire realism debate must be rejected as "metaphysics": the theses of both realism and antirealism are meaningless. Thus, the question of the existence of a reality "external" to human cognition is a pseudo-problem caused by misuse of language. Similar criticism of "metaphysical realism" or "objectivism" has also been presented on the basis of Edmund

---

[8] A classic example of the effects of setting pictures next to each other is Leo Kulesov's experiment, in which the expressionless face of an actor creates different impressions (hunger, delight, sorrow), if it is connected with a picture of a plate of soup, a playing child, or a dead woman. See Knight (1957) and Eisenstein (1965).

Husserl's phenomenology: phenomenology "brackets" the external world and focuses on studying human intentional acts and the "life-world" constituted through them and determined by human interests and ways of experiencing.

Analytic philosophers who have feared the infection of metaphysics have, as heirs of logical positivists, avoided the word *reality* itself. The only legitimate form of ontology was, for a long time, Carnap's and Quine's (1969) strategy, according to which questions of existence can meaningfully be talked about only internally to a given linguistic system — we can say that the language of arithmetic contains an "ontological commitment" to the existence of natural numbers, but we cannot claim that natural numbers "really" exist.

This way of thinking has still some influence among analytic philosophers, but the proponents of its later versions are today called, following Hilary Putnam (1981), *internal realists*. For a Putnamian internal realist, epistemology is prior to ontology: the world in itself cannot meaningfully be talked about; truth and reference are epistemic, internal to language; linguistic expressions do not refer to mind-independent reality *an sich*, but only to "world versions" (Goodman), to the-world-as-we-understand-and-linguistically-structure-it.

Inside analytic philosophy, there are, however, new trends which have aimed at rehabilitating the status of ontology. This view is manifested in *critical scientific realism*, which urges that the human interaction with the objective reality enables us to achieve more and more accurate and "truthlike" knowledge that "converges to the truth" about the world. Another phenomenon is the new wave of *analytic metaphysics*, understood as a systematic investigation of the most general concepts that can be used to describe reality.[9]

The realistic program has, in part, gained support from *possible worlds semantics*, developed by Jaakko Hintikka (1969) (and others) from the 1950s. From the 1970s, this method has been applied to, among other things, the semantics of fictive expressions. Along with this development, the philosophy of art has strongly arisen — including the ontology of art.

Analytic philosophy has excelled in enthusiastically aiming at conceptual clarity, but in taking up philosophical questions

[9] See Loux (2007).

it has not been as brave as Continental philosophy.[10] It is very interesting that precisely now that interesting tools and results concerning the realism debate — specifically, also the problem of realism of film — could be found in analytic philosophy, the Continental structuralist and post-structuralist, "postmodern" semiotics has arrived at a contrary situation.

Andrew describes the situation as follows. Contrary to the American trend relying on Peirce, the French semiotics of film "brackets reality": a sign or signifier (*signifiant*) refers to another sign or signified (*signifié*) in human mind or culture, but the discussion of the relation of language and reality has been eliminated. "The pictures and sounds of a film are not only denied to be fragments of reality, but now they don't even refer to reality."[11]

At the background of this situation, there is the manner, typical of European semiotics, of operating with a two-place semantic relation signifier-signified, instead of Peirce's triadic relation sign-object-interpretant.[12] Louis Althusser's Marxism-based thesis that film reflects the dominating bourgeois ideology — and even the film technology is "polluted by ideology" — was also influential in the French discussion of the 1970s.

In my own view, the world is not "lost" (as Richard Rorty says) — reality, reference, and truth are not empty or merely "ideological" notions. The basic question of film theory, the problem of realism, which both logical positivists and post-structuralists have for many times tried to close in the garbage bags of history, must be taken seriously in philosophy.

---

[10] A remarkable example from the 1980s is Gilles Deleuze's application of Henri Bergson's theories of duration (*la durée*) and movement to explain how a spectator perceives time and continuous movement in cinema.

[11] See Andrew (1984), 58.

[12] Cf. Chapter 4 above. Another influence for this attitude may come from the commitment to the conception of language as a "universal medium", instead of the calculus view of language (cf. Chapter 4). For Peirce's semiotics, it is important that he accepted the conception of language as a calculus. Despite his internal realism, Goodman (1976), for whom language is a calculus, seriously analyzes the reference relations of the language of art (denotation, exemplification).

The Ontology of Film

For the consideration of the ontology of film, it is useful to distinguish, following Karl Popper, three realms of reality: *World 1* includes physical entities, events, and processes; *World 2* mental phenomena, states, and events connected with individual psyche; *World 3* products of the mind created and maintained through human social action (cf. Chapter 1 above). Thus, World 1 is identical to the material nature; World 2 includes the consciousness of men and higher animals; World 3, the material and spiritual culture, material and abstract artefacts, cultural entities, and social institutions.

In my view, the most reasonable way of understanding the theory of the three worlds is *emergent materialism*: World 1 is temporally and ontologically the original reality, from which Worlds 2 and 3 have emerged through biological and cultural evolution. The events of World 2 can exist only as supported by World 1. World 3 is human-made, historically changing, and its entities remain real only in so far as they have been "recorded" or "preserved" in World 1 or 2.

For example, the sentence "Marlon Brando is a film actor" as spoken or as a written sequence of signs is a material entity of World 1; the mental image raised by it in my mind belongs to World 2; its semantic meaning, its objective informational content, that is, the proposition it expresses, belongs to World 3.

Materialistic theories of art have aimed at identifying works of art with physical objects of World 1, idealists, for their part, with mental phenomena of World 2. In my view, works of art are typical — though different from each other — examples of World 3 objects.

A painting, a sculpture, and a building are works with which a uniquely determined material object, a World 1 entity, is essentially connected. As a cultural object, da Vinci's *Mona Lisa* is not, however, identical with the physical object in Louvre — as a World 3 entity it possesses all the relational cultural properties, that is, the "meaning" of the object in question, such as the content of the work, its relation to its creator, to the audience, to art and museum institutions, and its economic and aesthetic value.

In the field of industrial art, the design work is the immaterial form or figure, the prototype, from which an arbitrary number of material copies can be reproduced. Tapio Wirkkala's *Vodka bottle* cannot be identified with any specific object of World 1; there is a class of World 1 objects similar to each other, and each of these bottles realizes the "type" (in the sense of Peirce's *type-token* distinction) in question belonging to World 3.

Musical compositions, too, are abstract artefacts that have material tokens in World 1 (a performance as sound waves, radio waves, partitures, records, tape recordings) and mental tokens in World 2 (the ideas of the composer, the experiences and feelings of the listener). The normal way of using language includes an ontological commitment that each musical work of art — to which the composer has a copyright — is one "entity": Jean Sibelius composed *one* violin concerto and *seven* symphonies. This entity belongs to World 3: proposals of identifying it with an object (or a system of objects) of World 1 or 2 are artificial and unsatisfying. For example, if Sibelius's violin concerto were identical with the original partiture, it would disappear as a work of art, if this partiture were destroyed.

Correspondingly, a novel, such as Leo Tolstoy's *Anna Karenina*, is an abstract World 3 entity created by its author. Related to it, there are again mental ideas and experiences of the author and the reader, as well as the linguistic text of World 1 expressing the content of the work, which can be printed in a book, copied, or even translated into another natural language.[13]

The ontological approach presented above can also be extended to the works of cinema art. In this case, the analysis should clarify how, among others, the following sentences can be true: "I have seen John Ford's film *The Searchers*", "I have seen *The Searchers* for nine times", "More than a million Finns were watching the film *The Godfather* on TV on March 3rd, 1987". It is, of course, possible to claim that these sentences are, literally interpreted, false: the work *The Searchers* does not exist as a singular entity. I myself believe, however, that postulating

---

[13] Roman Ingarden's ontology of novel, in which a novel is a multi-layered ideal intentional object that is constituted by the reader, can be mentioned as an interesting comparison. Cf. Mitscherling (1985).

cultural entities in World 3 is ontologically more economical, theoretically more fruitful, and practically more natural than the attempt to artificially reduce them to Worlds 1 or 2.

A set of physical objects and events in World 1 (filmstrips, shows, pictures and sounds projected onto the screen) and of mental phenomena in World 2 (the intentions and plans of the makers of the film, the perceptual experiences of the spectator, an experienced meaning, an understood content) is related to each film. But, again, we can claim that *The Searchers* is a work belonging to World 3. It is a collective product, and, among others, the screenplay writer, the director, the actors, the composer, and the editor have contributed to its creation. The work has become ready in 1956 in the moment when the film material shot by the direction of Ford has in the editing table been united into a singular whole, to which the title *The Searchers* has been given. The work cannot, however, be identified with the original filmstrip — in this case, it would be false to say that I have seen the film *The Searchers*, because, in reality, I have only seen different copies or reproductions of the original strip. It is equally impossible to say that each copy of the film, separately, is the work *The Searchers*: although the copies I saw in the 1960s and in the 1980s were different physical objects, I have seen the *same* film for nine times. The film *The Searchers* cannot be identified with the mental process of the spectator, either: the nine different watchings have given me nine *different* experiences of the same film. Correspondingly, in the same show, different spectators have different experiences of the same film. And the film as a work of art exists even in those moments when no one is watching it.

Film has its special ontological status, however, compared to novel, for example. A written text expresses its information content on the basis of conventions accepted by the linguistic community; hence, a novel published as a book expresses its content even at the moment when nobody perceives it. The novel's relative independence of perception is also apparent in the fact that its meaning can be mediated through visual, auditory, and (in the case of braille) tactile senses.

However, a film as a meaningful cultural object does not get expressed merely by means of the *pictures* of the filmstrip; rather, it is realized or "constituted" as a *motion*-picture only as shown and watched (and heard) in the right way in the right

circumstances: a complex process in Worlds 1 *and* 2 "supports" the film as a World 3 entity. Therefore, *The Searchers* cannot be a type of filmstrips identical to each other. On the other hand, *The Searchers* cannot be a type of events of World 2, either, because watching experiences differ from each other: the screen cannot be perceived in the "right" way in itself; each event of watching contains interpretation of the perceived picture. If we want to say that different spectators see and interpret the same film in different ways, the work in question must be placed in World 3.

Formalism and Realism

André Bazin presented in 1958 his famous distinction between film directors "believing in picture" and "believing in reality".[14] The former are usually called *formalists*, the latter *realists*. As Andrew (1976) notes, the corresponding distinction can also be made among film theoreticians.

The formalist tradition of film begun in the beginning of the twentieth century, when the directors of first dramatized films (George Méliès, Edwin S. Porter, and D.W. Griffith) started to develop a specific expressive language for film narration — angles of view, framings, and editions.[15] The schools of the silent film, according to Bazin, developed these tools into their full perfection. The German expressionists turned film into a decorated photographic art of stage settings, lights, oblique angles of view, close-ups, and plastic composition (Josef von Sternberg). The Soviet school (Sergei Eisenstein, Vsevolod Pudovkin) enriched the expressing power of the language of film, its ability of creating new meanings, through the montage technique and the utilization of metaphors.[16]

Rudolf Arnheim, a classic of formalist film theory, saw in 1932 the heart of film to lie in how picture distorts the perceived object: precisely the limitations of the picture (such as framing, two-dimensionality, colorlessness, soundlessness) make possible active selection and the "creative organizing of the raw sensual material", through which cinema art is

---

[14] Bazin (1967-71).

[15] See Martin (1968).

[16] See Lotman (1976).

realizable.[17] Arnheim's view is an extension of the formalist theory of photography, according to which photographs — through the choice of framing, angle, hole, lighting and paper — are works of art comparable to paintings, in which the conscious changing of reality, not its mechanical copying or representing, is the measure of artistry.[18]

Bazin opposes formalism with the realistic view that "the meaning of a picture lies in what it *reveals* from reality, not in what it *adds* to it". During the period of the silent film, the realistic trend was represented, for Bazin, by Robert Flaherty's documents and Erich von Stroheim, who was the most ardent opponent of "picture expressionism and artificial montage" and in his film *The Greed* (1924) "views the world so closely and intensively that its ugliness and cruelty are in the end revealed". The silent film was, however, a "cripple" compared to the more realistic sound film, whose victories of artistic maturity for Bazin are, among others, Jean Renoir's "poetic realism", John Ford's westerns, the deep and accurate description of Orson Welles's *Citizen Kane* (1941), and Italian neorealism (Roberto Rossellini).

Reality as a Passion

Film realism in Bazin's sense can be seen as a continuation of the realistic theory of photography: in photography, camera is the medium through which reality itself is reproduced or copied on the strip.

Edgar Allan Poe admired in 1840 the daguerreotype technique, which had been invented in the preceding year, since by means of it we reach "reality better than by any other

---

[17] See Arnheim (1958); Mast & Cohen (1979).

[18] In this context, Arnheim himself talks about differences between "those figures of (mental) images that we reach when watching the real world, and those that we realize when watching the screen". There is, however, no unique "given" way of "pure" watching of reality (cf. Husserl's thesis about the intentionality of perception, Wittgenstein's notion "seeing as"); the watching of pictures is a complex perceptual event which presupposes learning (cf. Gombrich, Hochberg and Black, 1972). In defining formalism, it would be better, in my view, to talk about how the camera and the photographer can intentionally select and manipulate the physical effects of the object.

means" — the picture is "*infinitely* more accurate in its reproducing ability than any painting by a human hand". The same magical sense of reaching reality, of capturing time and movement, was present in the Lumière brothers' presentation of a train arriving at a station in 1895. This sense of respecting and admiring reality is visible in Flaherty's document of Eskimos, in the news films of Vertov's "kinocs", in the poetic pictures of raindrops on apple blossoms in Alexander Dovzhenko's film *Zemlya* (1930), in Rossellini's film group's rides on the streets of Rome occupied for the last days, or in cinéma-vérité films describing life in cafés of Paris.

With respect to photography, a similar attitude toward reality has vividly been described by Roland Barthes in his book *Camera Lucida* (1981). He begins by telling about his astonishment when seeing the photograph taken in 1852 of Napoleon's brother Jerome: "I am watching eyes which have seen the Emperor". For Barthes, "each photograph is a piece of evidence of presence" — contrary to a painting or a language. The "noema" of photography may, according to him, be called "This-has-been".

Peter Wollen (1977) has criticized this kind of aesthetics of realism of a "monstrous fallacy" that "truth lives in a real world and can be picked by a camera". According to his insightful remark, this realism is an "outgrowth of romanticism" — lacking interest and respect toward scientific knowledge.

The difference between a romantic and a scientific realist can be illuminated by means of the French physicist Henri Poincaré's work *Science et hypothèse* (1902). Poincaré first quotes the romantic Thomas Carlyle's words of how the "God of Things as they are" must be worshipped:

> Nothing but the facts are of importance. John Lackland passed by here. Here is a reality for which I would give all the theories in the world.

The physicist, however, according to Poincaré, would take a different attitude:

> John Lackland passed by here. It is all the same to me, for he will not pass this way again.[19]

---

[19] Poincaré (1952), 141.

Bazin's and Barthes's realism is, without doubt, closer to Carlyle than to Poincaré. In accordance with the Christian-existentialist tradition, the world is, for Bazin, religiously interpreted: reality is the presence of God, and God is the real creator or subject of the film. A similar element of holiness going through reality is visible in Robert Bresson's films.[20]

A skillful maker of films may describe the world as interpreted through myths even in the case he or she does not believe in them. Pasolini's *The Gospel according to St. Matthew* (1964) is a description of the story of the gospel by a non-believer Marxist, together with a 2000-year tradition of myth, which produces a far more effective and genuine religiously charged picture of Jesus than any Hollywood Bible spectacle. Pasolini's films *Oidipus* and *Medea*, which picture the reality of ancient Greek myths, have the same effect.

Hence, realism in film is relative to the creator's conception of reality. Some forms of realism may set the task of art parallel to the one of science as seen by scientific realism: seeking knowledge about reality and controlling reality by means of knowledge.[21] Examples of this could be provided by Kracauer's theory of films "rescuing reality", socialist realism,[22] and the so-called informative conception of art. On the other hand, scientific realism typically sees science as aiming at the *correction*, through theoretical concepts irreducible to observations, of the conception of the world which we obtain in everyday experience and perceptions — thus, science transcends photography and film in so far as they are concerned with recording different experiences of reality related to sense perception. (Infrared, ultrasound, and X-ray pictures are a different matter, though.)

Relativity to the concept of reality makes it possible to interpret many different trends of art as "realistic". For example, impressionism in painting, whose influences are seen e.g. in Renoir's film *Une partie de campagne* (1936), is related to a

---

[20] A parallel but different trend in literature is the so called "magical realism" (e.g. Gabriel Garcia Marquez), which takes the real world to have an undercurrent of supernatural and fantasy.

[21] According to Susan Sontag (1977), photography aims at substitutional possessing, collecting, and controlling of reality.

[22] Cf. Basin (1979).

positivistic or phenomenalistic view of the world as a totality of perceptual experiences. Cubism is connected with the phenomenological doctrine of the constitution of the life-world by the "meanings" or "noemas" of intentional acts;[23] in the field of film, German expressionism can be regarded as a corresponding view.[24] Surrealism can also be understood as a form of radical realism inspired by psychoanalysis, which attempts to — in the manner of Buñuel's *Un chien andalou* (1929) — picture dream-like hallucinations springing from human unconscious mental life (in World 2).[25] Thus, drawing a boundary between formalism and realism in film theory is an extremely problematic matter, which is dependent on our philosophical conception of reality — of ontological and epistemological presuppositions. Bazin's distinction between directors believing in picture and reality is a useful simplification which needs additions and more sophistication.[26]

For example, Roy Armes (1974) divides film into three main lines: (1) realism, i.e. "revealing of reality"; (2) illusion, i.e. "imitation of reality"; (3) modernism, i.e.. "questioning of reality". Realism, in his sense, includes the documentary tradition from Louis Lumière, Robert Flaherty, and John Grierson to contemporary TV realism, Vertov's Kino-Pravda, von Stroheim, Renoir, Rossellini's neorealism, and the cinéma-vérité of the 1960s. Griffith, westerns, Charles Chaplin, the heritage of Hollywood, Alfred Hitchcock, and Walt Disney's animations belong to the tradition of illusion. Representatives of modernism (or, according to other thinkers, "postmodernism") include those who experiment with the expressive possibilities of film: Eisenstein, expressionism, Luis Buñuel's surrealism, Alain

---

[23] See Hintikka (1975).

[24] It has been said that film renders possible the pictorial representation of different perspectives towards an object, at which Pablo Picasso's and George Braque's cubism aimed within the framework of one work. At the background of cubism, there is the idea of the language of art as a calculus (cf. note 12).

[25] The idea of film as a "dream-like representation" advocated by Susan Langer in her *Feeling and Form* (1953) is better applicable to surrealistic films than, for example, ordinary dramatized films using third person narration.

[26] Cf. also Lotman (1976), who creditably pays attention to the "paradoxical" ability of film narration to express stories and ideas by means of a composition of successive iconic signs (framing, editing).

Resnais's films with fragmented time, Jean-Luc Godard, and underground films.

According to radical formalism, a film is a "text" produced by means of a camera (or cartoons); it is analyzed purely syntactically, independently of representational relations of reference. For a realist, the relation of this text to what it expresses externally to itself is important. In this case, we need not be concerned with a "recording" or revealing of the actual world, but with a "visualization" of an imagined world.[27]

In the following sections, I try to approach the problem of realism of film by utilizing the basic concepts of Charles S. Peirce's semiotics (cf. Chapter 4 above).

Film as Index of Reality

In Peirce's semiotics, signs have a triadic structure: they refer to something (the referent) for some persons (the interpreter) with respect to something (cf. Chapter 4 above). The reference of an interpreted sign to its referent may be based on three different grounds: in the case of *indices*, there is a natural causal relation (smoke is a sign of fire, a weather vane is a sign of wind, the odor of a cat is a sign of a cat); *icons* are in some sense similar to their objects (e.g. pictures, diagrams, metaphors); *symbols* rely on a convention accepted by the linguistic community (e.g. words of natural language; the word "cat" in English is a conventional sign of a cat).

Some scholars, like Peters (1981) and Wollen (1977), have tried to analyze the special nature of the language of film by means of Peirce's theory of signs. In my view, Peirce's theory does provide a good instrument for considering the referential relations of films.[28] Armes (1974), following Wollen, even bases his whole tripartition between realism, illusion, and modernism of film on what type of sign (index, icon, symbol) is dominant in each case.

Film realism in the sense of Bazin and Barthes — when animation and other artefactual ways of producing pictures are closed out — is essentially based on the indexical nature of the pictures of a film. The existence of a cinematic picture

---

[27] Cf. Peters (1981), 9.

[28] Cf. Mast and Cohen (1979).

presupposes a real causal relation between the object and the film. A photograph is evidence for the existence of the object in front of the camera at the moment of the shooting of the picture — for what has been here and now. A camera has a "factual present perspective". Hence, a film can be considered a "text", which is in an *indexical* relation to what has happened in front of the camera.

On the other hand, a photograph as a trace of its object, as a projection of reality for us, also expresses the *non-existence* of the object, as Stanley Cavell (1979) especially emphasizes. In this respect, John Wayne on the screen is in a similar status as the products of imagination: if I imagine what my friend David is doing in Berlin just now, David as imagined is, according to Sartre, a part of "nothing".[29] Film is, however, distinguished from imagination — from "fantasy" in Scruton's (1983) disparaging sense — by its indexicality: my perceptions of John Wayne as Ethan Edwards of the film *The Searchers* reduce through the camera to the causal influence of Wayne himself.

Film as Icon

An index need not in any way be similar to its object — smoke does not resemble fire. However, photographs and films have, in addition to their indexicality, the nature of Peirce's *icon*: they represent their objects in the form of *picture*. Through the "eye" of the camera, reality is reproduced on the film, which, when projected onto the screen, produces an observable picture in motion which is similar to its object. Iconicity is, thus, an essential feature of the possibility of film realism.

In semiotics, the concept of icon has been considered problematic because of the difficulties related to, among others, the notion of similarity. In his famous critique of picture, Umberto Eco proposed that icons are also culture-bound representations whose codes must be learned.[30] For analytic philosophy of language, Eco's thesis is not surprising: the picture theory of language sketched in Wittgenstein's *Tractatus* (1921) has been made more precise by utilizing the concept of isomorphic correspondence, which is always relative to a given "key". Such a key is used also in Tarski's model theory as the interpretation

[29] Cf. Niiniluoto (1985).
[30] See Mast and Cohen (1979).

function, which maps proper names to particular objects, one-place predicates to sets of objects and n-place predicates to relations between objects.[31] The notion of similarity (or resemblance), too, can — despite Goodman's nominalist criticism — be usefully defined, even so that we can sensibly talk about *degrees* of similarity.[32] Thus, iconicity is not a matter of either-or. Hence, the distinction between icons and symbols can still be maintained: an interpreted picture (contrary to a conventional sign) refers to its object on the basis of similarity.[33] (In this case, the object is an object or state of affairs in reality, not a perception of reality.)

Eco may be right in claiming that a two-dimensional photograph literally shares no properties with its three-dimensional object. Correspondingly, a filmstrip as a physical object of World 1 is not, in itself, similar to its object. Still, a film or a screen appropriately interpreted may function as an iconic representation similar to reality.

However, the watching experience also belongs to the pictorial nature of film: a film becomes a film only as being watched. Perhaps the perceived film in the mind of the spectator, which belongs to World 2, is an icon of reality? Answers to this question depend essentially on what kind of philosophical theory of perception we advocate. According to representational theories, in perceiving a tree I perceive a sense datum which, in a way or another, represents the tree. However, direct realism (which I myself consider better argued) regards the tree as an external physical object which is the real object of my perception. Depending on its truthfulness, my perception gives more or less adequate information about it.[34]

---

[31] This interpretation function was an innovation introduced by Rudolf Carnap in his formal semantics (see Chapter 16 above).

[32] See Niiniluoto (1988).

[33] A lot of confusion has been caused by the mistaken assumption that being an icon and a symbol would be mutually exclusive properties of a sign. A cartoon of a politician drawn by a caricaturist is a new symbol, a convention established with the reader, in which there is, however, enough similarity to the object, so that the cartoon functions as an immediately recognizable icon (cf. Chapter 4 above).

[34] Direct realism should be distinguished from epistemologically naive realism, since it holds that the content of my perception may depend on different conceptual, theoretical, and practical background conditions.

A realistic theory of perception may seem to support a thesis of the ontology of film, which could be described as *ultrarealism*. Semioticians have sometimes characterized film as a "language without a code" consisting of "natural signs". Pasolini (interviewed by Oswald Stack, 1969) expresses this idea as follows:

> By studying the cinema as a system of signs, I came to the conclusion that it is a non-conventional and non-symbolic language [*linguaggio*] unlike the written or spoken language [*lingua*], and expresses reality not through symbols but via reality itself. If I have to express you, I express you through yourself; if I want to express that tree I express it through itself. The cinema is a language [*linguaggio*] which expresses reality with reality. So the question is: what is the difference between the cinema and reality? Practically none. I realized that the cinema is a system of signs whose semiology corresponds to a possible semiology of the system of signs of reality itself. So the cinema forced me to remain always at the level of reality, right inside reality: when I make a film I am always in reality, among trees and among people like yourself; there is no symbolic or conventional filter between me and reality, as there is in literature. So in practice the cinema was an explosion of my love for reality.[35]

I can perceive a tree directly through my eyes, through a window, as reflected in a mirror, as a photograph through my own camera, or through a film shot by Pasolini's camera — as I can hear the voice of my friend directly, over the telephone, or on the radio. For an ultrarealist, film provides a mediated but still direct connection with reality. In the film *The Searchers*, I see John Wayne, not a picture of Wayne, as I see my wife at home over dinner — the only difference lies in the longer and more complex causal chain from the object of the perception to my perception here and now. The screen is not a painting representing the world or a picture together with its frame, but a "hole to reality" (Jean Mitry).

---

[35] See Stack (1969), 29. For Pasolini's article "The Written Language of Reality", and its critique by Eco and Metz, see Viano (1993). In Viano's post-structuralist reading of Pasolini's semiotics, "cinema is to reality what written language is to oral language".

The worship of reality belonging to ultrarealism is manifested in the documentary tradition of film and in Bazinian aesthetics. Even though it reaches something essential about the nature of film, as a theory of film it has remarkable limitations. A one-sided emphasis on the indexicality of film may lead to a denial of the possibility of the entire film art: for Scruton (1983), photography and film are based on a causal relation, not on intentionality; hence, as representative art they are as impossible as an art of mirror images.

Pasolini's ultrarealism does not take into account the fact that dramatized films are "pictures of pictures of pictures":[36] the camera records a picture of actors visualizing a story which describes a (real or fictive) history. In the film *October*, Eisenstein may denote the square of the Winter Palace by the square of the Winter Palace; however, in his film *Reds* (1981), Warren Beatty denotes St. Petersburg by Helsinki; in his film *The Gospel according to St. Matthew*, Pasolini does not denote Jesus by Jesus, but by a Spanish amateur actor Enrique Irazoqui.

The Untruthfulness of the Screen

In the early 1960s, as a school-boy, I wrote a composition on the subject "The Untruthfulness of the Screen". I paid mainly attention to the tools by means of which a film may "cheat" the spectator by creating an illusion of reality. When I think I see the bleeding villain falling with a thump on the street of a tumble-down western city penetrated by a bullet shot by the hero, as a matter of fact, there "really" was only a bang without a bullet, red tomato ketchup, a stuntman, and paperboard stage setting in Hollywood. The teacher was not satisfied, though he admitted that the subject could be dealt with in this way, too. He had himself thought of the title as referring to how the weepy happy end stories of films give a distorted and romanticized picture of the world and of "real life".

Both views of the untruthfulness of the screen — on the level of the expressing language of film and the content of stories — bring out limitations of ultrarealism.

---

[36] Peters (1981), 20.

Markku Tuuli has emphasized that the shooting of pictures always selects, limits, and distorts reality.[37] Even in the case of documentation, the presence of a camera brings its own unordinary significance to the situation. In film realism, what is essential is not what happens in front of the camera but what is recorded by the camera. For example, shots of different times and places may be united into the same scene on the filmstrip. *Verisimilitude* (in French, *vraisemblance*) is not dependent on what is real but on what the spectator feels as real. In order to achieve a sense of verisimilitude in front of a camera, one often has to manipulate the dramatized reality: a fight must be exaggerated, so as to make it seem genuine.

In Godard's film *Le Petit Soldat* (1963), it is stated that "film is the truth 24 times per second". It must be kept in mind, however, that a picture does not in itself state or assert anything about the real world. The verisimilitude of a fictive film is thus an illusion of reality, an iconic representation of a fictive world.[38] Joseph von Sternberg, a master of stage setting and lights, magically created — for example, in his film *Der blaue Engel* (1930) — by means of film an artificial "dream-world", whereby "the iconic feature of a sign independently of indexicality" is emphasized in film.[39]

The *vraisemblance* of film is thus different from the truthlikeness of scientific realists: the former is verisimilitude relative to a fictive world, the latter relative to the actual world (cf. Chapter 7 above). But even in art, the verisimilitude is reduced, if the story allows its characters to have supernatural powers.

Fiction, Symbols, and Reality

The basic nature of fictive art has impressively been expressed by the Russian filmmaker Andrei Tarkovsky (1987). According to Tarkovsky, a film begins from the moment at which the

---

[37] Unpublished radio broadcast "The truth 24 times per second" (1981). The film critic Markku Tuuli (1945-83) was my closest friend in the student years.

[38] This feature is brilliantly utilized by Akira Kurosawa's *Rashomon* (1950), which tells four different versions of the same events. Cf. Jarvie (1987), who in general is interested in the capabilities of film in expressing philosophical theses.

[39] See Wollen (1997), 86; Armes (1974).

director "sees through his mind's eye the figure of the future film":

> The artist starts, when an idea and a film give rise to a personal pictorial construction and an original way of experiencing the existing world. The director directs this to the judgment of the spectators and shares it with them as if revealing his most secret wishes.

The main idea of film art is "time sculpted in its concrete forms and phenomena", which is loyal to the "truth of life".

Let us consider, as an example of fiction, Francis Ford Coppola's film *The Godfather* (1972), based on Mario Puzo's novel. What kind of referential relations does it have to the world? In Figure 3, the situation is illuminated with regard to the section in which Vito Corleone dances with his daughter in her wedding in New York in 1947.

Firstly, it may be noticed that the pictures on the filmstrip and on the screen as well as the picture experienced by the spectator are in an indexical and iconical relation to the stage set by Coppola in 1971, on which the smiling, made-up Marlon Brando dances with Talia Shire. Thus, the spectator "really" sees Marlon Brando dancing amidst stage setting and other actors. In accordance with the illusion of the film, however, the spectator's experience is wholly different: she sees the Mafia boss Vito Corleone dancing in New York in 1947. The content of her experience thus partially describes a fictive possible world, in which a person named Vito Corleone is dancing in 1947. The spectator's perceptions are not in an indexical relation to this imagined world, which the director Coppola has seen "through his mind's eye". However, the picture Coppola has, by means of his film, conjured up into the spectator's mind is, in a trivial sense, an icon of the fictive New York — though the author Puzo may himself have originally imagined everything differently.

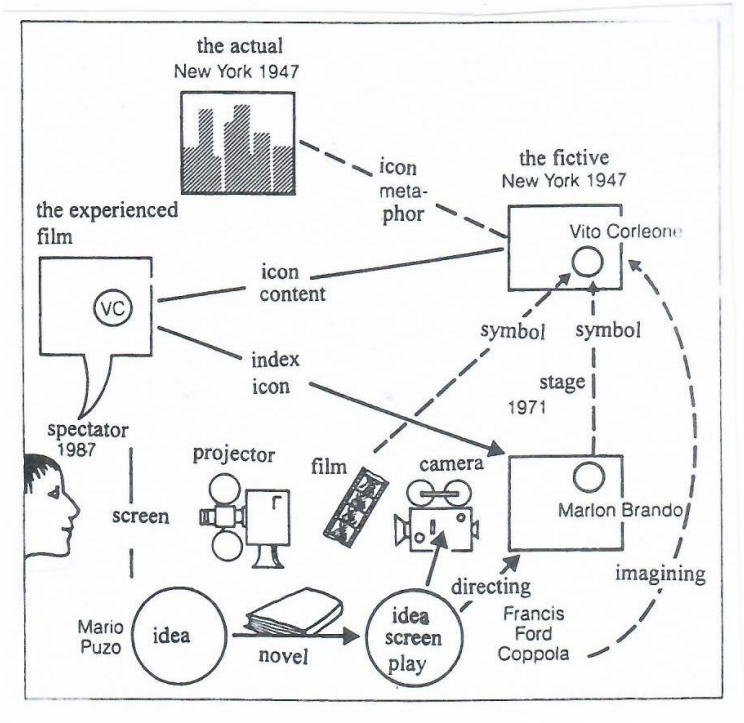

Fig. 3. *The Godfather* and reality.

How is it possible that, in watching the film *The Godfather*, I see Vito Corleone instead of Marlon Brando? It is not necessarily a matter of self-deception: I see Corleone although I know that Brando acts him. Jarvie (1987) emphasizes, following Hugo Munsterberg, that the picture of a film does not deceive the spectator: it is not a matter of a false belief, but of "voluntary illusion".[40]

[40] Conscious illusions that do not lead to false beliefs (for example, a stick in water seems to bend) play an important role in Hintikka's (1969) logic of perception (cf. Chapter 6 above). Because of the non-deceptive illusive nature of film, attempts to mix two levels of reality have been, to some extent, artificial and unsuccessful (e.g. Buster Keaton and Woody Allen have shown a spectator stepping among the events of the screen).

Hence, we may say that, on the basis of a speechless convention between the director and the spectator, Marlon Brando is a sign referring to a fictive Mafia boss. I can understand the film only if I accept this convention from the beginning. Thus, we can say that Marlon Brando — as well as Brando's picture on the filmstrip — works as a conventional sign, as a *symbol* in Peirce's sense, which refers to the fictive entity Vito Corleone. The language of dramatized films thus includes, in addition to indices and icons, the level of "codes" or symbols. Hence, all types of signs of Peirce's tripartition are relevant in film semiotics.[41]

In the language of film, narrative codes that the director may use as his or her personal means of expression have a special significance. Business negotiations in an obscure room and the daughter's wedding in the bright yard create a contrast, which expresses the dark background and the well-lighted facade of the godfather's double role. An analysis of a film is often precisely interpretation of these kinds of means. Even though they are usually based on different metaphors and indirect references — on "connotation" of signs — they may, if established, rise to the status of a generally adopted, in Peirce's sense conventional sign, i.e. a symbol. In this spirit, Peters (1981) emphasizes that each film creates a unique symbol system or code of its own through its multi-layered language of expression.

The above-mentioned symbols are included in the story: they refer to the fictive world of *The Godfather* instead of the actual world. What about the "realistic" nature of the film *The Godfather* in the sense of the relation between the imagined New York and the actual New York? It is clear that this kind of relation exists, since Mafia is a real phenomenon. Although Vito Corleone is a fictive character, he resembles in many respects the real godfathers of New York — and thus functions as their icon. More generally, we may say that Vito Corleone "symbolizes", in a metaphorical form, organized criminality as a whole. Furthermore, it is possible to consider the moral message of *The Godfather* by interpreting the Mafia boss leading

[41] My argument for this thesis is different from that of Wollen (1977) and Peters (1981).

and taking care of his "family" as a metaphor of the President of the United States.

Through fictive films, an artist has an enormous liberty of creating reality, of illuminating his or her own imagination by means of narration. The most absorbing questions in interpreting individual films are related to what the director meant (if he or she even knows it him- or herself) by the metaphors he or she invented. Those metaphors that connect the story with some actual or potential features of our own world are often the most exiting ones.[42] These kinds of metaphors relate fiction to reality and thus produce the core of film art, in Tarkovsky's sense:

> To make a man face his limitless environment, meet an innumerable number of other people, relate him to the world.

Epilogue: Art and Truth

As Jarvie (1987) points out, Plato anticipated the later cinema culture in the allegory of the cave in the 7th book of his work *Republic*: a group of people chained to the back wall of a cave watch shadows projected on the wall from various objects, which are passed in front of a fire burning behind them. According to Plato, the prisoners of the cave are misled, whereas the true objects are known only to a philosopher who breaks out and sees the reality in sunlight. In the same way, according to Plato's doctrine of ideas, the changing sensible world known by perception is only appearance, while the true reality is constituted by unchanging forms or ideas, known only by the light of reason.

In the 10th book of *Republic*, Plato distinguishes (i) the real couch, as an idea created by God, ii) the couch produced by a craftsman, as a dim shadow of the first couch, and (iii) the couch painted by a painter, as an imitation (Gr. *mimesis*) of the

---

[42] In this sense, Scruton (1983) distinguishes arbitrary "fantasy" from "imagination" which is in some way related to reality. In a film, imagination can of course be used in many ways. One of these is meta-fiction (e.g. François Truffaut's *La nuit americaine* tells about making a film in a group whose leader is acted by Truffaut himself; Carlos Saura's *Carmen* tells about the makers of a dance version of Carmen who realize the destinies of the story in their own relations).

second couch. The painted couch is "an appearance as it appears" rather than "imitation of reality as it is" (598b). Plato concludes that the imitator is "the producer of the product three removes from nature", so that "mimetic art is far removed from truth" (597e). Imitative poetry is for him as confusing as scene painting, witchcraft, and jugglery (602d), since it has the "power to corrupt, with rare exceptions, even the better sort" (605c). Therefore, poets and artists are not admitted to the ideal state (695a).[43]

Plato's condemnation of art is based on a severe criterion of realism: painting and poetry ought to be real and truthful – but, as they are not capable of this, art should be rejected as worthless and corruptive.[44]

One way of answering Plato's challenge is to abandon realism. Expressive theories maintain that the aim of art is not to represent reality: good art is the artist's creative and honest self-expression. It produces beautiful artefacts (cf. Chapter 3 above), which give aesthetic experiences to the author as well as the spectators. In particular, emotivism views art as the expression and communication of emotions from the artist to the audience.

Effect theories take the value of art to be in the ability of its works to influence the receivers: comedy gives joy, horror films fear, love stories tears, political art activity in changing the world.[45]

Another reply to Plato is the attempt to save realism. This was tried already by Aristotle, even though he accepted the mimetic theory of art. But, in *Poetica*, he allowed that poetry may "imitate" events even when they are only possible or something that ought to be. This kind of narrative about the possible is "more serious" that history about particular events, since it is able to convey "general truths". Art has a positive effect, when the imitation of compassion and fear in fictive

---

[43] Similar arguments in favor of censorship were launched by conservative religious groups in their campaign against Martin Scorsese's film *The Last Temptation of Christ* (1988).

[44] Postmodernism in art also suspects the possibility of artistic representation of reality, but draws a conclusion opposite to Plato: art is great fun as a play, and should not be taken too seriously.

[45] For expression and effect theories of art, see Dickie (1971).

tragedy causes the "purification" of the spectator and his emotions (Gr. *katharsis*). Following Aristotle, many theorists emphasize that documentary and fictive art should give *information* about reality - this could be the objective World 1 and the social World 3, but also the subjective World 2. Art provides a language which allows the artists (not only to express their momentary feelings but) to describe and represent emotions – and to tell about their conceptions of what joy, love, grief, jealousy, and agony really are. Thereby art involves (besides its emotional and affective meaning) a cognitive dimension: a work of art is a witness of the experiences and conceptions of its creator, and it may help others to see and structure the world in a novel way.

Goodman (1976) expresses this cognitive dimension with the words "right" and "understanding", instead of "truth" and "knowledge". One may say that the neo-pragmatist school of "new philosophy of science" (Kuhn, Goodman, Putnam, Rorty, and others) have tried blur the distinction between science and art, but in my view this is problematic in many ways. Even though scientific theories are products of creative imagination, and their evaluation may involve considerations of beauty,[46] aesthetic criteria are secondary to epistemic ones in scientific theorizing. Indeed, the relations of science and art to truth are different in characteristic ways.[47]

Let a be the *author* of a *fictional text* F (such as a short story, a novel, a play, a painting, a photograph, a film), conceived as a conjunctive sequence of sentences in the language of art. For simplicity, assume that F does not include any non-fictional sentences. Then in publishing the text F, its author a certainly is not *asserting* that F is true in the actual world. Searle (1979) suggests that a is (nondeceptively) *pretending to assert* that F. However, with explicitly fictional stories it does not seem to me at all plausible to say that the author "pretends" to assert anything or "pretends to refer" to something. Instead, a is making a sincere attempt to describe something that she has first imagined in her mind. She is trying to transform her private imaginary characters to public ones. So a is really engaging in the illocutionary act of *recommending* her readers to share

[46] See Kuipers (2019), Ch. 9.
[47] Cf. Niiniluoto (1985), 220-221.

her imagination: the customary opening of a tale "Once upon a time there was …" has the same force as "Let us imagine that F!", which as an imperative is neither true or false, and does not commit a to defending the truth of F. It is nevertheless possible that a intends to convey some interesting truths by her story: even if the actual world W is not a model of F, the text F may have logical consequences which are actually true.

Let Mod(F) be the class of possible worlds which are compatible with F. This is the class of models where F is true, and the actual world W is not included in it. The class of worlds Mod(F) is "projected" in Wolterstorff's (1980) sense. Mod (F) is the narrower, the more informative the text F is. In contrast to "external" truth and falsity in W, we may follow David Lewis (1978) by saying that everything that is true in Mod(F) is "internally true" in fiction F; everything that is false in Mod(F) is "internally false" in fiction F, and the rest of the sentences are indeterminate in F. For example, "Anna Karenina committed suicide" is internally true in Tolstoy's novel, even though Anna Karenina is a fictional entity.

As a description of a possible world, every fictional text is necessarily incomplete. For example, "Anna Karenina had a birthmark on her left shoulder" is left indeterminate by Tolstoy. One might regard fictional objects as "thin" characters, like ghosts or zombies, who have only a few properties and relations recorded in writing by the author. But it is much more plausible to allow the author to have a set of contextual *presuppositions* P_a which enrich the characters and events with the kind of features and activities that they would have in the actual world (the hero of the story has blue eyes, sleeps every night, eats breakfast etc.), even when they are not mentioned in the text F. Then the projected class of worlds Mod(F&P_a) is much narrower than Mod(F). The presuppositions are usually relative to the cultural context,[48] the genre and the artist's style.[49] The reader interprets the text with her own presuppositions P_b, and thereby has an active role in "constituting" for herself the imaginary world of F. The communication between

[48] In the old Viennese novels, the reader understood that the main couple has started an intimate relation when they started to call each other by first names or use the *Du* –pronoun.

[49] Surrealist art does not assume even the validity of the laws of nature.

a and b is successful to the extent that their projections Mod(F&P_a) and Mod(F&P_b) overlap each other. By guessing what the author's presuppositions P_a are, the reader b tries to understand the indented meaning of F. But, as the presuppositions P_a may be partly unconscious to a, an interpreter b may sometimes claim that she understands the text F better than its author. It is also important that, by the public nature of the text F, the author and readers are able to think about the same fictional objects.

The traditional distinction between science and art can be formulated by noting that the fictive text F is not even intended to be externally true in the actual word W, while a scientific theory T should be true in W. But this is a simplification, since many theories in science include idealizations and approximations, so that they are at best *truthlike* or *approximately true*. This means that the actual world W is at some positive but small distance from the class Mod(T), i.e. T is true in some world W´ which is similar to W.

On the one hand, as Aristotle understood, fictive texts may imply by concrete examples something generally true about the actual world. In the case of films, Coppola's *The Godfather* tells about organized crime, Ingmar Bergman's *Smultronstället* on aging and time consciousness, westerns about moral decisions, Buster Keaton's films about the absurdity of the technological world, Fritz Lang's *Metropolis* about the future dangers of robotics, and Charlie Chaplin's pilgrim about the eternal human striving for a better world.

However, the requirement of truth is not constitutive of art in the same way as science, where temporary idealizations are eventually removed or "concretized". A similar process is not applied in the works of art. A painting of a fictive character may have more aesthetic value than a poor portrait of a real person. Beauty and truth do not exclude each other, even though they are conceptually distinct values.

Note. This chapter is based on the article "Elokuva ja todellisuus", *Synteesi* 7 (1-2) (1987), 90-107. See also the extended version "Elokuva, kuvittelu ja todellisuus", in *Maailma, minä ja kulttuuri*, Helsinki: Otava, 1990, 247-282. Its shorter version has been published in English as "Film and Reality", in Eero

Tarasti (ed.), *Snow, Forest, Silence: The Finnish Tradition of Semiotics.* Bloomington: Indiana University Press, 1999, 248-261.

Literature

Andrew, Dudley (1976). *The Major Film Theories: An Introduction.* Oxford: Oxford University Press.

Andrew, Dudley (1984). *Concepts in Film Theory*. Oxford: Oxford University Press.

Aristotle (1924). *De Poetica*, transl. by I. Bywater. In *The Works of Aristotle*, vol. XI, ed. by W.D. Ross. Oxford: Clarendon Press.

Armes, Roy (1974). *Film and Reality: An Historical Survey.* Harmondsworth: Penguin Books.

Arnheim, Rudolf (1958). *Film as Art*. London: Fuller and Fuller.

Barthes, Roland (1980). *La chambre Claire*. Paris: Gallimard.

Basin, Y. (1979). *The Semantic Philosophy of Art.* Moscow. Progress.

Bazin, André (1967-71). *What Is Cinema?*, vol. 1 -2. Berkeley: University of California Press.

Cavell, Stanley (1979). *The World Viewed: Reflections on the Ontology of Film.* Enlarged Edition. Cambridge, MA: Harvard University Press.

Dickie, George (1971). *Aesthetics: An Introduction*. Indianapolis: Pegasus.

Eisenstein, Sergei (1965). *Film Form: Essays in Film Theory*. Cleveland: Meridian Books.

Freeland, Cynthia & Wartenberg, Thomas (1995). *Philosophy and Film*. New York & London: Routledge.

Gombrich, Ernst H., Hochberg, Julian & Black, Max (1972). *Art, Perception, and Reality*. Baltimore and London: The Johns Hopkins University Press.

Goodman, Nelson (1976). *The Languages of Art*. Indianapolis: Hackett.

Hintikka, Jaakko (1969). *Models for Modalities*. Dordecht: D. Reidel.

Hintikka, Jaakko (1975). "Concept as Vision: On the Problem of Representation in Modern Art and in Modern Philosophy". In *The Intentions of Intentionality and Other New Models of Modalities*. Dordrecht: D. Reidel, 223-251.

Jarvie, Ian (1987). *Philosophy of the Film: Epistemology, Ontology, Aesthetics*. New York and London: Routledge & Kegan Paul.

Knight, Arthur (1957). *The Liveliest Art: A Panoramic History of the Movies*. New York: Macmillan.

Kuipers, Theo (2019). *Nomic Truth Approximation Revisited*. Cham: Springer.

Langer, Susan (1953). *Feeling and Form*. London: Routledge and Kegan Paul.

Lewis, David (1968). "Truth in Fiction", *American Philosophical Quarterly* 15, 37-46.

Lotman, Juri (1976). *Semiotics of Cinema*. Ann Arbor: University of Michigan.

Loux, Michael (ed.) (2007). *Metaphysics: Contemporary Readings*. London: Routledge.

Martin, Marcel (1968). *Le langage cinématographique*. Paris: Éditions du Cerf.

Mast, Gerald & Cohen, Marshall (eds.) (1979). *Film Theory and Criticism: Introductory Readings*, 2nd ed. Oxford: Oxford University Press.

Mitscherling, Jeff (1983). "Roman Ingarden's 'The Literary Work of Art': Exposition and Analyses", *Philosophy and Phenomenological Research* 45, 351 — 381.

Niiniluoto, Ilkka (1985). "Remarks on the Logic of Imagination". In Ghita Holmström & Andrew Jones (eds.), *Action, Logic and Social Theory*, Acta Philosophica Fennica 38, Helsinki: The Philosophical Society of Finland, 183 — 202.

Niiniluoto, Ilkka (1986). "Imagination and Fiction", *Journal of Semantics* 4, 209-222.

Niiniluoto, Ilkka (1988). "Analogy and Similarity in Inductive Logic". In David Helman (ed.), *Analogical Reasoning*. Dordrecht: D. Reidel, 271-298.

Peters, Jan Marie (ed.) (1981). *Pictorial Signs and the Language of Film*. Amsterdam: Rodopi.

Plato (1956). *The Republic*, transl. by W. H. D. Rouse. In *Great Dialogues of Plato*. New York: The New American Library, New York, 118 — 422.

Poincaré, Henri (1952). *Science and Hypothesis*. New York: Dover. (Original in French, 1902.)

Poe, Edgar Allan (1840). "Daguerroetype". *Alexander's Weekly Magazine*.

Putnam, Hilary (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.

Quine, W. V. O. (1969). *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Scruton, Roger (1983). *The Aesthetic Understanding*. London: Methuen.

Searle, John (1979). *Expression and Meaning*. Cambridge: Cambridge University Press.

Sontag, Susan (1977). *On Photography*. New York: Farrar, Straus and Giroux.

Stack, Oswald (1969). *Pasolini on Pasolini*. London: Thames and Hudson Ltd.

Tarkovsky, Andrei (1987). *Sculpting in Time: Reflections on the Cinema.* London: Bodley Head.

Viano, Maurizio (1993). *A Certain Realism: Making Use of Pasolini's Film Theory and Practices*. Berkeley: University of California Press.

Williams, Christopher (ed.) (1980). *Realism and the Cinema: A Reader.* London: Routledge and Kegan Paul.

Wollen, Peter (1970). *Signs and Meaning in the Cinema.* London: Thames & Hudson.

Wolterstorff, Nicholas (1980). *Works and Worlds of Art*. Oxford: Oxford University Press.

Chapter 6: Virtual Worlds

The aim of this chapter is to raise and discuss some philosophical questions about Virtual Reality (VR). The most fundamental problem concerns the ontological nature of VR: is it real or fictional? Is VR comparable to illusions, hallucinations, dreams, or worlds of fiction? Are traditional philosophical categories at all sufficient to give us understanding of the phenomenon of VR? In approaching these questions, I shall employ possible world semantics and logical theories of perception and imagination as my philosophical tools. My main conclusion is that VR is comparable to a 3-D picture which can be seen from the inside.

What is Virtual Reality?

In his book *Virtual Realism* (1998), Michael Heim states that virtual reality is a "technology" or "an emerging field of applied science" (p. 4). This is, indeed, one way of looking at VR: computer programs, data gloves and helmets are used to produce artificial sensory inputs; these inputs resemble the participant's normal interface with the physical environment and thus she feels herself immersed in a new "reality". Hence, VR is a method of constructing and designing new kinds of artefacts.

As a technological activity, VR can be assessed by various criteria which include economy, efficiency, aesthetics, ergonomics, ecology, ethics, and social effects (see Chapter 20 above). For example, from the aesthetic viewpoint VR is a new form of media art: by using methods of interactive design it helps to produce works and experiences with interesting aesthetic qualities. From the ergonomic perspective, intense occupation with extraordinary sensory stimulations may lead to a state of nausea where images of virtual worlds and the actual

world are confused; this is called the Alternate World Syndrome (AWS) and Alternate World Disorder (AWD) by Heim (1998).

In *Virtual Reality* (1991), Howard Rheingold points out that VR can be employed for the purposes of entertainment, but it may also function as a way of escape and addiction. These ethical and social concerns are also discussed by Heim who suggests that "virtual realism" could mediate between "naive realists" (who blame computers for all social evils) and "network idealists" (who promote all new forms of computerized technology).

Instead of aesthetics, ergonomics, and ethics, my main philosophical concern in this chapter is *ontological*. Since Jaron Lanier coined the term "virtual reality" in 1986 and William Gibson spoke of "cyberspace" in his novel *Neuromancer* in 1984, this new area of human digital culture has been characterized in the United States and Japan by such terms as "virtual environment", "synthetic environment", "virtual worlds", "tele-presence", and "tele-existence". Notions like "reality", "world", "environment", "space", "presence", and "existence" are ontological categories in the sense that they refer to the most general structures of what is real or exists. On the other hand, they are here qualified by phrases like "virtual" and "tele" which imply some kind of deviation or distance from reality.

Today there is a tendency of calling "virtual" almost anything connected with computers: a "virtual library" allows multimedia works to be read in the Web, and a "virtual university" offers courses in the Internet in an electronic form. The original Latin term *virtus* meant human powers and potentiality, and later it referred to the "virtues" of human character. The *Concise Oxford Dictionary* defines "virtual" as "that is such for practical purposes though not in name or according to strict definition", and Heim defines it as "not actually, but as if" (*op.cit.*, 220). In this as-if sense, terms like "virtual velocity" are comparable to expressions like "semiofficial", "pseudoscience", and "artificial leg" which imply that something is only half, falsely, seemingly, non-naturally, or non-genuinely so-and-so. In the same way as we may ask whether artificial

intelligence merely pretends or "really" is intelligence,[1], we may raise the question whether virtual reality is "really" real or not.

Even though Heim warns that VR is "not synonymous with illusion or mirage or hallucination", and "not a state of consciousness or a simulated drug trip" (p. 4), the idea of "as-if reality" clearly has a connection to the old philosophical distinction between appearance and reality. Therefore, our analysis of the ontological status of VR has to proceed by comparing it to some important types of phenomena discussed in traditional epistemology.

Reality vs. Fiction

According to Karl Popper's (1979) useful classification, the domain of reality can be divided into three parts (cf. Chapters 1 and 5 above). *World 1* consists of physical objects and processes, *World 2* contains mental states and events within a human mind, and *World 3* includes human-made artefacts and socially produced institutions. In the traditional terminology, the Popperian three-fold ontology corresponds to the division between nature, consciousness, and culture & society.

Physical entities exist in space and time, and they are in causal interaction with each other. *Physicalists* take these features to be the general criteria of reality or actual existence.[2] Therefore, they either eliminate World 2 entities or reduce them to physical states of human brains or bodies. Similarly, most physicalists are *nominalists* who either reject all World 3 entities as philosophically illegitimate abstractions or attempt to reduce them via subjective World 2 entities to World 1. Conversely, the *subjective idealists* (e.g. solipsists and phenomenalists) suggest that the whole of reality should be reduced to World 2, while the *objective idealists* take as the primary reality some abstract entities like Plato's forms or Hegel's objective spirit. In contemporary philosophy, idealism has gained some support in new linguistic and social forms: the *social constructivists* claim that the world is a "social construction" arising from human practices and discourses.

---

[1] Cf. the distinction between weak and strong AI in Searle (1984).

[2] See Devitt (1993).

In my view, the most plausible interpretation of Popper's ontology is *emergent materialism* which differs both from physicalism and idealism.[3] Worlds 2 and 3 have historically developed from the primary World 1 through biological and cultural evolution, and their existence continues to be dependent on the support provided by the material reality. However, as emergent levels of reality, Worlds 2 and 3 have gained relative independence from World 1: they are able to be in a complex mutual interaction with World 1, and they have their own characteristic features and lawlike regularities that cannot be derived from true theories about the physical world.

Emergent materialism accepts ontological *realism* in two different senses. In the first place, against nominalists, it admits the reality of some abstract entities (like concepts, propositions, numbers, symphonies) which are different from their physical documentations (like written and uttered words and sentences) and mental representations (like ideas and thoughts). However, against Platonist versions of realism, such World 3 entities are regarded as human-made social constructions. Secondly, a realist may accept that human beings can causally interact with physical nature, and that World 1 can be structured or "carved up" in different ways by means of various conceptual frameworks; in this sense one may speak about "worldmaking" with Nelson Goodman (1978). However, unlike Goodman and the recent versions of social constructivism, the realist asserts that World 1 exists independently of human minds, languages, and societies.

Charles Peirce defined "the real" as "that whose characters are independent of what anybody may think them to be".[4] This definition allows for the existence of *real possibilities*: for example, fragility is a real dispositional property of a glass window, since it would be manifested in a regular way under certain conditions – quite independently of what we may think about the matter. In this sense, powers, dispositions, tendencies, and propensities may be real even when they are not actualized.

As Peirce himself suggested, his definition gives us a criterion for distinguishing reality and *fiction*. For example, if I am thinking about a golden mountain, then my mental state of

---

[3] See Niiniluoto (1999, 2006).

[4] *Collected Papers* 5.311, 5.405.

thinking-a-golden-mountain is a real fact about World 2, but the golden mountain is only a figment of my imagination. When Leo Tolstoy published his novel *Anna Karenina*, a public work in World 3 was created, but Anna Karenina as a character in the novel remains a fictional entity. Her properties include only those implied by Tolstoy's novel (cf. Chapter 5 above). In contrast, natural numbers as mathematical constructions are real in Peirce's sense: any sufficiently large number which has never been investigated in arithmetic still has properties like being prime or not – even when we don't yet know that.

Around the year 1900 a famous controversy took place between Alexius Meinong and Bertrand Russell. Meinong suggested that all names and definite descriptions, including empty ones (like "the present king of France") and fictional ones (like "Donald Duck"), should be understood to have referents, where the entities serving in this role are "real" but not necessarily actual or existing. Russell showed how discourse with such empty descriptions can be semantically interpreted as typically consisting of false statements – without assuming strange Meinongian entities. In terms of the *possible world semantics,* there may be non-actual and unreal possible worlds where presently France has a king or Anna Karenina is living. Fictional entities are thus denizens of possible worlds. What is today called "Meinong's jungle" by Routley (1980) is a composite of all objects and states of affairs which are logically or conceptually possible.

Perception and Imagination

The classical problem of *epistemology* concerns our possibility of obtaining knowledge about external reality. Given Plato's definition of knowledge as justified true belief and Descartes' distinction between mind and matter, we may ask on what conditions our beliefs are reliable and correct representations of material facts. Stated in other terms, this is a problem about relations between mental states in World 2 and facts in World 1. Another problem of epistemology concerns our knowledge of the human mind – this involves relations between states in World 2.

A paradigm case of knowledge is *perception*. Suppose that I see a tree. In the case of *veridical* perception, this means that

there is really a tree in front of me, that it causally influences my sensory apparatus, and the received visual data give rise to a perceptual judgment of the form "This is a tree". The tree exists in World 1, the mental state of seeing the tree in World 2. This causal account of perception can be combined with Ludwig Wittgenstein's famous thesis that all seeing is *seeing as*. In order to see the thing in front of me as a tree, and thus to form the judgement "This is a tree", it is presupposed that I have already learnt the concept "tree".

In Jaakko Hintikka's (1975) logic of perception, a statement of the form "I see b as an F" can be interpreted as claiming that in all possible worlds compatible with what I see there is an object of type F in front of me, and this F-perception is caused by the object b existing in the actual world. Such a F-perception is veridical just in case the object b is really of type F.

A visual *illusion* obtains when I see a real object but make a mistake in its character: I see b, which is not an F, as an F. To see a bush as a bear is a typical case of illusion. Many philosophers have defined illusions as false beliefs caused by sensations, but as Hintikka points out, there are conscious visual illusions where we are not deceived by what appears to us. I know that an oar does not bend in water, but I cannot help seeing it as bent when I put it in water.[5]

When I see a star, the causal chain from the remote object to my perception may take millions of years. But even more complicated cases are obtained by allowing that the causal link is mediated by artificial technologies. Perhaps no one objects, if I claim to see my wife through spectacles, a window, or even a mirror. But could I see her through a picture? If I am looking at a photo of John Wayne, or watching John Ford's western *The Searchers*, do I see John Wayne? And if I am arranging a teleseminar with my colleagues in London, do I see them from my video-conference room in Helsinki? As long as the causal chain is more or less mechanical and the signals causally transmit sufficiently correct information about the source, I am inclined to answer these questions positively – in the same way as we are already accustomed to saying that we hear other persons (not only their recorded voice) on the telephone or radio. But if the picture is a painting of John Wayne made by an artist,

[5] Cf. Niiniluoto (1982).

then it seems more natural to construe the situation to be such that the object of my perception is the painting – and add that under certain circumstances a picture of X serves as an iconic sign of X in Peirce's sense (see Chapter 5 above). Icons are signs which are similar to their objects in some respects (cf. Chapter 4 above). Thus, seeing a picture of X may indirectly give information about X as well. In particular, seeing a picture of X and directly seeing X may involve perceptual experiences closely resembling each other.[6]

Hallucination is usually treated as a limiting case of perception, since the person is convinced that she sees or hears something. If I am experiencing a hallucination, there is no object in front of me, or the "normal" causal link between reality and what I seem to perceive is missing. If it seems to me that I am seeing a pink elephant in my room, this deceptive appearance may be due to extreme nervous exhaustion or drug excitement of my brain functions.

The logic of *imagination* can be developed along the same lines as the logic of perception by employing the possible world semantics.[7] The statement "I imagine that p" means that in all possible worlds compatible with what I imagine it is the case that p. Acts of imagination may be voluntary (fantasy, daydreaming) or involuntary (dreaming). Again we have statements of the form "I imagine b as an F". Here b may be a real object, but then no direct causal link from b to the contents of my imagination exists: for example, I may imagine of my wife (who is in Helsinki) that she is coming to meet me during my trip to Rome, or in my dream I may discuss with my late father. But b may also be an imagined object. I may use private phantasy to create in my mind a person and then imagine that she is dancing with me. Or I may imagine that I am walking with some publicly known fictional character like Mickey Mouse.

Appearance and Reality

The traditional distinction between *appearance* and *reality* can be directly applied to cases involving perceptual illusions (the real bush appears to me as a bear) and hallucinations (the pink

[6] Cf. Gombrich, Hochberg, and Black (1972).
[7] See Niiniluoto (1986, 2020).

elephant that I seem to see is not real). From my perceptual experience alone I cannot judge whether it is veridical or not. The *sceptical* question immediately arises: how can we ever be certain or justified in thinking that our perceptions correspond to reality?

Some philosophers have found idealism attractive, since it abolishes the distinction between appearance and reality: if no reality is hidden behind our observations, then scepticism becomes obsolete. The ancient sceptics had a more pragmatic attitude: follow the appearances in your everyday life and withhold all judgments about their reality.[8] Immanuel Kant's critical idealism accepts things in themselves behind phenomena, but asserts that nothing beyond their existence can be known by human beings. Edmund Husserl's phenomenology adopts the research programme of studying the-world-as-conceived-by-us and putting the external world-as-it-is-in-itself into brackets.

However, a realist who accepts the three worlds ontology outlined above cannot avoid facing the problem of scepticism.[9] Moreover, she has to be ready to consider its most radical version: how can I know that I am even perceiving something rather than only seeing a dream?

World literature contains many touching descriptions of situations where a person feels uncertain whether she is dreaming or not. "We are such stuff as dreams are made on, and our little life is rounded with a sleep", Shakespeare exclaimed in *The Tempest*. In *Hamlet*, he described an alienated outsider, beset with a weakened sense of reality and a melancholic feeling of a shady dream-like existence. This ambiguous mood of mind was expressed by romantic poets of the 19th century in well-known verses – Samuel Taylor Coleridge in *Reality's Dark Dream* (1803):

> I know 'tis but a dream, yet feel more anguish
>
> Than if 't were truth. It has often been so:
>
> Must I die under it? Is no one near?

[8] See Niiniluoto (2000).
[9] See Niiniluoto (1999).

> Will no one hear these stifled groans and wake me?

and Edgar Allan Poe (1845):

> All that we see or seem
>
> is but a dream within a dream.

The same theme, treated as an epistemological rather than existential problem, has been discussed by philosophers ever since Plato's dialogue *Theaetetus*. To refute the attempted definition of knowledge as perception, Socrates raises a question "you must often have heard persons ask":

> How can you determine whether at this moment we are sleeping, and all our thoughts are a dream; or whether we are awake, and talking to one another in the waking state?

Theaetetus replies:

> Indeed, Socrates, I do not know how it could be determined, for in both cases the facts precisely correspond; and there is no difficulty in supposing that during all this discussion we have been talking to one another in a dream.

This thesis – viz. merely illusory, imagined, or dreamed experiences do not contain any observable feature that would distinguish them from "real" presentations – was called the *Theaetetus theorem* by Eino Kaila in 1958.[10] Perhaps the most famous formulation of this "theorem" was given by René Descartes in his *Meditations on the First Philosophy* (1641). In exercising his method of universal doubt, Descartes ponders in his chamber:

> However, I must here consider that I am a man, and consequently that I am in the habit of sleeping and of representing to myself in my dreams those same things, or sometimes even less likely things, which insane people do when they are awake. How many times have I dreamt at night that I was in this place, dressed, by the fire, although I was quite naked in my bed? It certainly seems to me at the moment that I am not looking at this paper with my eyes closed; that this head that I shake is not asleep; that I hold out this hand intentionally and deliberately, and that I am aware

[10] See Kaila (1979), 261.

> of it. What happens in sleep does not seem as clear and distinct as all this. But in thinking about it carefully, I recall having often been deceived in sleep by similar illusions, and, reflecting on this circumstance more closely, I see so clearly that there are no conclusive signs by means of which one can distinguish clearly between being awake and being asleep, that I am quite astonished by it; and my astonishment is such that it is almost capable of persuading me that I am asleep now.[11]

The Theaetetus theorem does not deny that sometimes in dreaming I may have a strongly felt conviction that "this is only a dream". Plato and Descartes were looking for a general criterion which would exclude all doubt about my state. But if a property F is proposed as a criterion of waking, in particular cases it is always possible to claim that I only dream that my experience has this property F.

Kaila concluded that the Theaetetus theorem is valid. However, he argued that Descartes failed to distinguish *logical doubt* from *empirical uncertainty*: Even if it is always logically possible to doubt the reality of our impressions, we may still in fact be in some weaker sense empirically certain about their reality.

Many philosophers have accepted the Theaetetus theorem for momentary experiences, but still they have suggested that the interrelations of sequences of experiences provide a criterion of reality. In the Sixth Meditation, Descartes concluded that "our memory can never connect our dreams with one another and with the general course of our lives, as it is in the habit of connecting the things which happen to us when we are awake" (*op.cit.*, 168). This consistency requirement is hardly so conclusive as Descartes implied, since sometimes a single dream at least seems to cover a whole life.

G. W. Leibniz admitted in 1704 that "it is not impossible, metaphysically speaking, for a dream to be as coherent and prolonged as a man's life", but he regarded this as highly improbable.

> Consequently I believe that where objects of the senses are concerned the true criterion is the linking together of phenomena, i.e.

---

[11] Descartes (1968), 96-97.

> the connectedness of what happens at different times and places and in the experience of different men.[12]

In his works in the 1930s, Eino Kaila accepted and elaborated on Leibniz's idea that the defining character of reality is *invariance* – regularity, lawlikeness, and the possibility of prognosis.[13] He proposed that different types of things can be placed on a scale of levels according to their *degree of reality* defined by their degree of invariance: perceptual experiences, everyday physical objects, and objects postulated by scientific theories. Dream experiences clearly have a low degree of invariance and should be placed on the lowest levels of Kaila's hierarchy.

One way of supporting Kaila's argument can be based on the theory of evolution: the evolution of life and the human species would not have been possible in an irregular dream world. But at the same time, we should acknowledge the fact that our actual world is not completely lawlike in all of its respects, but includes chance and irregularity as well.

Verisimilitude and Virtual Reality

The classical problem of realism has received new impetus from the techniques of *representation* that have been developed in the history of art – from poetic language to pictures, cinema, television, CD-roms and virtual reality (cf. Chapter 4). The relation of an artistic representation to reality can be discussed in the same way as the relation between perceptions and reality. But just as the products of imagination, works of art may also be intended as representations of fictional possible worlds.

Rheingold (1991) starts the history of virtual reality from the wall paintings in the caves of Lascaux. Paintings in medieval churches were understood as "windows onto other worlds". At the same time, there was the Roman tradition of poetics and rhetoric which demanded that even fictional narrative stories should have *verisimilitude*: their characters should not have queer or supernatural powers, but rather be plausible relative to the reader's expectations.[14]

---

[12] Leibniz (1982), 374.

[13] See Kaila (1979), 102.

[14] See Mehtonen (1996). Cf. Chapter 7 below.

Besides the school of realism which seeks accurate representations of external or inner reality, romantic poets and artists have always been attracted by irregularity, unpredictability, and space-time-discontinuity (i.e. lack of invariance in Kaila's sense). These analogies of dreams (phantasms, hallucinations, myths, absurdities) were consciously employed in the theoretical writings and artistic experiments of the dadaist and surrealist schools.

Susanne Langer presented in *Feeling and Form* (1953) her famous thesis that film as a poetic art uses "the dream mode". According to Langer, visual arts like painting create an artificial or "virtual space" that can be seen but not touched. Cinema is like a dream: it creates an illusion of reality, a virtual present where the moving camera takes the place of the dreamer.

The dramatic character of dreams – without a specific reference to films – was discussed already by Jean-Paul Sartre in *L'Imaginaire* (1940). Sartre argued against Descartes and the Theaetetus theorem that unlike perceptions, dreams are associated with a special type of "belief" or "fascination without existential assumption": my dreams are adventures like stories in novels, they close on my consciousness in an imaginary world without presenting themselves as apprehensions of reality:

> The dream is not fiction taken for reality, it is the Odyssey of a consciousness dedicated by itself, and in spite of itself, to build only an unreal world.[15]

Sartre's argument is important, since it explains the haunting and often frustrating character of dreams: even if my dreams are authored by *my* subconsciousness and there is often *me* playing a central role in these stories, dream-events occur to me without my full control and frequently my dream-plans fail or change in disturbing ways.

In this sense, I have less control over the contents of my dreams than over my daydreams or waken imaginations. But in compensation dreams have a much stronger apparent verisimilitude or illusion of reality.

---

[15] Sartre (1972), 206.

However, films need not be compared to dreams in order for us to argue that they are systematically based upon visual illusions.[16] In seeing a film, a documentary or a fiction, I am in fact looking at a screen onto which pictures are projected 24 times per second, and the impression of continuous movement is created in my mind. Moreover, by using the technique of editing the film material, dramatic scenes may be composed of pictures taken in different places at different times, actors replaced by stuntmen, etc. Still, by filling in missing details and by combining different sequences, I see the events as taking place in a "virtual space" in Langer's sense. This virtual space is not actualized unless the film is perceived by someone, but the film as an artefact has the dispositional capacity to produce these audiovisual perceptual experiences in spectators.

These issues have gained new significance in the "postmodern" communication society, where we live in the middle of various kinds of electronic signs, neon lights, radio waves, TV screens, movies and videos – and reality seems to be replaced by a web of representations of reality. These representations (especially when they are transformed into digital form) can easily be manipulated and distorted by "image processing". Jean Baudrillard (1984), a radical commentator of postmodern culture, is claiming that reality itself ceases to exist and is transformed to a *hyperreality* or a *simulacrum*, an apparent copy intended to deceive us. In an exaggerated but amusing way he urges that our cultural products or "hyperreal" signs do not any more reflect a basic existing reality or even mask or pervert it, but rather "mask the absence of reality". Thus, for example, "Disneyland is there to conceal the fact that it is the `real´ country, all of `real´ America, which *is* Disneyland." (To Baudrillard's delight, a new European Disneyland has been opened near Paris.)

Observations of this sort suggest that the basic Cartesian question of dream vs. reality could be replaced by another question: am I at this moment *dreaming or seeing a film*? Here I think the most plausible answer is the one that applies more generally to attitudes towards fictional "texts". As Searle (1979) points out, the author and readers of a fictional text F do not assert that F is true, nor do they non-deceptively pretend

[16] See Andrew (1984), cf. Chapter 5 above.

to assert that F. Rather, as claimed by Niiniluoto (1986), more or less openly they ask us to imagine that F. As Kendall Walton (1990) states, a fictional work is "a prop in games of make-believe". In so far as films create visual illusions, they are conscious illusions which may entertain, thrill and amuse us but (*pace* Baudrillard) in general they do not deceive us.

But perhaps the make-believe character of audiovisual signs is based upon the contingent historical fact that the old methods of representation have not been true-seeming enough? From ancient China to Woody Allen, there are stories of artists who have entered their own paintings or films. Virtual reality seems finally to bring to a completion the technological utopia of creating a perfect illusion of reality. This applies especially to VR in the strong sense defined by Heim (1998), characterized by immersion (i.e. the experience that you-are-there), interactivity (i.e. you are not any more an external observer but also an actor or a participant moving in a synthetic cyberspace) and information intensity (i.e. the ability of the program in the memory of a digital computer to create in us experiences of telepresence).

But as Jaron Lanier remarks, we enter this virtual world awake. The environment is given to us but we can choose how to move in cyberspace. So how do we know whether we just now are living *in the real world or in virtual reality*? Is the Leibniz-Kaila criterion of invariance still applicable? At least in the present stage of technology, the answer still seems to be clear. The objects that "we" can encounter by moving in virtual reality are shadowy figments like the "toons" in the Toontown of *Roger Rabbit*. In this sense, virtual reality does not yet have complete verisimilitude, but still has some characters of dreams and phantasms.

Virtual Reality as a Picture

Suppose that I put on a data helmet and enter a virtual city. As the city that I appear to see is not really there in front of me, my perception is not veridical. On the other hand, my visual experience of a city is not merely a product of my imagination, since it is based upon visual data provided by a computer – and, moreover, these data depend partly on my movements

recorded by the computer. Hence, my visual experience should be classified as a case of a visual illusion.

It is important to add that this illusion is not merely a subjective experience, existing mentally in World 2, but the virtual environment is "out there" for anyone who places himself or herself into the appropriate position. In this sense, the virtual city is a public artefact, belonging to World 3. It can be understood as a complicated *three-dimensional picture* that we are able *to see from inside* by using the technological apparatus. Seeing a virtual city is thus an extension of situations that have been discussed by philosophers and psychologists studying the perception of two-dimensional pictures.[17]

Just like any picture, a virtual city can be constructed in three different ways. First, it may be intended as a simulation of some real city: virtual Helsinki is an icon of Helsinki. In such cases, we may ask how realistic (i.e. accurate and comprehensive) a representation of its intended referent the virtual city is. Secondly, VR may be an expression of a city which so far has existed only in the mind of an architect. Then the virtual city may be understood as a description and elaboration of a World 2 entity. Thirdly, VR may represent some fictional city (e.g. Batman's Gotham City). In this case, the question of realism does not arise, and the virtual environment provides a window to a possible world.

Note. This article has originally appeared in *IO Internet Magazine*, vol. 3: *Virtual Environments*, ed. by Pauline von Bonsdorff, summer 2006.
www.helsinki.fi/iiaa/io/io_3_index.html. It has been published, as "Virtual Worlds, Fiction, and Reality", also in *Discusiones Filosoficas* 12, No 19 (2011), 13-28.

Literature

Andrew, Dudley (1984). *Concepts in Film Theory*. Oxford: Oxford University Press.

Descartes, René (1968). *Discourse on Method and the Meditations*. Harmondsworth: Penguin Books.

Devitt, Michael (1991). *Realism and Truth*. Oxford: Blackwell.

[17] See Gombrich, Hochberg, and Black (1972).

Gombrich, Ernst H., Hochberg, Julian & Black, Max (1972). *Art, Perception, and Reality.* Baltimore*:* The Johns Hopkins University Press.

Goodman, Nelson (1978). *Ways of Worldmaking*. Hassocks: Harvester Press.

Heim, Michael (1998). *Virtual Realism.* Oxford: Oxford University Press, Oxford.

Hintikka, Jaakko (1975). *The Intentions of Intentionality.* Dordrecht: D. Reidel.

Kaila, Eino (1979). *Reality and Experience.* Dordrecht: D. Reidel.

Langer, Susanne (1953). *Feeling and Form*. London: Routledge & Kegan Paul.

Leibniz, G.W. (1982). *New Essays on Human Understanding*. Cambridge: Cambridge University Press.

Mehtonen, Päivi (1996). *Old Concepts and New Poetics: Historia, Argumentum, and the Twelfth- and Early Thirteenth-Century Latin Poetics of Fiction.* Helsinki: Societas Scientiarum Fennica.

Niiniluoto, Ilkka (1982). "Remarks on the Logic of Perception". In Ilkka Niiniluoto & Esa Saarinen (eds.), *Intensional Logic: Theory and Applications*. Helsinki: The Philosophical Society of Finland, 116-129.

Niiniluoto, Ilkka (1986). "Imagination and Fiction". *Journal of Semantics* 4, 209-222.

Niiniluoto, Ilkka (1999). *Critical Scientific Realism*. Oxford: Oxford University Press.

Niiniluoto, Ilkka (2000). "Scepticism, Fallibilism, and Verisimilitude". In Juha Sihvola (ed.), *Ancient Scepticism and the Sceptical Tradition.* Helsinki: The Philosophical Society of Finland, 145-169.

Niiniluoto, Ilkka (2006). "World 3: A Critical Defence". In Ian Jarvie, Karl Milford & David Miller (eds.), *Karl Popper: A Centenary Assessment, vol. II. Metaphysics and Epistemology*. Aldershot: Ashgate, 59-69.

Niiniluoto, Ilkka (2020). "Perception, Memory, and Imagination as Propositional Attitudes". *Logical Investigations* 26: 1, 36-47.

Peirce, Charles S. (1931-35). *Collected Papers*, vols. 1-6. Cambridge, MA: Harvard University Press.

Popper, Karl R. (1979). *Objective Knowledge: An Evolutionary Approach.* Oxford: Oxford University Press.

Rheingold, Howard (1991). *Virtual Reality.* New York: Summit Books.

Routley, Richard (1980). *Exploring Meinong's Jungle and Beyond*. Canberra: Australian National University.

Sartre, Jean-Paul (1972). *The Psychology of Imagination*. London: Methuen.

Searle, John (1979). *Expression and Meaning.* Cambridge: Cambridge University Press.

Searle, John (1984). *Minds, Brains, and Science.* Cambridge, MA: Harvard University Press.

Walton, Kenneth (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.

Science

Chapter 7: The Roots of Verisimilitude

Introduction to Conceptual History

A central task of analytic philosophy is to make conceptual distinctions, i.e. to separate ideas which only appear to be similar. Successful distinctions help to remove conceptual confusions and thereby improve our understanding. But it is equally significant to combine or associate ideas whose similarity has not yet been noted. The importance of this task of combination arises from the fact that unperceived systematic connections create artificial boundaries between specialties: scholars work in isolated teams and groups failing to communicate with each other. Such an isolation leads also to spoiled opportunities in rational reconstruction, and to misleading results in historical reconstruction.

In this chapter, I try to illustrate this dual task of conceptual separation and combination in connection with the notion of verisimilitude. Here the failures of communication are especially dramatic, since the same word has been used even at the same time in several different contexts. As a result, the history of this concept is both long and notorious - and so far there are no systematic accounts of its development that would even attempt to cover its various roots.[1]

At least five different groups have used or studied this term, but they have largely failed to note the existence of each other. (i) Karl Popper introduced "verisimilitude" into the contemporary philosophy of science in 1960 by his qualitative criterion for one theory to be more truthlike than another. After the refutation of Popper's definition in 1974, the quantitative non-probabilistic concept of verisimilitude (and related concepts like "degree of truth" and "approximate truth") has been

[1] Cf. Niiniluoto (2000), however.

explicated by using the notion of similarity between states of affairs. (ii) As the historians of probability theory have recognized, the Latin word *verisimilitudo* was often used as a synonym for the modern mathematical concept of probability. (iii) In the contemporary theories of literature and film, "verisimilitude" indicates either the illusory and truth-seeming character of poetic representation or the conformity of a narrative to what is generally accepted and expected by the public. This usage has also been adopted by some champions of qualitative methods in cultural and social research. (iv) In classical rhetoric, from the Romans to their medieval and renaissance successors, a special kind of narrative was called "possibly true", "probable" or *veri simile*. (v) The historians of ancient philosophy have debated on the interpretation of Cicero's terms *probabile* and *veri simile,* their roots in the Greek term *pithanon* of the Academic sceptic Carneades, and their later treatment in Augustine's *Contra academicos*.

It is remarkable that, while (iv) and (v) are the historical traditions behind the themes (iii) and (ii), respectively, they meet each other in Cicero's work. The verisimilitude of narratives was discussed in the works of poetics and rhetoric by Cicero, Horace, and Quintilian. Carneades' *pithanon* and Cicero's *veri simile*, with the principle that a sceptic may follow what-is-like-truth in practical action, have been interpreted in two main senses: as an anticipation of epistemic decision-theoretic probability (cf. ii), or as a guide to follow the appearances or what is commonly accepted (cf. iii). In this paper, I argue that ideas connected with the similarity approach to truthlikeness (cf. i) also played a role in the ancient debates. The roots of contemporary fallibilism, based upon the concept of truthlikeness, thus go back to the ancient school of Academic scepticism.

Philosophy of Science: Truthlike Theories

Karl Popper (1963, 1979) introduced the notion of verisimilitude in 1960 in order to make sense of the idea that one scientific theory may be closer to the truth than another. In his falsificationist campaign against inductivism, especially Rudolf Carnap's inductive logic, Popper urged that the scientists can critically test and falsify scientific theories, but they can never prove that a theory is true or even probable. According

to Popper, typical scientific theories are false, but still the succession of theories may be progressive in the sense that e.g. Einstein's theory is a better or closer approximation to the truth than Newton's mechanics.

While Popper did not succeed in demolishing inductive logic as a research programme in the philosophy of science, he convinced many critical scientific realists of the importance of the notion of truthlikeness.[2] He emphasized the distinction between logical and epistemic notions of verisimilitude: the former tells how close to the truth the rival theories really are, and the latter how close they can be claimed to be given our evidence. As truth is not a manifest property of theories, Popper's fallibilism allows that comparative claims about truthlikeness are themselves only conjectures, to be assessed by indicators like the "corroboration" of theories in experimental tests.

Popper's proposal for a comparative notion of truthlikeness was refuted in 1974 by David Miller and Pavel Tichý, who proved that all false theories are incomparable on Popper's account. Some philosophers have still followed Popper's attempt to explicate the logical notion of truthlikeness by using the notions of truth value (truth and falsity), logical entailment, and information content. But already in 1974 a new "similarity program" in defining truthlikeness was started by Tichý, Risto Hilpinen, and Ilkka Niiniluoto. They employed the notion of similarity (resemblance, likeness) between states of affairs (or their representations like possible worlds, structures, constituents etc.).[3] A further important distinction can be made between *approximate truth* (i.e. closeness to being true, defined by the condition that at least one of the states of affairs allowed by a theory is close to the truth) and *truthlikeness* (reflecting also the amount of information that a theory gives about the whole truth, defined by the condition that all of the states of affairs allowed by a theory are close to the truth).

While the details of these works are still open to debate, at least they indicate that Popper was right in distinguishing the logical concept of truthlikeness from the concept of probability. Niiniluoto's (1987) proposal for estimating unknown degrees of truthlikeness by their expected values relative to

---

[2] See Niiniluoto (1984, 1999).

[3] See Oddie (1986), Niiniluoto (1987, 1998a), Kuipers (2000).

posterior probabilities, given some empirical evidence, suggests also that the epistemic notion of verisimilitude has properties that are different from those of probability. For example, we may know that a hypothesis is false, and hence has the posterior probability zero, but still it may be judged to be close to the truth.

The explication of the concepts of approximate truth and truthlikeness is part of the contemporary debate about scientific realism.[4] But it also allows us to assess earlier, often confused ideas about "degrees of truth", "partial truth", and "approach to the truth". Comments on Hans Reichenbach's attempt to link truth and probability are given in Niiniluoto (1998b). Criticism of the British Hegelian F. H. Bradley is presented in Niiniluoto (1987), with remarks on the Marxist views about absolute and relative truth (Friedrich Engels, V. I. Lenin).

If "nearness to truth" is an idea which makes sense - *pace* the attacks by Ewing (1934), Quine (1960), and Laudan (1984)[5] - then we can also understand Charles S. Peirce's fallibilism as a form of "convergent realism": since the 1870s, Peirce emphasized the self-corrective nature of the scientific method, and upheld a dynamic view of truth as the "limit" of scientific inquiry (see Peirce, 1931-35). Earlier advocates of such a *strong fallibilism*, which takes science to approach to the truth via false theories, include Nicolaus Cusanus in the 15th century (see Cusanus, 1954) and Robert Boyle in the 17th century (see Boyle, 1996). Boyle's *The Sceptical Chemist* in 1661 used Carneades as a spokesman of critical inquiry.[6] It is clear that Popper's account of scientific progress in terms of increasing truthlikeness continues this tradition.

Probability

The mathematical theory of probability was created in the 1660s by Blaise Pascal and Pierre Fermat.[7] The primary applications of classical probability calculus were games of chance,

[4] Cf. Niiniluoto (1999), Kuipers (2019).
[5] Cf also Sellars (1968).
[6] See van Leeuwen (1970).
[7] See Hacking (1975).

where the probability of an event has some connection to its frequency of occurrence in repeated experiments. In this sense, as stated by G. W. F. Leibniz, probability is a non-epistemic "degree of possibility". But, in applications to scientific inference, probability was also linked with the incompleteness of our knowledge. In this sense, as stated by the Bayesian school, probability is an epistemic or doxastic concept expressing the "degree of belief" in the truth of some hypothesis on the basis of the available evidence.

In the 20th century, the frequentist theories of scientific inference were mainly based upon the Peircean idea that some methods are "reliable" in the sense that they produce true conclusions with a high frequency (Neyman-Pearson statistics, Reichenbach).[8] Bayesian approaches included both theories of subjective or personal probability (Bruno de Finetti, Frank Ramsey, L. J. Savage) and inductive logic (Rudolf Carnap, Jaakko Hintikka). A major trend of the contemporary philosophy of science can thus be characterized as *weak fallibilism* which rejects the idea of complete certainty in human knowledge but allows that scientific hypotheses have different degrees of probability relative to the evidence. In this view, scientific progress typically means that new empirical or experimental evidence makes our favourite theory more and more certain.

When Latin was still the language of science, mathematical treatises of probability often used the terms *probabilitas* and *verisimilitudo* as synonyms. It is also interesting to note that in many modern languages the concept of probability is expressed by terms that in some way include a reference to truth and similarity. In English, this connection is still seen in the terms *likelihood* and *likely*, whereas *probable* has connotations with "probing", "proving", "approving", and "accepting". For example, in German *Wahrscheinlichkeit* combines *Wahr* (= true) and *Schein* (= appearance). In Finnish, the adjective *todennäköinen* means "true-looking" or "truth-seeming", and in Swedish *sannolik* means "like the truth".

The terminological connection between probability and verisimilitude is much older than the mathematical theory of probability. It can be found among the Renaissance humanists,

[8] Cf. Goldman (1986).

like Lorenzo Valla in the sixteenth century, who studied "dialectical arguments" with "probable" premises.[9] Medieval scholastic philosophers assigned the concept of probability to a belief (Gr. *doxa*, Lat. *opinio*) on the basis of the number of authorities approving it; the debates on "probabilism" in ethics were conducted in these terms between the Jesuits and the Protestants in the 16th century.[10]

The roots of these discussions go back to Cicero who used in his *Academica* (45 BCE) the Latin terms *probabile* and *veri simile* as translations of the Greek term *pithanon* of Carneades. The historians of probability theory usually remember that Bishop Joseph Butler in 1736 claimed that "to us probability is the very guide of life", but Cicero is sometimes mentioned in this connection as well. Among professional historians of probability, Ivo Schneider (1977) has ventured to consider the views of Carneades. Some 20th century philosophers have also recognized this ancient connection. Richard Jeffrey (1984), one of the leading weak fallibilists in the subjectivist Bayesian school, has drawn a straight line from Carneades and Cicero via Montaigne to de Finetti, Ramsey, and Carnap. Popper - who preferred to consider Xenophanes as a precursor of his strong fallibilism[11] - once mentioned Cicero as the philosopher who introduced the terms "probability" and "verisimilitude", but he thought that Cicero used them in the subjective sense.[12]

Verisimilitude in Poetics

In the film *The Moderns* (1988) by Alan Rudolph, the heroine, after being unfaithful to her husband with a painter of art forgeries, outbursts: "How I hate the verisimilitude of Parisian life!". It is quite remarkable that the French term *vraisemblance*, translated as "verisimilitude" or "seeming-real", became fashionable in modern and postmodern film theory in the 1960s - around the same time when Popper introduced it in the philosophy of science. According to Dudley Andrew (1984), verisimilitude expresses the artificial and deceiving nature of

[9] See Jardine (1983).
[10] See Byrne (1968), Kantola (1994).
[11] Cf. Feyerabend (1984).
[12] See Popper (1963), 263, 404.

cinema as an illusory representation. Thus, this term does not indicate anything like an objective or truthlike presentation of reality, as with Popper, but rather the appearance (German *Schein*) of reality, i.e. our difficulty in recognizing a distinction between real and apparent representations in films.

In modern theories of literature, verisimilitude indicates the conventional character of literary stories: they should not include incredible and supernatural elements, but rather conform to what is generally accepted and expected by the audience.[13] As Päivi Mehtonen (1996) shows in detail, this requirement has its roots in the ancient theories of rhetoric and poetics. In his main work on rhetoric, *De Inventione* (81-80 BCE), Cicero distinguished three kinds of narratives: *historia* gives a true account of actual events, and *fabula* a purely fictional story, but *argumentum* should be "possibly true", at least according to the expectations of the public. Here Cicero follows Aristotle's *Poetics* (II.9): the poet's function is not to describe the thing that has happened, but a kind of possible thing that might happen.

According to Glucker (1995), when Cicero translated the Greek term *pithanon*, which in rhetoric means "persuasive" or "convincing", by *probabilis*, he was following the Latin verb *probare* in the meaning of "to convince the audience". In the first century CE, Horace in *Ars Poetica* required that a poetical work has to be *proxima veris* (close to the truth), and Quintilian in *Institutio Oratorio* characterized *argumentum* as a narration which is *vero simile*. As shown by Mehtonen (1996), this usage was continued in the 12th century theories of poetics by Thierry of Chartres and William of Conches.

The poetic notion of verisimilitude has recently inspired also scholars who develop "constructive" approaches in the humanities. The psychologist Jerome Bruner makes a distinction between well-formed arguments and good stories as two "modes of thought" or "ways of ordering experience, or constructing reality": the narrative mode does not establish truth but "lifelikeness" or "verisimilitude".[14] According to Bruner, art and the humanities are not constrained by the requirement of truth and testability in the scientists' sense, but they aim at

[13] See Todorov (1977).
[14] Bruner (1986), 11.

hypotheses that are recognizable as "true to conceivable experience" or "have verisimilitude" (*ibid*., 52). Appealing to Nelson Goodman's writings about "worldmaking", Bruner argues that there is no "aboriginal" reality against which one may compare a constructed possible world (*ibid*., 46).

A recent *Handbook of Qualitative Research* (1994) uses the notion of verisimilitude in defending the view that "objective reality" and "neutral observations" can never be captured in cultural and social research. Methodologies are viewed as "storytelling traditions". In rejecting "positivist" and "postpositivist" methods, "many members of the critical theory, constructivist, poststructural, and postmodern schools of thought" are seeking "alternative methods of evaluating their work", among them "verisimilitude, emotionality, personal responsibility, an ethic of caring, political praxis, multivoiced texts, and dialogues with subjects".[15] Verisimilitude is characterized as "a style of writing that draws the reader so closely into subjects' worlds that these can be palpably felt".[16] Even in the social sciences there are multiple genres, and each of these writing forms has its own standards. Truth and verisimilitude are different standards. A text is "a site for political struggle over the real and its meaning". Verisimilitude is "the mask a text assumes as it convinces the reader it has conformed to the laws of its genre".[17]

Academic Scepticism

One of the main topics in the study of Hellenistic philosophy is the debate between Stoic epistemology and the school of Academic scepticism[18]. According to the Stoic view, the "wise man" assents to impressions which are infallibly true (Gr. *phantasia kataleptike*), i.e. "arise from what is" and "could not arise from what is not". Thus, in the strong sense, such impressions are true, caused by the object, and cannot be false.[19] Arcesilaus, who became the head of Plato's Academy in 273 BCE, argued that there is no criterion for distinguishing true

[15] Denzin & Lincoln (1994), 5.
[16] Adler & Adler (1994), 381.
[17] Lincoln & Denzin (1994), 579-580.
[18] For the principal sources, see Long & Sedley (1988).
[19] Cf. Brittain (2001), 19.

impressions from the false ones, so that for every true impression there is an identical but false one; hence, the best attitude is *epoché*, i.e. suspension of judgment on all questions.

The Stoic counterargument pointed out that a person who withholds of assent on everything cannot act.[20] Arcesilaus replied that sufficiently "reasonable" (Gr. *eulogon*) impressions can serve as a criterion of action, but following such impressions need not involve any assent in the Stoic sense. His most important successor Carneades (c. 214-129 BCE) applied the term *pithanon* to impressions that are adequate for the conduct of life. He further developed the idea that *pithanai* impressions may have at least three different levels of "convincingness", ranging from vivid to unimpeded (undiverted) and thoroughly tested or scrutinized ones.[21] This doctrine was passed on to Philo of Larissa who was the main teacher of Cicero (106-43 BCE). Cicero translated *pithanotes* as *probabilitas* and *verisimilitudo*, thereby gaining the reputation of a thinker who treats probability as the guide of life. At the same time, Aenesidemus protested against the development of the Academy, and revitalized the Pyrrhonian school of scepticism, later exposed by Sextus Empiricus around 200 CE. When Augustine discussed scepticism in *Contra Academicos* in 385 CE, he took Cicero's *Academica* as his focus.

For the historical scholars, it is highly important to ask whether Arcesilaus and Carneades really approved the doctrines of *eulogon* and *pithanon*, or only expressed them as dialectical moves against the Stoic position.[22] Many commentators think that, as a consistent sceptic, Carneades did not make any concessions to dogmatism and thus did not himself assent to the doctrine of convincing impressions. Thus, one might suggest that Philo, who allowed the wise man to hold opinions, was a revisionist who misunderstood the real message of scepticism.[23]

In my view, the question about the dialectical interpretation is secondary to the problem of interpreting the doctrine formulated by Carneades. It may happen that he did not approve the

---

[20] See Stiker (1980).

[21] See Bett (1989).

[22] Cf. Couissin (1983), Schofield (1999).

[23] Cf. the discussion of Philo in Brittain (2001).

thesis that *pithanon* is the basis of human life without assent, but nevertheless this is an extremely interesting view, and whoever was the first to outline it deserves credit in the history of fallibilist epistemology. Long and Sedley (1987, 449, 460) may be right in concluding that Carneades' doctrine "was so brilliantly constructed that it gave Philo the foundation for a philosophy of fallibilism".

But what was the doctrine formulated by Carneades and adopted by Cicero? Several interpretations can be suggested.[24]

(a) As Cicero translated *pithanon* by *probabile*, Carneades has often been interpreted as a "probabilist",[25] comparable to modern empiricists and pragmatists.[26] Long and Sedley (1987, 458) argue that Carneades' classification of the levels of convincingness concerns impressions that are "apparently true" (relative to the percipient) rather than "true" (relative to the perceived object). In this sense, these levels of credibility are comparable to the degrees of belief of the Bayesian school of subjective probability, so that Carneades and Cicero appear as forerunners of weak fallibilism. This interpretation can be complemented by the observation that, at least for Cicero, there is a weak sense of accepting an impression as probable which does not involve a commitment to its truth.[27] In this view, the rule of following *pithanai* impressions in practical action means that many of our actions have to be based upon assumptions that are not certain but only more or less probable.

(b) Several scholars of ancient philosophy deny that Carneades was a probabilist. Myles Burnyeat (1983) argues that *pithanon* should not be mistranslated "probable", as its normal meaning is "persuasive" or "convincing". This is supported by the early use of this term in rhetoric: *peithestai* means to "be persuaded".[28] Malcolm Schofield (1999, 350) claims that there is "little in the evidence" supporting the probabilistic reading of Carneades, since the process of examining and testing of convincing impressions is "not articulated as a form of calculation of the likelihood that they are actually true". R. J.

---

[24] See Niiniluoto (2000).
[25] See Copleston (1966), 417; Popkin (1979), xiv.
[26] See Long (1986), 96.
[27] Cf. Frede (1984), Hankinson (1995).
[28] See Glucker (1995); Brittain (2001), 15.

Hankinson (1995, 111) states that the probabilistic interpretation of Carneades is a "fantasy". Burnyeat suggests that the controversial reading of *pithanon* as "a positive criterion of life" was due Philo of Larissa, and Philo was opposed by the Pyrrhonian Aenesidemus who argues that the sceptic's way to "live without belief" is "by keeping to appearances". In this view, acting on *pithanoi* impressions means "to follow the appearances". Using the term of Jonathan Barnes (1982), a sceptic can be an "urbane Pyrrhonist", who pays due regard to ordinary beliefs and customs, but applies *epoché* to philosophical and scientific claims about reality transcending the appearances. The alternative interpretation of Sextus Empiricus as a "rustic Pyrrhonist", who has no beliefs whatsoever, is debated in Burnyeat & Frede (1997).

(c) According to Sextus, Carneades claims that impressions that are convincing to the highest degree "tell the truth for the most part".[29] This might be a reference to Aristotle's frequentist notion of probability (Gr. *eikos*) as that which usually occurs. In this interpretation, to follow "convincing" impressions has a high degree of truth-frequency or reliability in Peirce's sense, as it leads to truth or success in many cases.

(d) One of Cicero's translations for *pithanon* was *veri simile*. This term had been used in rhetoric by some of his contemporaries (Plautus, Terence) as a translation of the Greek term *eikos* (likely, credible) which refers to arguments that are weaker than true or certain.[30] In tracing the roots of verisimilitude, this is a highly interesting fact. It leads us to ask whether any of the Academic sceptics used a concept which in some way indicates "closeness to the truth" - either objective similarity with the truth or an estimate of such similarity? Instead of the weakly fallibilist idea of probability as a corrigible criterion of more-or-less-convincing-truth, this would be a strongly fallibilist idea of falsities that are at some distance from the truth. It seems to me that a positive answer can be given in the cases of Augustine and Cicero.[31]

In his arguments against the Academic sceptics, Augustine (1950, 82, 85) appeals to an example:

---

[29] See Long & Sedley (1987), 452.

[30] Glucker (1995).

[31] See Niiniluoto (1987), 161; Niiniluoto (2000).

> "I beg you", said I, "give me your best attention. If a man who never saw your father himself, yet, on seeing your brother, asserted that he was like your father, would you not think that he was mentally affected or talking at random?" ... "It is obvious that in the same way are your Academics to be laughed at, since they say that in practical matters they follow 'what-is-like-truth', although actually they do not know what truth itself is."

In other words, the sceptic cannot know what-is-like-truth unless he already knows the truth, and the denial of the latter possibility excludes the former as well. This is, indeed, still today a standard argument against the theories of truthlikeness.[32] It was used against Cicero by Giulio Castellani in 1558 and Johannes Rosa in 1571.[33] What is more, Augustine's dialectical reconstruction of the sceptic's position involves a similarity relation. No one mistakes the brother for the father, i.e. the brother is not treated as being probably (convincingly, credibly) the father, but he is claimed to resemble his father. In the same way, our current theory may be known to be different from the true theory, but still it may be truthlike by resembling the true theory in important respects.

In Augustine's dialogue, the sceptic Licentius defends the view that a man who seeks the truth can be happy even if he never finds it (*ibid*., 40-42). Cicero's *Academica* repeatedly points out that the goal of such truth-seeking is to approach the truth, or to find something that is similar to the truth. For example, "the wise man" explores a position in order to advance "towards the actual truth, or indeed come as near to it as possible", but it may happen that his conclusions only appear "to resemble truth" but are "really far remote from truth" (II,xi). The purpose of philosophical discussions is to find "results that may be either true or the nearest possible approximation to the truth" (.. exprimant aliquid quod aut verum sit aut ad id quam proxime accedat) (II,iii). The difference between a Stoic and a sceptic is summarized as follows:

> ... while if a notion comes to us that appears to bear a likeness to the truth, the mind is filled with the most humanizing kind of pleasure. These researches therefore will be pursued both by your

---

[32] See Niiniluoto (1987), 265.

[33] See Schmitt (1972), 122, 150.

> wise man and by this sage of ours, but by yours with the intention of assenting, believing and affirming, by ours with the resolve to be afraid of forming rash opinions and to deem that it goes well with him if in matters of this kind he has discovered that which bears a likeness to truth. (II,xli)

Concluding Comparisons

Let us try to summarize our picture. The concept of verisimilitude was introduced by Roman scholars, especially Cicero and Quintilian, in two different contexts that they inherited from the Greeks. These two contexts have been historically largely separate. While Cicero was one of main philosophers for the Renaissance humanists through his writings on poetics and rhetoric, his epistemological work *Academica* was not well-known during the medieval times and became a topic of a lively debate and opposition during the sixteenth-century revival of ancient scepticism.[34]

The first context was rhetoric and poetics. Here *veri simile* was originally a supporting partial argument which is not convincing alone. But it also came to mean the special quality of fictional stories that are realistic, plausible, or possibly true by depicting persons and events that are at least partially similar to those that actually happen. This notion of verisimilitude is still alive in contemporary theories of literature and film. It has its subjective or pragmatic side: a plausible narrative does not violate the expectations of the audience. But it also has its objective side, grounded in the requirement of partial similarity with actual reality. This latter aspect links this notion to the recent similarity approach to truthlikeness within the philosophy of science.

Another context for the Roman scholars, again for Cicero in particular, was epistemology. Here *probabile* and *veri simile* were introduced as translations of the term *pithanon* of the leading Academic sceptic Carneades. A conceptual distinction between probability as credibility and verisimilitude as closeness to the truth was not made - this is understandable, as clarity in this matter has been achieved by philosophers only after

---

[34] See Schmitt (1972), Powell (1995).

Popper's work in the 1960s. In fact, even the scholars studying ancient epistemology never refer to Popper, or to the recent debates about truthlikeness, and Schofield (1999) does not mention Cicero's term *veri simile* in his discussion of probabilism.

The ancient discussions of sceptical doctrines, including the late commentary by Augustine, seem to contain three different ideas: (1) probability as epistemic credibility, (2) probability as statistical reliability, and (3) probability as what-is-like-the-truth. Augustine used arguments and examples where the third meaning involves the idea of similarity or resemblance, and similar locutions about proximity or closeness to the truth are employed by Cicero, both in relation to philosophical doctrines and sense impressions. The alternatives (1) - (3) anticipate later attempts in epistemology and philosophy of science to find "fallibilist" middle ways between dogmatism and scepticism.[35]

In the debates between the Stoics and the sceptics, the notions of probability and verisimilitude were also linked with human action. The Pyrrhonian sceptics (like Sextus Empiricus, 1985) defended life without beliefs which is achieved by following the appearances - by conforming to the common perceptions, the laws and customs of one's own culture, and the instruction of arts. This idea is connected with the use of "verisimilitude" as a public standard in literature. The instruction to follow the appearances, and to suspend judgment about reality behind them, is continued by the French postmodernists and some American neo-pragmatists.[36]

The probabilistic reading of *pithanon* as a guide of life anticipates the branch of the Bayesian school which denies the rationality of any inductive rules of acceptance (*pace* Levi, 1967), but allows the use of degrees of belief in practical decision making (see Carnap, 1971). The verisimilitude reading anticipates Popper's (1979) recommendation that what is judged as the most truthlike theory is used as a tool in problems of pragmatic preference between alternative actions.

---

[35] Cf. e.g. Popkin (1979); van Leeuwen (1970); Pappas and Swain (1978); Sihvola (2000).

[36] See Lyotard (1985) and Rorty (1989).

Verisimilitude is used also in the contemporary constructivist and postmodern approaches to literature, art, and qualitative social research. Here we again see a problem of communication: Bruner (1986) refers to Popper's *Objective Knowledge*, but in using the notion of verisimilitude he fails to note that Popper applies this same term in a quite different sense. More generally, within qualitative research the notion of verisimilitude is used in seeking alternatives to the realist view of science, while philosophers of science have followed Popper in employing this term in the defence of scientific realism.

It might seem that the connection between verisimilitude and action is lacking in poetics, as in normal situations we are hardly inclined to act upon a fictional story. However, in the postmodern culture there is an interesting phenomenon of fashion that Oscar Wilde called "life imitates art".[37] Fictional heroes of narratives, if sufficiently similar to us, may persuade us to follow their example in our everyday activities.

I hope that the conceptual map drawn in this chapter helps all of us who wander among the perplexing applications of the notion of verisimilitude.

Note. An earlier version of this chapter was presented at the Conference of the German Society for Analytic Philosophy, in Munich, on September 16, 1997.

Literature

Adler, P.A. and Adler, P. (1994). "Observational Techniques". In Denzin & Lincoln, 377-392.

Andrew, Dudley (1984). *Concepts in Film Theory.* Oxford: Oxford University Press.

Augustine, St. (1950). *Against the Academics.* Westminster: The Newman Press.

Barnes, Jonathan (1982). "The Beliefs of a Pyrrhonist", *Proceedings of the Cambridge Philological Society* N.S. 28, 1-29. Reprinted in Burnyeat & Frede (1997), 58-91.

Bett, Richard (1989). "Carneades' Pithanon: A Reappraisal of its Role and Status". In Julia Annas (ed.), *Oxford Studies in Ancient Philosophy* 7. Oxford: Oxford University Press, 59-94.

---

[37] Cf. Bruner (1987).

Boyle, Robert (1996). *A Free Enquiry into the Vulgarly Received Notion of Nature.* Cambridge: Cambridge University Press.

Brittain, Charles (2001). *Philo of Larissa: The Last of the Academic Sceptics.* Oxford: Oxford University Press.

Bruner, Jerome (1986). *Actual Minds, Possible Worlds.* Cambridge, MA: Harvard University Press.

Bruner, Jerome (1987). "Life as Narrative", *Social Research* 54, 11-32.

Burnyeat, Myles (1980). "Can the Sceptic Live His Scepticism?". In Schofield *et al.*, 20-53. Reprinted in Burnyeat (1983), 117-148.

Burnyeat, Myles (ed.) (1983). *The Skeptical Tradition.* Berkeley: University of California Press.

Burnyeat, Myles & Frede, Michael (eds.) (1997). *The Original Sceptics: A Controversy.* Indianapolis: Hackett.

Byrne, Edmund (1968). *Probability and Opinion: A Study of Medieval Presuppositions of Post-Medieval Theories of Probability.* The Hague: M. Nijhoff.

Carnap, Rudolf (1971). "Inductive Logic and Rational Decisions". In R. Carnap & R. Jeffrey (eds.), *Studies in Inductive Logic and Probability* vol. I. Berkeley: University of California Press, 5-31.

Cicero (1956). *De natura deorum, Academica.* London: William Heinemann.

Copleston, Frederick (1966). *A History of Philosophy* vol. I. London: Burns and Oates.

Couissin, Pierre (1983). "The Stoicism of the New Academy". In Burnyeat (1983), 31-63.

Cusanus, Nicolaus (1954). *Of Learned Ignorance.* London: Routledge and Kegan Paul.

Denzin, N. K. & Lincoln, Y. S. (eds.) (1994). *Handbook of Qualitative Research.* London: SAGE Publishers.

Ewing, A. C. (1934). *Idealism: A Critical Survey.* Strand: Methuen.

Feyerabend. Paul (1984). "Xenophanes: A Forerunner of Critical Rationalism?". In G. Andersson (ed.), *Rationality in Science and Politics.* Dordrecht: D. Reidel, 95-109.

Frede, Michael (1984). "The Sceptic's Two Kinds of Assent and the Question of the Possibility of Knowledge". In R. Rorty, J. B. Schneewind & Q. Skinner (eds.), *Philosophy in History: Essays on the Historiography of Philosophy.* Cambridge: Cambridge University Press. Reprinted in Burnyeat & Frede (1997), 127-151.

Glucker, John (1995). "*Probabile, Veri Simile*, and Related Terms". In Powell, 115-143.

Goldman, Alvin I. (1986) *Epistemology and Cognition.* Cambridge, MA: Harvard University Press.

Hacking, Ian (1975). *The Emergence of Probability.* Cambridge: Cambridge University Press.

Hankinson, R. J. (1995). *The Sceptics.* London: Routledge.

Jardine, Lisa (1983). "Lorenzo Valla: Academic Skepticism and the New Humanist Dialectic". In Burnyeat (1983), 253-286.

Jeffrey, Richard (1984). "De Finetti's Probabilism", *Synthese* 60, 73-90.

Kantola, Ilkka (1994). *Probability and Moral Uncertainty in Late Medieval and Early Modern Times.* Helsinki: Luther-Agricola-Society.

Kuipers, Theo (2000). *From Instrumentalism to Constructive Realism.* Dordrecht: Kluwer.

Kuipers, Theo (2019). *Nomic Truth Approximation Revisited.* Cham: Springer.

Laudan, Larry (1984). *Science and Values.* Berkeley: University of California Press.

Leeuwen, Henry G. van (1970). *The Problem of Certainty in English Thought 1630-1690.* Amsterdam: Nijhoff.

Levi, Isaac (1967). *Gambling With Truth.* New York: Alfred A. Knopf.

Lincoln, Y. S. & Denzin, N. K. (1994). "The Fifth Moment". In Denzin & Lincoln, 575-586.

Long, A. A. (1986). *Hellenistic Philosophers: Stoics, Epicureans, Sceptics*, 2nd ed. Berkeley: University of California Press.

Long, A. A. & Sedley, David (eds.) (1988). *The Hellenistic Philosophers* (Cambridge: Cambridge University Press).

Lyotard, Jean-Francois (1985). *The Postmodern Condition: A Report on Knowledge.* Minneapolis: University of Minnesota Press.

Mehtonen, Päivi (1996). *Old Concepts and New Poetics: Historia, Argumentum, and Fabula in the Twelfth- and Early Thirteenth-Century Latin Poetics of Fiction.* Helsinki: Societas Scientiarum Fennica.

Niiniluoto, Ilkka (1984). *Is Science Progressive?* Dordrecht: D. Reidel.

Niiniluoto, Ilkka (1987). *Truthlikeness.* Dordrecht: D. Reidel.

Niiniluoto, Ilkka (1998a). "Verisimilitude: The Third Period", *The British Journal for the Philosophy of Science* 49, 1-29.

Niiniluoto, Ilkka (1998b). "Truth, Probability, and Simplicity - Comments on Hans Reichenbach's Probabilistic Empiricism". In Hans Poser & Ulrich Dirks (eds.), *Hans Reichenbach - Philosophie im Umkreis der Physik.* Berlin: Akademie Verlag, 69-87.

Niiniluoto, Ilkka (1999). *Critical Scientific Realism.* Oxford: Oxford University Press.

Niiniluoto, Ilkka (2000). "Scepticism, Fallibilism, and Verisimilitude". In Sihvola, 145-169.

Oddie, Graham (1986). *Likeness to Truth.* Dordrecht: D. Reidel.

Pappas, George & Swain, Marshall (eds.) (1978). *Essays on Knowledge and Justification*. Ithaca: Cornell University Press.

Peirce, Charles S. (1931-35). *Collected Papers* vols. 1-6, ed. by C. Hartshorne and P. Weiss. Cambridge, MA: Harvard University Press.

Popkin, Richard H. (1979). *The History of Scepticism from Erasmus to Spinoza.* Berkeley: University of California Press.

Popper, Karl R. (1963). *Conjectures and Refutations.* London: Routledge and Kegan Paul.

Popper, Karl R. (1979). *Objective Knowledge,* rev. ed. Oxford: Oxford University Press.

Powell, J. G. F. (ed.) (1995). *Cicero the Philosopher*. Oxford: Oxford University Press.

Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: The MIT Press.

Rorty, Richard (1989). *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.

Schmitt, Charles B. (1972). *Cicero Scepticus: A Study of the Influence of the Academica in the Renaissance.* The Hague: Martinus Nijhoff.

Schneider, Ivo (1977). "The Contributions of the Sceptic Philosophers Arcesilaus and Carneades to the Development of an Inductive Logic Compared with the Jaina-Logic", *Indian Journal of History of Science* 12, 173-180.

Schofield, Malcolm (1999). "Academic Epistemology". In K. Algra, J. Barnes, J. Mansfeld & M. Schofield (eds.), *The Cambridge History of Hellenistic Philosophy.* Cambridge: Cambridge University Press, 323-351.

Schofield, Malcolm, Burnyeat, Myles & Barnes, Jonathan (eds.) (1980). *Doubt and Dogmatism: Studies in Hellenistic Epistemology*. Oxford: Oxford University Press.

Sellars, Wilfrid (1968). *Science and Metaphysics.* London: Routledge and Kegan Paul.

Sextus Empiricus (1985). *Selections from the Major Writings on Scepticism, Man, & God*, ed. by P.P. Hallie. Indianapolis: Hackett.

Sihvola, Juha (ed.) (2000). *Ancient Scepticism and the Sceptical Tradition.* Helsinki: The Philosophical Society of Finland.

Striker, Gisela (1980). "Sceptical Strategies". In Schofield *et al*., 54-83.

Todorov, Tzvetan (1977). *The Poetics of Prose.* Oxford: Blackwell.

Chapter 8: Is It Rational to Be Rational?

For the classical Greek philosophers, the cultivation of human rationality was a central ingredient of education and edification (Gr. *paideia*). But the notions of reason and rationality have received a great variety of interpretations in philosophical discussions about human life, thinking, and action – and, apart from the educational context, they have been applied to many special areas in society (such as science, law, and politics). These different developments lead us directly to the general philosophical queries: *What is rationality? Why should we be rational?*

In this chapter, I consider only briefly the first question by distinguishing three aspects of rationality. Then I shall use three notions to give nine reformulations of the puzzling question "Is it rational to be rational?". My main task lies in the analysis of the relevant questions, not in their detailed answers. I hope this approach helps us to understand in a clearer way the nature and importance of human rationality.

The Rationality of Rationality: A Puzzling Question

Why be rational? It is tempting to offer the straightforward reply: because it is rational! Black is black, and rationality is rational. Indeed, the question

(Q) Is it rational to be rational?

sound so strange for everyday thinking that it must be philosophical – and perhaps one that should be eliminated or dissolved as a pseudo-problem.

An argument along these lines has been put forward in connection with Hume's problem of induction, G. H. von Wright (1957) maintains that the rationality of induction is a "disguised tautology", since the "inductive character of a policy is the very criterion by means of which we judge its goodness"

(p. 175). Hence, Hume's worry about the impossibility of justifying induction is in some stronger sense "grammatical" and again a "disguised tautology" (p. 178). Peter Strawson (1952) claims that the rationality of induction is a matter of what we mean by the word "rational": there is not much point in asking whether it is reasonable to place reliance on inductive procedures, since in such contexts "being reasonable" means that convictions are proportioned to the strength of inductive evidence. To ask for a justification of induction in general is, according to Strawson, like asking the question: Is the law legal? While a particular action may be legal relative to a system of law, "it makes no sense to inquire in general whether the law of the land, the legal system as a whole, is or is not legal. For to what legal standards are we appealing?".

Still, it is well-known that the surface form of a statement does not always reveal its hidden structure. For example, the statement "Boys are boys" has a non-tautological reading, as the same words may have different meanings even within one sentence. Similarly, questions which have the same form as (Q), e.g. "Is it happy to be happy?" and "Is it fun to be funny?", may be perfectly legitimate. Being funny for others can be sad for oneself, as witnessed by many tired and tragic clowns and comedians.

In the next two sections, I shall consider the traditional dichotomy between reason, emotion, and will and then propose three basic meanings of the term "rationality". This allows us to distinguish several interesting readings of the question (Q).

Reason, Emotion, and Will

The English word *rational* comes from the Latin *ratio*, which means the faculty of human *reason*. The verb "to reason" refers to the process of thinking or "reasoning", and "reason" is also a ground presented in favor of the conclusion of an argument. A *reasonable* person is one who is guided by reason or appropriately moved by reasons.[1]

The Greek terms for theoretical reason (thinking) we *nuus* and *logos*, and for practical reason (deliberation) *phronesis* – in Latin *prudentia*. Related concepts include knowledge (Gr.

[1] See Siegel (1988). Cf. also Aarnio (1987).

episteme, Latin *scientia*), wisdom (Gr. *sophia*, Lat. *sapientia*), understanding and intelligence. When Aristotle gave his famous definition of man as "a rational animal", descriptively he was asserting that the essence of human beings as a species is the faculty of reason. Normatively he implied that man ought to be rational or reasonable in the sense of being guided by reason, i.e. the humans ought to think and act in accordance with rational principles of *logos* and *phronesis*. Carl von Linné's species name for the humans, *Homo sapiens*, is an expression of this intellectualist conception. In his *Topics*, Aristotle further stated that humans differ from other animals in their ability to learn grammar – this view was expressed by Ernst Cassirer (1944) with his notion of *animal symbolicum*. The use of reason makes possible the endeavors in science and technology, culture and art - as well as the realization of individuals as self-conscious persons who are morally responsible for their actions. This is the grand idea of Enlightenment, expressed by Immanuel Kant in his slogan *Sapere aude!*, and continued today by the teaching of critical thinking.[2]

Aristotle knew that sometimes we humans act against our better knowledge - he called this phenomenon *akrasia* (weakness of the will). But he was convinced that the "nutritive" and "sensitive" parts of the human soul learn to obey reason in an educated person. The intellectual virtues – wisdom, understanding and practical reason – can be reached by experience and education. Similarly, other virtues – justice, prudence, and bravery – can be learned by practicing virtuous actions.

A similar view was formulated by Plato in his dialogue *Phaedrus,* where he elaborated the contrast between temperance (judgment guides rationally toward what is best) and wantonness (desire drags us irrationally towards pleasure) (258a). According to Plato, the human soul can be compared to a charioteer driving two horses, one good (spiritedness, Gr. *thumos*) and the other not (erotic desire, Gr. *eros*) (253d). The philosopher-psychologist Eino Kaila (1943) interpreted this allegory so that human action - even in our highest "deep-mental" aspirations - receives its powers and energy from the lower needs of the "human animal". But the aim of education is to encourage the more noble urge or passion (i.e. *thumos*) to

[2] See Siegel (1988).

join the logos, so that the drive of the wild desire comes to serve the valuable aims of good life. He also compared this dynamic multi-layered conception of the human mind-and-body to Sigmund Freud's psychoanalytic theory of the tripartite composition of man (super-ego, ego, id).

An exaggerated faith in reason was visible in the modern trend of epistemology known as "rationalism" and in its culmination within German idealism. Its modern champions (Descartes, Spinoza, Wolff, Hegel) claimed that knowledge can be obtained by the a priori use of human reason. The British school of empiricism (Bacon, Locke, Hume, Reid) was a healthy opposition in its defense of the epistemological significance of observation and experiment. Kant attempted to reconcile rationalism and empiricism with his doctrine of synthetic a priori truths, but a better compromise was developed by the logical empiricists, as summarized by Eino Kaila in his 1939 book on human knowledge: knowledge in the formal science (logic, mathematics) is analytic and a priori, while factual knowledge in the real sciences is synthetic and a posteriori.[3]

It would be naïve to assume that human beings always and everywhere, by their intrinsic nature, are moved by reasons.[4] Being worried about uncontrollable passions, the ancient Stoic philosophers regarded emotions as false judgments, and proposed a therapy of desire, with the goal that the wise man should get rid of all emotions (Gr. *apatheia*).[5] But David Hume argued in 1739 that all sensible creatures are slaves of their passions (Hume, 1969). Blaise Pascal had already in the 1650s exclaimed that "the heart has its reasons, which reason does not know". For the 19th century romantic thinkers and poets, who devalued the Enlightenment commitment to scientific reasoning, urged that the highest forms of human culture were based upon sentiments and imagination. The same contrast is continued in the postmodernist revolt against discursive reason in

---

[3] See Kaila (2014).

[4] The illusion of intellectualist psychology, which simply equates descriptive and normative rationality (cf. Elster, 1988), has been challenged by modern psychologist from Sigmund Freud to Amos Tversky & Daniel Kahneman.

[5] See Knuuttila (2004).

favor of the free flow of human desires. The sharp dichotomy between cognition and emotion was fashionable in the mid-twentieth analytic philosophy with its emotivist theories of meaning, art, and ethics.

A more balanced way of treating the interplay between cognition and emotion was developed in the latter part of the last century. With inspiration from phenomenology (Husserl and Sartre), analytic philosophers stared to talk about the "logic", "intentionality", and "rationality" of emotions.[6] Emotions were recognized as powerful motives for knowledge-seeking, and sometimes sensations and feelings may directly contribute to our knowledge about our bodies and minds.[7]

Kant's *Anthropologie* in 1798 consolidated the threefold division of human faculties into reason, emotion, and will (Lat. *voluntas*). The 19th century philosophy was a battle-field between rationalists (like Hegel) and their voluntarist critics (like Schopenhauer and Nietzsche). The notion of *thumos* can be regarded as a precursor of the concept of will, as mental power without a definite direction, while the directed notion of will is related to the concepts of want, wish, and intention. Important aspects of the "conative" attitudes were presented in Spinoza's deterministic system. The classical debate about the "freedom of the will" concerns the conditions and possibilities for choices in human behavior, which are autonomous with respect to external effects and internal affects.

Three Notions of Rationality

The opposite to rational is *irrational* or anti-rational which is in conflict or contradiction with reason. Beyond the two opposites there is the neutral area of *arationality* or non-rationality, where the distinction between rational and irrational is irrelevant. It would be unrealistic and too strained to demand that we ought to maximize our rationality in all of our affairs: there is a domain of arationality in everyday thought and actions, such as impulsive playful behavior and expressions of emotions (e.g. I let my mind play with an idea in daydreaming, I whistle while walking, in passing I rumple my wife's hair),

[6] See Rorty (1980).
[7] See Niiniluoto (2006).

where the question of rationality dos not, or need not, arise at all.[8]

Nicholas Rescher (1988) states that "rationality consists in the intelligent pursuit of appropriate ends", and "deliberately doing the best one can with the means of our disposal". Thus, rationality is systematic use of reason to give grounds for beliefs, actions, and values. For Rescher, the three forms of rationality are "cognitive" (What to believe or accept?), "practical" (What to do or perform?), and "evaluative" (What to prefer or prize?). A similar three-fold division is used in this chapter.

First, *cognitive rationality* (or C-rationality) concerns human beliefs and knowledge claims. Plato's classical definition equates knowledge with justified true beliefs – and thereby separates cognition from errors and mere opinions. The standards for rational beliefs are studied in epistemology. Rational belief systems (such as scientific theories and world views) are studied in the philosophy of science. Rational changes of belief by means of inference are investigated in logic. Thus, logical rationality is a part of cognitive rationality. The mainstream view in epistemology is Charles Peirce's *fallibilism*, which maintains that all factual claims are uncertain and corrigible by new evidence – and yet the critical methods of inquiry guarantee that science is the most reliable source of rational beliefs. In particular, it is cognitively rational to replace supernatural and religious doctrines by natural scientific explanations (cf. Chapter 9 below).

Secondly, *instrumental rationality* (or I-rationality) concerns the use of effective or optimal means for realizing the given goals. The appropriate relationship between means and ends is the basic idea of Max Weber's *Zweckrationalität*. The key concepts in this field include effectivity, efficiency, utility, instrumental goodness, technical norm, and practical syllogism. Instrumental rationality is studied in decision theory, game theory, utility theory, management science, operations research, praxeology, action theory, planning theories, and the various branches of technology and applied research.

Thirdly, *value rationality* (or V-rationality) concerns the appraisal of the goals of human action. If we ask, whether a goal

---

[8] See Hursthouse (1991).

is realizable by available means or useful for some other purpose, we are still within the domain of instrumental rationality. But if we ask, whether the goal as such is intrinsically valuable, not only accessible or instrumentally valuable, we come to Weber's *Wertrationalität*. The study of value rationality in this sense belongs to value theory and ethics.

Classical utilitarianism (Bentham, Mill) assumed that happiness is the ultimate goal of human life, and then evaluated particular acts by their instrumental effectivity in promoting this goal.[9] A larger domain of intrinsic values can be included in those theories of ethics that are based on the ideas of virtues, rights, and duties.[10] V-rationality is also defended by theories which assume value realism (Plato) or objectivism (Kant, Chaim Perelman, Jürgen Habermas). Radical relativism, which takes values to be merely expressions of personal wants and emotions or social interests, does not leave room for value rationality. Also many naturalist philosophers have denied the possibility of normative ethics: the only viable approaches are moral psychology and sociology, which study the valuations that people have in different times and cultures, and practical ethics, which clarifies real-life ethical problems by investigating the hidden premises in dialectical argumentation about morality. However, moderate value relativism, which sees values as social constructions in the Popperian World 3, does not exclude the possibility and even necessity of rational discussion of values.[11]

Habermas (1990) attempted to unify cognitive and value rationality in his *discursive* approach to epistemology and ethics. I am skeptical about the possibility of justifying the validity of values and norms by means of an ideal communication situation. For factual truth, the consensus theory is clearly insufficient, as it does not contain non-discursive rules for the causal interaction between the investigator and external reality. Such a causal interaction between the scientific community and the

---

[9] An improvement of act-utilitarianism is *rule-utilitarianism* which asks which moral rules (e.g. truth-telling instead of lying) will promote the greatest general good for everyone (see Frankena, 1963, 30).

[10] Virtue ethics belongs to the Aristotelian tradition. For a modern version, see von Wright (1963).

[11] See Niiniluoto (2009).

"real things"- e.g. by means of experiments as "questions put to nature" – was a crucial element of Peirce's characterization of truth as the limit of inquiry.

According to value constructivism, absolute value statements of the form "A is good/right" are incomplete and lack truth values. But relativized value statements of the form "A is good/right for X", where X is a community, and "A is good/right in a value system V", where V might be defined by written public documents and doctrines, have truth values. Such values can be investigated by empirical means or by hermeneutical methods of interpretations.[12]

In English discussions, a distinction is often made between "rational" and "reasonable".[13] It seems that "rational" often refers what we have called instrumental rationality, while "reasonable" refers to the nature of the goals and thereby to value rationality. Indeed, according to John Dewey's *Logic* (1938, 9), rationality is "an affair of the relation of means and consequences". Dewey (1939) also rejected the distinction between instrumental value and intrinsic values, but allowed that science should study (besides objects of actual appraisals) "ends-in-view" or "plans" which function as "directive means" for the future.[14]

It is important to emphasize, though, that the distinction between means and ends is not absolute. In evaluating a particular human act, we may ask about its instrumental rationality by investigating the link between its means and intended end, but also its links to unintended consequences (i.e. the so called side effects). An end may bad, if it is utopian or its achievement requires too much time, energy, and inconvenience. We may also evaluate both the ends and the means for their value rationality. The intended end may be useful for some other purpose, so that intrinsic value and instrumental value do not exclude each other: true theories in science and beautiful

---

[12] For value constructivism, see Niiniluoto (2009).

[13] Cf. Perelman (1979), Aarnio (1987).

[14] Dewey (1939, 28) rejects the notion of intrinsic value as a fallacy of the "nonnaturalistic school". But here he fails to see that intrinsic values need not be Platonist entities but could be understood as human social constructions in Popper's World 3 (see Niiniluoto, 2009).

works of art have intrinsic value, yet they may have economic value in practical applications and the art market of galleries.

In environmental philosophy and business ethics, we frequently encounter situations where the demands of I- and V-rationality are in conflict, since the most effective means for some end may be morally harmful. The conflicts between science and religion (e.g. Darwin's theory of evolution) also show that C-rationality may lead to results which challenge socially accepted doctrines of V-rationality.

Rescher (1988) argues that cognitive, practical, and evaluative rationality are "inseparably intertwined". I agree that the three types of rationality may support each other, but there are also potential conflicts between them. This gives a special flavor to the analysis of the question (Q).

Is it Rational to be Rational?: Nine Questions

The distinction between cognitive (C), instrumental (I), and value (V) rationality gives us a method of reformulating the perplexing question (Q), "Is it rational to be rational?", in nine different ways, obtained by replacing the first and second occurrences of "rational", respectively, by C, I, or V. Brief comments on each of the nine questions are given below.

(Q1) Is it C-rational to be C-rational?

Comment: It has been often claimed, by Michael Polanyi for example, that the method of critical doubt cannot be justified by non-dogmatic doubt, but only by faith. Karl Popper (1945) argued that his "critical rationalism" as a form of scientific world view cannot be adopted by a rational choice, but has to be based upon moral principles and "irrational faith in the power of reason". Now of course it is possible that someone uncritically assumes or accept science as the only rational world outlook, which is expected to give final solutions to all cognitive problems. In this sense, "scientism" can be an irrational or religious position. But still Polanyi's and Popper's claim is incorrect, as argued by William W. Bartley, III (1987) in his "comprehensive" or "pancritical rationalism". For a Peircean fallibilist, progressive science is self-corrective, so that both scientific knowledge and method are C-rational in their own ways: not only the scientific world view at any

moment, but also the method of science is always open to criticism and improvement. The scientific method should not be adopted in a dogmatic way, but as a cognitive procedure which can be critically studied and revised.

(Q2) Is it I-rational to be C-rational?

Comment. It is an important idea of William James (1907) that "truth pays". If cognitive rationality increases our chances of finding truth, then it also improves our chances of practical success in instrumental action. For example, research-based medical treatments are more reliable than unscientific tricks of charlatans. This is also Rescher's (1988) rationale of rationality, although (as a fallibilist methodological pragmatist) he points out that practical success cannot be proved in a strict sense: at best it can be claimed that "the counsel of reason represents the most promising prospect of realizing our objectives" (p. 33). Siegel (1988) argues that our possession of knowledge helps to protect us from manipulation. It may be added that rational and public methods of knowledge-seeking save us from various kinds of miseries (persecutions, fights, wars) that may result from dogmatic and irrational belief systems.

(Q3) Is it V-rational to be C-rational?

Comment. We are here asking whether commitment to knowledge is value rational. According to Aristotle, life with reason is good and valuable as the essence of humanity. This idea also expresses the core of the Enlightenment ethics: it is always better to know than to be ignorant. Siegel (1988) defends this view by suggesting the C-rationality demands public justification of beliefs and thereby is constitutive of the respect for persons and the preparation for adulthood and democratic forms of life. Von Wright (1993) argues that the search for rational reasons enhances our freedom.

(Q4) Is it C-rational to be I-rational?

Comment: This is the basic idea Francis Bacon's program of the scientification of technology: instrumental rationality necessarily presupposes cognitive rationality, since the links between means and ends have to be established by scientific research. Applied science in the late 19th century is the result

of this insight.[15] The same principle is behind the Marxist principle that "practice is the criterion of truth": even though erroneous beliefs may accidentally lead so successful consequences, practical success is typically a reliable indicator of the truth or truthlikeness of a theory.

(Q5) Is it I-rational to be I-rational?

Comment: The answer is trivially yes in many cases, but not always. Instrumental rationality may be understood in different ways relative to various criteria. For example, it is known from Technology Assessment (TA) that a tool or a procedure may be effective relative to its intended use, and also efficient relative to economic cost-benefit considerations, but still due to their overall side-effects harmful for aesthetic, ergonomical, ecological, and social purposes (cf. Chapter 20 below). Indeed, many central institutions of instrumental rationality (e.g. economic and urban planning, bureaucratic administration, new public management) can be criticized for being too narrow and one-sided – and thus failing to be sufficiently I-rational.

(Q6) Is it V-rational to be I-rational?

Comment. An unqualified positive answer is given by the technocratic ideology, which limits the domain of rationality to I-rationality, and similarly by those capitalists who take money to be an intrinsic value, to be pursued and accumulated for its own sake. Such views have been effectively criticized by the Frankfurt School. According to Max Horkheimer (1947), "man's urge to dominate nature" is the "disease of reason", and "denunciation of what is currently called reason is the greatest service reason can render". In spite of Horkheimer's demand that the two concepts of reason (i.e. I and V) should be "reconciled", these formulations seem to suggest that for him I-rationality as such is something to be rejected. I can agree that I-rationality is dangerous when effectiveness suggests actions which are criminal or unethical (e.g. robbing a bank in order to gain money quickly) or when it is combined with V-irrational ends (e.g. nuclear weapons in warfare, face recognition systems in oppressive surveillance of citizens). But one may also join Rescher (1988) in thinking that, when the goals

[15] See Niiniluoto (1993).

are appropriate, it is for us an important responsibility to attune the means to the ends. For example, we should rationally prefer planes whose security systems have been tested by critical scientific methods.

(Q7) Is it C-rational to be V-rational?

Comment. This is thesis of rationalistic ethics, from Plato (virtue is knowledge) to Perelman (agreement in universal audience) and Habermas (discourse ethics). Some moral realists have suggested that values can be defined by natural conditions, which concern e.g. human needs.[16] Even for a moderate moral relativist, knowledge is relevant to the rational discussion of values, but – due to Hume's guillotine and the fact-value distinction - appraisals of goals are not reducible to purely cognitive considerations.

(Q8) Is it I-rational to be V-rational?

Comment. It may be sometimes, but not always. Positive examples are provided by products (e.g. leadless gas, energy without carbon dioxide wastes) which satisfy our ecological values and at the same time, due to their prize and demand, are useful and profitable both for the producer and the consumer. But in many cases deep commitments to intrinsic values (e.g. to one's occupation, country, or god) are held even at the risk of one's health and life.

(Q9) Is it V-rational to be V-rational?

Comment: Finally we ask whether it is valuable to be value rational. This is denied by opportunist and situationist ethical views: intrinsic value commitments are not needed anywhere, and each value conflict is solved as a power struggle between the interested parties. But those who subscribe to some principles of V-rationality usually accept the meta-statement that it is good and humane to have intrinsic values, even when they are not instrumentally useful.

The conclusion from these nine questions is that, with some important qualifications, it is generally reasonable to be rational – and rational to be reasonable.

---

[16] Cf. Pihlström (2005) for an interesting defense of "pragmatic moral realism".

Note. This chapter is an updated version of a contributed paper in the XX World Congress of Philosophy ("Paideia") in Boston in August 1998. It has been published as "Is It Rational To Be Rational?", in R. Cobb-Stevens (ed.), *The Proceedings of the Twentieth World Congress of Philosophy, Vol. 5, Epistemology*, Bowling Green, OH: Philosophy Documentation Center, 2000, 115-122.

Literature

Aarnio, Aulis (1987). *The Rational as Reasonable: A Treatise on Legal Justification*. Dordrecht: D. Reidel.

Bartley, W. W. III. (1987). "Theories of Rationality". In Gerard Radnitzky & W. W. Bartley, III (eds.), *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*. La Salle, IL: Open Court, 205-214.

Cassirer, Ernst (1944). *An Essay on Man: An Introduction to a Philosophy of Human Culture*. New Haven: Yale University Press.

Dewey, John (1939). *Theory of Valuation*. Chicago: The University of Chicago Press.

Burnyeat, Myles (ed.) (1983). *The Skeptical Tradition.* Berkeley: University of California Press.

Elster, Jon (1988). "Rationality". In Guttorm Fløistad (ed.), *Contemporary Philosophy, vol. 2*. The Hague: Nijhoff.

Frankena, William K. (1963). *Ethics*. Englewood Cliffs, NJ: Prentice-Hall.

Habermas, Jürgen (1990). *Moral Consciousness and Communicative Action*. Cambridge, MA: The MIT Press.

Horkheimer, Max (1947). *The Eclipse of Reason*. New York: Oxford University Press.

Hume, David (1969). *A Treatise of Human Nature*. Harmondsworth: Penguin Books.

Hursthouse, Rosalind (1991). "Arational Actions", *The Journal of Philosophy* 88, 57-67.

James, William (1907). *Pragmatism*. New York: Longmans, Green & Co.

Kaila, Eino (1943). *Syvähenkinen elämä*. Helsinki: Otava.

Kaila, Eino (2014). *Human Knowledge: A Classic Statement of Logical Empiricism*, transl. by Anssi Korhonen. Chicago: Open Court.

Knuuttila, Simo (2004). *Emotions in Ancient and Medieval Philosophy*. Oxford: Oxford University Press.

Niiniluoto, Ilkka (1993). "The Aim and Structure of Applied Science", *Erkenntnis* 38, 1-13.

Niiniluoto, Ilkka (2006). "Cognition and Emotion". In Vesa Hirvonen, Toivo J. Holopainen & Miira Tuominen (eds.), *Mind and Modality: Studies in the History of Philosophy in Honour of Simo Knuuttila*. Leiden: Brill, 371-382.

Niiniluoto, Ilkka (2009). "Facts and Vales – A Useful Distinction". In Henrik Rydenfelt (ed.), *Pragmatist Perspectives* (Acta Philosophica Fennica 86). Helsinki: Philosophical Society of Finland, 109-133.

Perelman, Chaim (1979). "The Rational and the Reasonable". In *The New Rhetoric and the Humanities: Essays on Rhetoric and its Applications*. Dordrecht: D. Reidel, 117-123.

Pihlström, Sami (2005). *Pragmatic Moral Realism: A Transcendental Defense*. Amsterdam: Rodopi.

Rescher, Nicholas (1988). *Rationality*. Oxford: Oxford University Press.

Rorty, Amelie (1980). *Explaining Emotions*. Berkeley: University of California Press.

Siegel, Harvey (1988). *Educating Reason: Rationality, Thinking, and Education*. London: Routledge.

Strawson, Peter (1952). *Introduction to Logical Theory*. London: Methuen.

von Wright, Georg Henrik (1957). *The Logical Problem of Induction*. Oxford: Blackwell.

von Wright, Georg Henrik (1963). *The Varieties of Goodness*. London: Routledge & Kegan Paul.

von Wright, Georg Henrik (1993). *The Tree of Knowledge and Other Essays*. Leiden: Brill.

Chapter 9: The Origin of Life as a Problem for the Philosophy of Science

The doctrine of spontaneous generation, supported by everyday experience, and belief in special creation, deriving its authority from the *Bible*, were for a long time the main rival "theories" about the origin of life. The thesis of spontaneous generation and its negation were empirically testable only relative to problematic auxiliary assumptions, and therefore they could not be settled by the experimental method. Special creation was supported, besides the story of *Genesis*, also by the design argument: the apparent design and purpose in nature presupposes a Designer. While many 19th century naturalists exercised their imagination to maintain a role for God in the emergence and development of life on Earth, Darwin attacked creationism on scientific, methodological, and theological grounds. Darwin's empiricist and naturalist research program on biological evolution was continued in the 1920s with Oparin's theory of the gradual and lawlike emergence of life through chemical evolution. The recent trend in Christian fundamentalism which calls itself "scientific creationism" has tried to challenge neo-Darwinist thinking, but frequently confuses evolution as a fact, as a theory, as a research program, and as an ideology. Besides using dishonest tricks in their political campaign, these creationists misuse or misunderstand modern philosophy of science, probability theory, and information theory. Their basic claims against the chemical evolution of life are simply new variants of the old – and inconclusive – design argument for the existence of a Great Designer.

Introduction: The Philosophy of Biology

Biology has been one of the most rapidly progressing fields of science in the late 20th century. At the same time, philosophical problems associated with the scientific study of life have received more and more serious attention. With the appearance of the first systematic textbooks, *The Philosophy of Biology* (1973) by Michael Ruse and *Philosophy of Biological Science* (1974) by David Hull, the philosophy of biology has been recognized as an important subdiscipline of the philosophy of science – comparable to specialties like the philosophy of physics, the philosophy of medicine, and the philosophy of the social sciences.

Philosophy of biology is not a branch of what used to be called "philosophy of nature" – an a priori attempt to grasp the essential nature of life through speculative thinking. It is a supplement – rather than a rival – to biological science: it applies the methods of philosophical analysis and argumentation to a critical study of the basic concepts and theories of biology.[1] Examples of relevant concepts include life, organism, species, function, gene, selection, adaptation, and fitness. Relevant theories and approaches include Darwin's theory of evolution,[2] genetics, teleology, and reductionism. The philosophy of biology has also an important role in the evaluation of the relevance of current biological theories to rival world views. This function of philosophy has been vividly illustrated by recent controversies about human sociobiology, the scientific status of Darwinism, vitalism, and creationism. Philosophical analysis is also indispensable for an adequate treatment of ethical problems associated with modern gene technology. The problem of the origin of life, with its fascinating history and systematic connections to the main issues in the philosophy of biology, is the focus of this chapter.

---

[1] For the philosophy of biology, see Nagel (1961), Ch. 11-12, Ruse (1973, 1982), Hull (1974), Rosen (1991), Sober (1993), and Hull & Ruse (2007).

[2] For the Darwinian revolution in biology, see Manier (19789, Gillespie (1979), and Ruse (1982).

Everyday Experience and Divine Revelation

Before the scientific study of life emerged, there were only two ways of approaching the puzzling questions concerning the existence and origin of life on Earth. Why is it the case that there are living entities in the world? What is the purpose of life and death? How did it all begin?

One way of answering these questions relies on *everyday experience*. In many cases, we observe living things to generate from other living things: plants grow from seeds developed by other plants, chicken are born from eggs laid by hens, human embryos grow in the womb of their mother. However, sometimes living entities make a sudden appearance in places where no generating mechanisms are observable: eels and frogs seem to arise from mud, insects and worms from decaying flesh, and fungi from earth. The doctrine of *generatio spontanea* takes it to be a common observation that new life may be generated at any time from matter through natural processes.

Everyday experience is obviously insufficient to tell anything about the emergence of the first living beings in the world. Various imaginative stories about the origin of the world and life are included in the prehistoric *myths*, preserved in the oral tradition of folklore and the holy scripts of religions. In some of these myths, the world and the living things develop through strange and complicated processes out of some original element (in many cases the sea), while in others they are created by the miraculous intervention of a supernatural being or God. Religious myths also typically contain claims about the ultimate aims of creation, and hence about the purpose of life. For example, in the Finnis epic *Kalevala*, the Virgin of Air descends to the sea, a bird[3] lays eggs and one iron egg on her knee, these eggs break in pieces and form the Earth, Heaven, Sun, Moon, and clouds. Then the Mother of Water gives birth to a man, Väinämöinen, who later finds another man to plant all the trees on the ground. In the Jewish *Old Testament*, God creates grass, herb, and fruit trees on the third day, fishes and birds on the fifth day, cattle, creeping thing, beasts on the earth, man and woman on the sixth day (*Genesis* 1).

---

[3] The Finnish name of this bird (*sotka*) corresponds to the genus *Authya* (tufted duck and pochard). The first poem of *Kalevala* never tells where and how the first *sotka* was born.

A literal interpretation of religious myths has often been defended by the assumption that they are based upon *divine revelation*. In the Christian tradition, the doctrine of *special creation* claimed that God, by exercising his divine powers, created the first pair of each species of plants and animals. According to Archbishop James Ussher's widely advertised estimate, published in 1650, the moment of the creation was in the year 4004 BC, Dr. John Lightfoot from Cambridge claimed further that the creation of man took place on the 23rd of October at 9 o´clock a.m.

Spontaneous generation, supported by common sense experience, and special creation, deriving its divine authority from the *Bible*, were for a long time the main rival "theories" about the origin of life. In their original forms they were clearly incompatible: the former claimed that life has emerged several times without the intervention by any agent. However, in their attempts to cope with the results of biological science, both of these theories went through various transformations and modifications which also came closer to each other. For example, it was suggested that spontaneous generation has occurred only once in the history of the world, or that God still continues to create miraculously new living things on earth.

The Spontaneous Generation Controversy

In his excellent historical study, John Farley (1977, 1) defines the doctrine of spontaneous generation (SG, for short) as the thesis that "some living entities may arise suddenly by chance from matter independently of any parent". If such entities are thought to arise from inorganic matter, we speak of *abiogenesis*; and if from organic matter which itself was alive or derived from a living organism, we speak of *heterogenesis*.

The controversy over SG in biology has an important lesson to teach to all those who still believe that science grows linearly by accumulating conclusively established truths. Various forms of SG have been defended by scientists since Aristotle to our days: the "living entities" which have been claimed to arise spontaneously range from eels, frogs, flies, and maggots to parasitic flukes, tapeworms, microscopical animals, infusorians, blobs of protoplasm, bacteria, viruses, genes, enzymes, and molecules of DNA. Experimental evidence against SG was

presented by Francesco Redi and Marcello Malpighi in the 17th century, Lazzaro Spallanzani in the 18th century, Theodor Schwann and Louis Pasteur in the 19th century.[4] Perhaps the most celebrated biological laws are inconsistent with SG: *Ex ovo omnia* (Every living thing from an egg: William Harvey in 1651); *Omnis cellula e cellula* (Every cell from a cell: Rudolf Virschow in 1858); *Omne vivum e vivo* (Every living thing from a living thing: Louis Pasteur in 1862). Still, the belief in SG flourished in the late 18th century (Joseph Needham, Denis Diderot), in 1800-30 (Jean Baptiste Lamarck, German *Naturphilosophie* represented by Lorenz Oken and Karl Ernst von Baer), in 1860-80 in England (Henry Bastian) and in Germany (Ernst Haeckel), and again in the 1920s (H. J. Muller).

Why is the case that the opinions of biologists have for such a long time fluctuated between the acceptance of SG and its rejection? Several interrelated factors have contributed to this situation. One of them was the practical inseparability of SG from religious, metaphysical, and political debates. Opposition to SG was often based upon purely religious or semireligious (Spallanzani) grounds. In the late eighteenth century and in the nineteenth century, SG was generally associated with "dangerous" views – such as materialism, atheism, and political radicalism. Farley (1977) argues that the Pasteur-Pouchet-Bastian controversy never received a neutral treatment in the politically and religiously conservative atmosphere of the French Third Republic. But the debate about SG did not end with Pasteur's famous experiments in 1862. The principle *Omne vivum e vivo* still raised the question of the ultimate origin of the first living things. This question became urgent also as a result of the appearance in 1859 of Charles Darwin's *The Origin of Species*, since the evolution and transmutation of the species of plants and animals rules out special creation as a scientific hypothesis. Even though Darwin himself did not support SG, many Darwinists – like Haeckel in his monism –

[4] Redi showed that maggots in meat are in fact larvae of flies. Spallanzani demonstrated that no microscopic organisms appear in heated and sealed bottles. Pasteur argued that fermentation and putrefaction were consequences of the activity of bacteria and not vice versa. See Pasteur (1960), Leikola (1985).

regarded abiogenesis as a necessary requirement for the evolutionary biology.

Attempts to develop a naturalist explanation for the emergence of life led finally to a theory which is an alternative to the traditional thesis of SG. It is not satisfactory to claim that the original birth of the first living entities was a unique, miraculous event which is the only exception to the otherwise valid principle of biogenesis. This would make the origin of life almost as mysterious as the doctrine of special creation. The panspermia hypothesis of Svante Arrhenius in 1908, i.e. the travel of living spores to the Earth by asteroids or comets, is likewise unsatisfactory, as it does not solve the riddle of the origin of life[5] – and the assumption that the world and life have existed eternally without a beginning is also beset with serious problems. Alexander Oparin and John Haldane independently suggested in the 1920s that life arose under such conditions (without free oxygen in the atmosphere) which, as a result of the photosynthetic activity of living entities themselves, do not exist anymore. In a possibly too optimistic statement, Farley (1977, 171) says that the general acceptance of Oparin's views in *The Origin of Life* (1938) has "in a very real sense" led to "the final abandonment" of the SG controversy. Oparin denies that life arose at once by an improbable but not impossible combination of molecules, and claims instead that "the simplest living organisms originated gradually by a long evolutionary process".[6]

The status of SG has depended on the historical development of other chemical and biological theories (e.g. theories of sexual generation, embryology, cells, evolution of species, diseases, composition of air, fermentation, colloid chemistry, molecular genetics). Already for this reason, the experimental method did not and could not settle the SG controversy. As the physicist-philosopher Pierre Duhem argued convincingly already in 1906, there are no decisive or crucial tests for theories,

---

[5] It is remarkable that Japanese researchers have recently found important ingredients of life in meteorites: amino acids, sugars, and the four nucleobases of DNA (adenine, guanine, cytosine, thymine). As these meteorites are about 5 billion years old, they may have played a role in the early chemical evolution of life on Earth.

[6] Oparin (1953), 60.

since the derivation of test-implications from theories as a rule requires auxiliary assumptions. The Spallanzani-Needham and Pasteur-Bastian debates are excellent illustrations of this thesis: Is the air in the container spoiled, when the infusions are kept at the boiling point for 45 minutes? At what point of temperature do the bacterial spores die? The determination of such questions by experimental means without at the same time begging the question about SG was not possible.[7]

Farley concludes his book by saying that the present state of the SG controversy "has not resulted from disproof by infallible experimental evidence" (p. 184). Indeed, he argues that SG is a thesis which cannot be falsified but can be proven, while the negation of SG can be falsified but "cannot be proven with absolute certainty" (p. 4). However, the situation is even more complex than Farley thinks. He obviously has in mind the fact that unrestricted existential statements are verifiable but not falsifiable.[8] But SG is in fact an existential-universal statement and its negation ~SG is therefore a universal-existential statement. To see this, write L(x) for 'x is a living entity' and P(x,y) for 'x is a parent of y'. Then, by using ordinary logical notation, SG can be formulated by

$$\exists x(L(x) \,\&\, \sim\exists y(L(y) \,\&\, P(y,x)))$$

which is logically equivalent to

$$\exists x\forall y(L(x) \,\&\, (L(y) \rightarrow \sim P(y,x))),$$

and ~SG by

$$\forall x(L(x) \rightarrow \exists y(L(y) \,\&\, P(y,x))).$$

It follows that SG as well as ~SG are both non-verifiable and non-falsifiable by observational evidence.

---

[7] Farley (1977), 5.

[8] To falsify the unrestricted statement "There are lions" one should be able to prove the generalization "All entities are non-lions", which is in principle impossible.

Darwinism vs. Creationism

Charles Darwin's book on the origin of species described a great number of *facts* about evolution and a *theory* which explained these facts in terms of variation and natural selection. What is more, Darwin was also a founder of a biological research program with a deep commitment to an empiricist and naturalist methodology.

In his wish to employ the best canons of scientific reasoning, Darwin was influenced by the views of John Herschel, William Whewell, Charles Lyell, and Auguste Comte.[9] Of special importance for him was the Newtonian *vera causa* principle: scientific explanation should refer only to such kinds of causes that have a "real existence in nature", i.e. are observable in some circumstances. This principle was also the ground for Lyell's uniformitarianism in geology: geological events should be explained by causes which by their nature and intensity are similar to those observable at present. In applying this idea to biology, Darwin wanted to make a clear separation of theology and science. Neal C. Gillespie (1979) therefore argues that the Darwinian revolution in essence consisted of a step from "creationist" to a "positivist" science.

Darwin deliberately chose not to give in the *Origin* any account of the ultimate origin of life on the Earth. At least publicly he agreed with Thomas Henry Huxley's cautious statement (in 1859) that such issues admit "neither proof nor disproof" and therefore are "no subjects for science". It remained as a task for other biologists to complement Darwinism by developing a naturalist theory of the emergence of life.

On the other hand, Darwin did not conceal his aversion to theologically grounded thinking in science. In his *Autobiography*, Darwin tells that in writing the *Origin* he had a firm belief in "a personal God", and even in his later years of growing agnosticism towards Christianity he had "moments of theistic reflections" about the First Cause of the universe and its laws.[10] Thus, Darwin was not a convinced philosophical or metaphysical materialist; rather he was – more consistently than most of his contemporaries – committed to a new naturalistic ideal of

---

[9] For Darwin's philosophy of science, see Hull (1973), Manier (1978), Gillespie (1979), and Ruse (1982).

[10] See Gillespie (1979), Ch. 8.

scientific methodology. This attitude led him to attack the doctrine of special creation on three different levels: methodological, scientific, and theological. First, references to the "will of the deity" or the "plan of the Creator" do not add anything to our knowledge of natural phenomena. It violates the *vera causa* principle and inhibits the progress of science. Secondly, if special creation were treated as a scientific hypothesis, there would be a number of facts about the descent of plants and animals, their similarities and differences, and their migration and geographical distribution, which serve to disconfirm it empirically.[11] Thirdly, God could not be thought to be responsible for the cruelties and waste seen in nature, or for creating a world which deceives and misleads honest inquiry.[12]

In Darwin's time, there was still a minority group of scientists who advocated special creation in the literal sense. One of them was Louis Agassiz, a leading "catastrophist", who in 1862 defined the study of nature as an attempt "to trace the connection between all created things, to discover, if possible, the plan according to which they have been created, and to search out their relation to the great Author".[13] Against this doctrine of *miraculous creation*, some naturalists – like Lyell, the anatomist Richard Owen, and the geologist Adam Sedgwick – favored a view that Gillespie (1979, 21) calls *nomothetic creationism*: God created new species by manipulating the order of nature through some "unknown lawful process" involving causes known only God himself. A third variant of creationism was presented by Robert Chambers who argued, in his *Vestiges of the Natural History of Creation* (1844), that God established in the beginning a system of laws which then operate to generate new species without God's further intervention. Nomothetic creationism was often combined with a theologically motivated attitude of *nescience* (ignorance) concerning God's modes of operation. Nescience on the origins of life could also be based upon the positivist methodology which regarded attempted solutions to this riddle as illegitimate metaphysical speculations.

---

[11] See Gillespie (1979), Ch. 4.
[12] See Gillespie (1979), Ch. 7.
[13] See Gillespie (1979), 52.

In the cultural climate of the nineteenth century, most natural scientists understood that appeal to divine revelation as support of creationism is circular. However, many of them were convinced by the classical design argument for the existence of God,[14] as developed by William Paley in his *Natural Theology* (1802). The argument takes the existence of the order and harmony in nature, and the perfect adaptation of organisms to their environment, as evidence of God's creation: as a watch requires a watch-maker, the design can be explained only by assuming the existence of a Designer. This idea could be easily applied to nomothetic creationism as well. Combined with the fact of evolution, it leads to a position called *providential evolutionism* by Gillespie (1979, Ch. 5): the process of evolution has a direction which is determined by God's will or plan. Among naturalists, this view was supported by Chambers, Owen, the Duke of Argyll, and St. George Jackson Mivart.[15] Another variant was advocated by Asa Gray who thought that the variations on which natural selection works are designed in a divine way.[16] In contrast to all these attempts to preserve a role for God in the theory evolution, Darwin argued that random variation and natural selection are sufficient to scientifically solve the problem of design – without making any assumption about the end or purpose of evolution.

Chemical Evolution and its Philosophical Interpretations

Farley (1977, 183) summarizes the modern view on the emergence of life as follows:

> Life did not arise by spontaneous generation. That is to say, that a functional living entity, whether that be a mouse, maggot,

---

[14] David Hume in the 1750s and Immanuel Kant in 1781 already argued that the "physico-theological" (or "teleological") argument is insufficient to prove the existence of a Creator. Hume and Kant also rejected the "cosmological" argument – which still had some appeal to Darwin in his momentary speculations about the First Cause.

[15] The idea of the directional evolution of the universe towards a pre-established goal is also a foundation for many philosophical systems of evolutionary metaphysics (Herbert Spencer, Charles Peirce, Henri Bergson, A. N. Whitehead, Teilhard de Chardin).

[16] See Gillespie (1979), Ch. 6.

> bacterium, virus, or 'living molecule', did not make an all-at-one appearance from material with no lifelike qualities. Life emerged slowly as a part of a long developmental process, all stages of which were highly probable at the time they occurred. As such, it becomes meaningless to draw a line through these stages and to call stages below the line nonliving and those above living. It, therefore, also becomes meaningless to speak of a spontaneous generation of life, either today or in the past. Oparin's scheme allows for evolutionary process to be a continuous one, but the very existence of life itself renders it impossible for the early stages to be repeated.

In the 20th century, biologists have gained detailed knowledge about the molecular structure of all living organisms. The most important spark was the discovery of the double helix structure of DNA by Crick and Watson in 1953. Together with mostly microscopic fossil records, it can be used as the basis for hypotheses concerning the chemical evolution of life before the Cambrian explosion 600 million years ago.[17]

The oldest fossils of eukaryotes (i.e. organisms consisting of cells with a nucleus) are about 1500 million years old, and the oldest prokaryotes (i.e. unicellular primitive organisms without a nucleus) at least 3500 million years old. Before these developments, the chemical evolution probably has started with the formation of amino acids from inorganic ingredients (carbon, nitrogen, hydrogen, oxygen), and then continued with the emergence of chains of amino acids and thereafter more complex macromolecules (nucleic acids, proteins) which are able to function as reproducing and self-replicating individuals. The macromolecules are the essential building blocks of the first cell-like organisms, and they (enzymes, DNA, RNA) also govern all chemical processes within a cell by the genetic code.

The Miller-Urey experiment in 1952 showed that amino acids can be formed in water condensation from inorganic matter by powerful energy sparks. The conditions of such "primordial soup" tried to simulate what Oparin and Haldane assumed about the early oxygen-poor atmosphere, but they have been modified in further experiments. About the next steps in the emergence of complex polymers, "the RNA world"

---

[17] See Oparin (1958), summary in Ruse (1982), Ch. 6, Korhonen (1985), Rosen (1991), and Luisi (2006).

hypothesis of Cech and Altman states that the first replicators were RNA molecules. An alternative is the "metabolism first" hypothesis of Morowitz. The lively debate still continues: the "everything first" hypothesis claims that life emerged fully formed in a "chemical big bang".[18] Since the development of gene technology in 1980s, the multi-disciplinary program of *synthetic biology* has successfully produced new systems and forms of life by genetic engineering, but it is questionable whether the manipulation of organic material helps to understand abiogenesis.[19] Another interesting program from the same period is *Artificial Life* (or A-life), which simulates living organisms by computer programs and models.

The scheme of chemical evolution is still open and tentative in many respects, but it has already removed much of the "mystery" concerning the emergence of life. The nature of life can no longer be regarded as a forever insoluble *Welträtsel*, as some nineteenth century scientific agnostics (like Emil du Bois-Reymond) thought. As a research program, Oparin's approach obviously continues Darwin's theory of biological evolution: it claims that life has arisen gradually through a completely lawful process involving only natural causes. For us it is interesting to ask what kind of philosophical interpretation can be given to chemical evolution.

Oparin (1953, 31-33) himself refers to the *Dialectic of Nature* by Friedrich Engels (written in 1883, published in 1925). For a scientist in the Soviet Union it was almost obligatory to cite some classic of Marxism. But in fact Oparin's views fit quite well with some of the basic theses of *dialectical materialism*. This type of materialism is anti-reductionist: even though all organisms are constituted by matter, as systems with a complex organization they have properties and laws which cannot be derived from the basic laws of mechanics. In other words, new qualities and regularities emerge when a new level of increasing complexity is reached.[20] These emergent features are the characters that distinguish living and non-living entities (i.e.

---

[18] See Smith (2020).

[19] See Luisi (2006).

[20] Oparin (1953, 250-251) talks about the gradual evolution of "properties and laws of higher order", attained in "the next and more advanced phase in the organization of matter".

the self-regulation and teleonomy of organism as exhibited by their metabolism, growth, development, reproduction, and heredity).

A classical formulation of *emergent materialism* was given C. D. Broad (1925), who distinguished it from radical (eliminative) and reductive materialism in the mind-body-problem. This view claims that the human mind is an evolutionary product of material nature, not capable to exist without a material basis, but still it has acquired a relatively independent and real status. Among others, this view has been supported by Karl Popper.[21]

Emergent materialism is sometimes expressed by talking about non-additive wholes which are more than their parts. To make this idea more precise, let us follow Nagel (1961) by saying that a property or a regularity P of a material whole w is *emergent* with respect to physics and chemistry if there is no true theory T in physics or chemistry such that the applicability of P to w can be derived by T from knowledge about the parts of w. Similarly, P is *epistemically emergent* at time t if no such theory T is known at time t. Emergent materialism then claims that biological organisms - from the simplest organic molecules to human persons - are material wholes with some truly (not only epistemically) emergent properties or regularities. As this claim refers to true physical theories – which may be unknown to us – emergent materialism is a philosophical (ontological) rather than a scientific thesis. The development of natural science is nevertheless highly relevant to our evaluation of its validity.

Emergent materialism is not the only philosophical interpretation that has been given to the theory of chemical evolution. One of these lines is *mechanistic materialism*: organisms and their parts are mere machines governed by the deterministic laws of mechanics, and the birth of life is attributable to a historically unique, utterly improbable chance event. Jacques Monod's *Chance and Necessity* (1971) argued in favor of Descartes and against Hegel that "the cell is after all a machine" so that the system of life as a product of chance is "a frozen accident" – and therefore was taken by his Marxist opponents to

[21] Cf. Beckermann et al. (1992). See Niiniluoto (1994) for a defense of emergent materialism.

represent mechanistic materialism. Another line attempts to revitalize *vitalism* by arguing that the new qualities of organic wholes have somehow pre-existed in their parts in a hidden form waiting for their realization. In other words, matter – which was able to develop living organisms – was already in some sense alive, and organisms are expressions of "entelechies" (Driesch) or creations of *élan vital* (Bergson).[22]

Both mechanistic materialism and vitalism turn back to the view that life arose by spontaneous generation (abiogenesis or heterogenesis). Our definition of emergence makes superfluous the – in any case untestable – assumption of latent properties activated by vital forces, thereby leading to the rejection of vitalism. The debate between emergent and mechanistic materialism concerns primarily the possibility of reduction and therefore it is a metaphysical or ontological issue.

On the other hand, the question whether life arose "by chance or necessity" is not unambiguous. Some biologists hope eventually to find a how-necessary explanation – rather than a how-possible explanation – of the origin of life on the Earth. Korhonen (1985) argues that the "abiotic synthesis" was "inevitable" under the conditions of the primitive Earth, even though it took 1000 million years to occur, but allows that chance played a role at some stages (e.g. the choice of the 20 amino acids in our genetic code). He concludes that "since replication molecules, their genetic code, protein synthesis and the basic structure of cells are similar in all living organisms, there are good reasons to assume that all these came into being only once, and that all the living things today are descendants of only one replicating molecule and one primitive cell" (p. 86). One the other hand, in biological evolution variation by mutations is a random probabilistic process, and the same may be true of the origin of primitive life – especially if it took place in "a nuclear geyser", as the Japanese team of Ebisuzaki and Mariyama has recently suggested. Thus, emergent materialists may assume that the world is ultimately indeterministic and the chemical evolution of life was governed by probabilistic laws. The values of the relevant probabilities in such models shows then to what extent the same process can be expected to

[22] For a critique of vitalism, with reference to Eino Kaila's anti-reductionist monism, see Niiniluoto (2010).

repeat itself if similar circumstances occur some part of the universe. This would be highly relevant for the so far unsolved question about the possibility of extraterrestrial life.

"Scientific" Creationism

A Gallup survey published in *Christianity Today* in 1982 told that more than 50% of the adults in the USA believe that the Earth and humans have been created within the last 10.000 years and that Adam and Eve were real people. With the trend of secularization, in 2020 the share of the Christians has declined to 65%, while the share of unaffiliated (among them agnostics and atheists) is over 25%. Similar world-wide surveys show that there still are many people who openly adopt a *religious* rather than a scientific *word view*: when science and the *Bible* contradict each other, alleged divine revelation wins. Another attitude, which has been increasing among the Christians since the end of the nineteenth century, is to adopt a metaphorical interpretation of the *Genesis* and thereby to make room for the acceptance of all the results of scientific research. As the question concerning the existence of God cannot be solved by the empirical sciences, this compromise view is able to combine Christian beliefs and scientific knowledge without any explicit contradiction – indeed, it may accept the chemical evolution of life but, with some imagination, interpret these facts in terms of nomothetic creationism and providential evolution. However, in endorsing claims about reality which are neither testable nor even indirectly grounded on scientific inquiry, this position is in conflict with the epistemological and methodological assumptions characteristic to a *scientific world view*.[23] A third alternative, which tries to remove even this epistemological conflict, is to regard religion only as a matter of personal attitude without any cognitive content. According to such *fideism*, religious faith is not in need of any reasons or evidence. In this case, the sentence "God created life on Earth" is

[23] Logical empiricists (Carnap, Ayer) thought that theism and atheism are both meaningless statements. Indeed, it may be doubted whether the notion of omnipotent God is coherent. Some scholars support scientific theism (Richard Swinburne) or scientific atheism (Richard Dawkins). I favor philosophical atheism on the ground that one cannot foresee any empirical or rational evidence for the existence of a supernatural God.

not an assertion with a truth value, but still it may have a "meaning", since it functions in certain religious and ethical language-games and rituals.[24]

Some Christian fundamentalists have launched a counter-attack against science-based secularization and metaphorical interpretation of "holy texts" with the program of "scientific creationism". With a carefully planned aggressive campaign against Darwinism, the representatives of this movement urge that the best scientific account of the origin and development of life coincides with the Biblical story of special creation. At the same time, they claim that neo-Darwinism has become a dogma among scientists which is accepted only upon faith grounded on materialist metaphysics.

As the method of science is non-dogmatic and self-corrective, no scientific theory is immune to criticism. In particular, specific theories of the chemical and biological evolution of life are tentative and may eventually be replaced by improved theories. Therefore, the creationists certainly have the right to present critical questions about these theories. But unfortunately their style of argumentation is not free from dishonest tricks. Michael Ruse (1982, 303) – who served as a witness in the court which decided in 1982 that the Arkansas law for giving equal time in biology classes for "creation-science" and "evolution-science" violates the constitution of the USA – concludes that scientific creationism is "a grotesque parody of human thought, and a downright misuse of human intelligence".

One of the creationist strategies is to present neo-Darwinism and special creation as the only possible alternatives – which is false, as we have seen – and then to interpret all the anomalies or open questions of modern biology as supporting their own position. At best such difficulties – which every scientific theory ever proposed has had – justify the need to search for better scientific theories which correct the preceding theories but also preserve their main virtues.

Another trick of the creationists is to misuse modern philosophy of science for their own purposes. They keep reporting

---

[24] Fideism goes back to Tertullian's dictum: *credo quia absurdum est*. Its different forms have been advocated by Søren Kierkegaard (faith as an irrational leap), William James (the will to believe), and some followers of Ludwig Wittgenstein.

Karl Popper's thesis that the theory of evolution is a unfalsifiable tautology[25] and yet claim that this theory has been falsified by the fossil records. They rely on the so-called new philosophy of science (Thomas Kuhn, Imre Lakatos, and others) in claiming that all science is based upon dogmatic or metaphysical background assumptions, but disregard the requirement that a scientific research program should be theoretically and empirically progressive. What new explanatory theories have the creationists proposed? Where are the successful empirical predictions of the creationist program?[26]

Larry Laudan (1982) has suggested that, as the demarcation criteria for distinguishing science from non-science are controversial, it would be reasonable to treat creationism as a science – and then show how bad it is in this respect. This would in fact correspond to one of the strategies that Darwin used in the *Origin*. Ruse thinks otherwise: even if the borderline were vague, there may still be clear examples of pseudoscience – and surely creationism is a case in point.[27] And this a good reason for keeping it outside the biology classes.[28]

One further strategy is the deliberate confusion between evolution as a fact, as an explanatory scientific theory, and as an ideology. When the members of the Creation Research Society claim that Christian children should be "protected" from having to read in school such "atheist teachings" as the theory of evolution, they fail to appreciate the fact, in explaining the evolution of life through natural causes, neo-Darwinism does not make any assertion about the existence or non-existence of God. The same mistake is made, when Darwinism is taken to claim that "only chance really exists" or that "there exists

---

[25] Popper's claim, which he later regretted, has been criticized by Ruse (1982), 140-142, and Kitcher (1982).

[26] Sober (1993), Ch. 2, evaluates scientific creationism from the viewpoint of general philosophy of science.

[27] In spite of its pretentious name, "scientology" classifies itself as "a religion in its highest meaning". For the demarcation problem, see Niiniluoto (1984).

[28] In Finland the controversy between creationism and Darwinism in schools in not so pressing as in the United States: teaching of the *Bible* takes place in the classes of religion, whereas students without religious denomination have since 1985 studied life stance education, so that biology is taught purely on the scientific basis.

neither God of love or morality, nor a God of amorality".[29] And to intimidate the readers with the dangers of Darwinism by asserting that Hitler was a "social Darwinist"[30] is itself an act of questionable moral value.[31]

Variations of Old Themes

A. E. Wilder-Smith book *The Natural Sciences Know Nothing of Evolution* (1981) is an illustrative example of literature on the origin of life, based upon creationism or theory of Intelligent Design. The author is advertised as an "internationally renowned scientist"; he avoids mentioning scientific creationism at all,[32] and usually replaces – without any conceptual analysis – the word "God" by *ersatz* phrases like "teleonomy", "planning", "idea", "logos", "know-how", "intelligence", "information", "concept", "expertise", "pattern recognition", "programming", "coding", or "something outside time and matter". However, the book is published by a division of CLP (Creation-Life Publishing) in San Diego. Wilder-Smith presents a number of specific objections to the views of Darwin, Oparin, Eigen, and Monod – and it is a task for chemists and biologists to evaluate them. Here I comment only on two important features of his argumentation. Similar objections can be presented against the later "ID-theorists" who try to hide their commitment to creationism.[33]

First, Wilder-Smith seems to think that his opponents believe in spontaneous generation. He follows G. A. Kerkut by formulating the first assumption of the theory of evolution as follows: "nonliving matter spontaneously produced living matter at biogenesis" (p. 149), and brings Pasteur and canning

---

[29] Wilder-Smith (1981), 5.

[30] Wilder-Smith (1981), 148.

[31] For the misnamed "social Darwinism" and evolutionary ethics, see Ruse (1982), Ch. 12, and Kitcher (1982).

[32] Wilder-Smith (1915-1995) had PhDs in organic chemistry, pharmacology, and technology. When he visited Finland in 1981, he refrained from using the term "creationism" (in spite of his support of young Earth creationism) and claimed to follow strictly "science and philosophy of science".

[33] Other scientific creationists and ID-theorists (Henry Morris, Michael Behe, Philip Johnston, and William Dembski) are criticized in Ruse (1982), Kitcher (1982), and Sarkar (2007).

factories as evidence against this theory (p. ix). Some creationists have gone so far in this respect that they have tried to calculate the probability of the momentary spontaneous arrangement of the 25.000 enzymes in a human body – and, to no one's surprise, have ended with extremely small numbers.[34] It has also been argued that even the birth of one protein with 100 amino acids would be statistically "impossible" or too improbable to occur, since from 20 amino acids one can build 20^{100} or circa 10^{130} such proteins. Further, mathematicians are claimed to regard as "impossible" events with a probability less than 1 in 10^{50}.[35] Such arguments prove nothing: draw randomly with replacement 80 cards from an ordinary deck of 52 cards, and the probability of the result that you have just obtained is less than 1 to 10^{130}! In this respect, William Paley in *A View of the Evidence of Christianity* (1794) was wiser than present creationists: in discussing Hume's attack on miracles, he in effect pointed out that improbability does not imply impossibility.[36]

Secondly, Wilder-Smith argues that the genesis of the genetic code and the increase of genetic information is "untenable according to the tenets of modern information theory" (p. x). The crux of his argument is the following:

> Evolution is thus basically an attempt to explain the origin of life from matter and energy without the aid of know-how, concept, teleonomy, or exogenous (extramaterial) information. It represents an attempt to explain the formation of the genetic code from the chemical components of DNA without the aid of a genetic concept (information) originating outside the molecules of the chromosomes. This is comparable to the assumption that a text of a book originates from the paper molecules on which the sentences appear, and not from any *external* source of information (external, that is, to the paper molecules). (p. 4)

However, the whole of this reasoning is ruined by the ambiguity of the concept of information (cf. Chapter 3 above). Wilder-

---

[34] Cf. Hitching (1982), 70-71.

[35] Cf. Hitching (1982), 71, 119,

[36] See Manier (1978), 70-71. Dembski's *The Design Inference* in 1998 repeats the same mistake about events of low probability. See Sarkar (2007), 50-57, 127.

Smith is well aware that human languages are based on conventions about the meaning of their symbols. Only through such semantic conventions strings of letters possess some information. This kind of information was called *semantic information* by Rudolf Carnap and Yehoshua Bar-Hillel who developed its theory in the early 1950s. It is clear, moreover, that the existence of semantic information presupposes human minds: language as a semiotic system of symbols is a human artefact (cf. Chapter 4 above), and its existence is as human-dependent as the existence of cars or the play *Hamlet*.[37] So far so good. The problems start with Wilder-Smith's misinformed assumption that Claude Shannon's information theory deals with semantic information – or "teleonomic information", as he calls it (p. 140). This was explicitly denied by Shannon in 1948, and the confusions between semantic information and Shannon's statistical information were cleared out by Bar-Hillel (1964) soon afterwards. Shannon's concept of information has nothing to do with the content of a "message", but depends only on the stable frequency in which it as an input in a communication channel.

Wilder-Smith further assumes that information in biology means Shannon's information,[38] and concludes that genetic information is semantic. This is clearly false. A DNA molecule is able to guide, through a messenger RNA molecule, the formation of new proteins – so that an ordered sequence of three bases always serves as a "code" for an amino acid. This genetic code is not a symbolic convention, but an expression of complex causal regularities which exist in nature without any

---

[37] These are Wilder-Smith's own favorite examples. The question whether meaningful words could be generated by chance was discussed in probability theory by P. S. Laplace in the early 19th century, and it goes back at least to Cicero.

[38] This assumption is also problematic – in spite of the fact that one may apply some concepts of Shannon's communication theory and Wiener's cybernetics to biological organisms. In biology "information" often means the same as the *physical complexity* of an organism, and a formal algorithmic measure of such complexity has been developed by Kolmogorov. ID-theorists like Michael Behe have attempted to argue that complexity cannot be increased by evolution and natural selection (see Sarkar, 2007, Ch. 6). For mistakes in Dembski's related argument in terms of "the law of conservation of information", see Sarkar (2007), Ch. 7.

human activity. No one has made an agreement that e.g. the combination guanine-cytosine-adenine corresponds to alanine. This is a regularity in nature – not at all comparable to the stipulation that the string of three points correspond to the letter "s" in the Morse code.

Wilder-Smith has been misled by the fact the so-called genetic code is not a system of conventional symbols (like natural languages) but rather a system of causal signals which govern "control and communication" in organisms. What is remarkable about this code is that its regularities are emergent in the sense that they cannot be derived from basic physical and chemical laws – so that the code is generated in the emergent evolution of life. Thus, all that is left of Wilder-Smith's attack on neo-Darwinism is an old Paley-type design argument: the complex structure of living entities requires a Designer. This argument is now as weak as ever.

Conclusion

Modern molecular biologists have been able to develop a detailed theory of the structure or organisms and a promising theory of the emergence of life on earth. For this purpose, they have not needed assumptions of vital forces or supernatural beings. On the other hand, their vitalist and creationist opponents have neither made theoretical or empirical progress nor invented any essentially novel philosophical points against the naturalist research program of biology.

Science does not assign any purpose to the phenomenon of life in the universe. In particular, no theological or philosophical account of the meaning of life can be derived from the results of scientific inquiry. However, this does mean that, according to the scientific world view, life is meaningless. On the contrary, as persons with self-consciousness, all human beings have the ability to find a purpose to their own lives – by reflecting on the conditions their rights and duties, welfare and happiness (cf. Chapter 18 above). As members of society, they have the prerogative to try to make the life of their fellow humans worth living. This view thus serves to emphasize humanism: the ultimate responsibility for the way in which the torch of life is carried forward belongs to human beings themselves.

Note. This chapter is based on my Studia Generalia lecture in the 25th anniversary of the University of Oulu, May 10, 1983, published in Sakari Piha (ed.), *The Origins and Purpose of Life*. Oulu: Acta Universitatis Ouluensis, 1985, 163-177.

Literature

Bar-Hillel, Yehoshua (1964). *Language and Information: Selected Essays on Their Theory and Application*. Reading, MA: Reading.

Beckermann, Ansgar, Flohr, Hans & Kim, Jaegwon (eds.) (1992). *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin: De Gruyter.

Broad, C. D. (1925). *The Mind and Its Place in Nature*. London: Routledge & Kegan Paul.

Dawkins, Richard (1986). *The Blind Watchmaker*. New York: Norton & Co.

Farley, John (1977). *The Spontaneous Generation Controversy from Descartes to Oparin*. Baltimore: The Johns Hopkins University.

Gillespie, Neal C. (1979). *Charles Darwin and the Problem of Creation*. Chicago: The University of Chicago Press.

Hitching, Francis (1982). *The Neck of the Giraffe or Where Darwin Went Wrong*. London: Pan Books.

Hull, David (1973). *Philosophy of Biological Science*. Englewood Cliffs, NJ: Prentice-Hall.

Hull, David & Ruse, Michael (eds.) (2007). *The Cambridge Companion of the Philosophy of Biology*. Cambridge: Cambridge University Press.

Kitcher, Philip (1982). *Abusing Science: The Case Against Creationism*. Cambridge, MA: The MIT Press.

Korhonen, L. Kalevi (1985). "The Language of Life". In Sakari Piha (ed.), *The Origins and Purpose of Life*. Oulu: Acta Universitatis Ouluensis, 1985, 61-88.

Laudan, Larry (1982). "Science at the Bar. Causes for Concern", *Science, Technology and Human Values* 41, 16-19.

Leikola, Anto (1985). "The Problem of the Origin of Life in the History of Science". In Sakari Piha (ed.), *The Origins and Purpose of Life*. Oulu: Acta Universitatis Ouluensis, 1985, 51-162.

Luisi, Pier Luigi (2006). *The Emergence of Life: From Chemical Origins to Synthetic Biology*. Cambridge: Cambridge University Press.

Manier, Edward (1978). *The Young Darwin and His Cultural Circle*. Dordrecht: D. Reidel.

Martin, Michael (ed.) (2007). *The Cambridge Companion to Atheism*. Cambridge: Cambridge University Press.

Monod, Jacques (1971). *Chance and Necessity*. New York: Alfred A. Knopf.
Nagel, Ernst (1961). *The Structure of Science*. London: Routledge and Kegan Paul.
Niiniluoto, Ilkka (1984). "The Nature of Science". In *Is Science Progressive?* Dordrecht: D. Reidel, 1-9.
Niiniluoto, Ilkka (1994). "Scientific Realism and the Problem of Consciousness". In Antti Revonsuo & Matti Kamppinen (eds.), *Consciousness in Philosophy and Cognitive Neuroscience*. Hillsdale, NJ: Lawrence Erlbaum Associates, 33-54.
Niiniluoto, Ilkka (2010). "Kaila's Critique of Vitalism". In Juha Manninen & Friedrich Stadler (eds.), *The Vienna Circle in the Nordic Countries*. Dordrecht: Springer, 125-134.
Oparin, A. I. (1938/1953). *The Origin of Life*. New York: Dover.
Pasteur, Louis (1960). "On the History and Theories of Spontaneous Generation". In L. Clendening (ed.), *Source Book of Medical History*. New York: Dover, 379-388.
Rosen, Robert (1991). *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. New York: Columbia University Press.
Ruse, Michael (1973). *The Philosophy of Biology*. London: Hutchinson.
Ruse, Michael (1982). *Darwinism Defended: A Guide to the Evolution Controversies*. Reading: Addison-Wesley.
Sarkar, Sahotra (2007). *Doubting Darwin? Creationist Designs on Evolution*. Oxford: Blackwell.
Smith, Michael Marshall (2020). "A Radical New Theory Rewrites the Story of How Life on Earth Began", *New Scientist*, August 5, 2020.
Sober, Elliott (1993). *Philosophy of Biology*. Boulder: Westview Press.
Wilder-Smith, A. E. (1981). *The Natural Sciences Know Nothing of Evolution*. San Diego: Master Books.

Chapter 10: The Human Sciences are Useful in Many Ways

After a decade of successful work in the humanities, social sciences, theology, law, and behavioral sciences, the 10th Anniversary Symposium of the Helsinki Collegium for Advanced Studies (HCAS) invited us to reflect upon the role of the human sciences in contemporary society by raising the question "What's the Use?".

The Value of Basic Research

The question "What's the use?" reminds me of a statement which I cited in my Finnish textbook on the philosophy of science in 1980.[1] John Herschel, British astronomer and philosopher, in his work *Preliminary Discourse on the Study of Natural Philosophy* (1831), made the following remark:

> The question, *Cui bono?* to what practical end and advantages do your researches tend? - is one which the speculative philosopher, who loves knowledge for its own sake, and enjoys, as a rational being should enjoy, the mere contemplation of harmonious and mutually dependent truths, can seldom hear without a sense of humiliation.

But he added:

> But if he can bring himself to descend from this high but fair ground, and justify himself, his pursuits, and his pleasures in the eyes of those around him, he has only to point to the history of all science, where speculations, apparently the most unprofitable,

[1] Niiniluoto (1980), 76.

> have almost invariably been those from which the greatest practical applications have emanated.

With remarkable foresight, Herschel was here defending *basic* or *fundamental research* during a time when there was not yet systematic discussion of applied research and innovation chains:[2] pursuit and contemplation of true scientific theories has intrinsic value as "knowledge for its own sake", but almost invariably it is also the route to most profitable practical applications.[3]

The same argumentation strategy can be applied in the case of the human sciences: knowledge about individual and social actions, their history and their cultural results, has intrinsic value. All scholars who have given contributions in these fields with their studies - and likewise their supporting institutions, such as the Helsinki Collegium for Advanced Studies and the University of Helsinki - can be proud of their achievements.

The intrinsic value of the humanities is related to their crucial role in the academic life of a university - *universitas* as a forum for multidisciplinary interactions. Wilhelm von Humboldt's idea of *Bildung durch Wissenschaft*, which covers sciences and scholarly studies in the broad sense, is today realized in research-based education (cf. Chapter 13 below). Thereby progress in basic research leads to improvements in the education of independently thinking critical university students.

Humanities: How Useful?

In Herschel's spirit we may "descend" to consider the question *Ciu bono?* by pointing out that the human sciences have been

---

[2] For the conceptual distinction between basic and applied research, and between science and technology, see Chapter 11 below.

[3] A similar slogan was put forward by the Austrian physicist Ludwig Boltzmann in 1890: "Theory is the most practical thing conceivable" (see Boltzmann, 1974, 35). In Finland, the same argument about beautiful mathematical theories and their potential applications was made by Rolf Nevanlinna.

"useful" in many ways. In Finland this is indeed most appropriate, since we know that the great humanist scholars in the University of Helsinki in the mid-nineteenth century helped to establish Finland as a nation state with its own culture, history, language, folklore, literature, music, education system, and legislation. In the next century, the jurists and social scientists helped to build the independent Finland as a democratic welfare state with rule of law. Being "useful" is a modest understatement to describe the enormous value of these developments.

The American philosopher Martha Nussbaum, Honorary Fellow of the Helsinki Collegium, in her recent book *Not for Profit: Why Democracy Needs the Humanities* (2010), gives a convincing argument to show why the humanistic education of critical and emphatic citizens is significant for democracy.

Other illustrations of the surprising practical dimensions of humanist research can be mentioned. In the 1840s the linguist and explorer Georg August Wallin made pioneering travels in Arabia. After his short period as Professor of Oriental Languages in Helsinki, the Arabic studies emerged as a flourishing branch of cultural research. No one could guess at that time that the Arabic countries will become the leading producers of oil, so that knowledge about their languages and cultures will be politically and economically of crucial importance.

In the first decades of the twentieth century, the Finnish linguist Gustav John Ramstedt made several expeditions to Siberia and Mongolia. As a side result, he wrote the first grammar of the Korean language. In the early 1950s with the outbreak of the Korean war, this grammar was the only source in English for this still unknown language, and a great number its copies were distributed to the U.S. army.

An influential attempt to analyze the utility of different kinds of research was given by the German philosopher Jürgen Habermas in his *Erkenntnis und Intresse* in 1968.[4] Habermas argued that the natural sciences are governed by the "technical" interest of controlling nature, the humanities with the

[4] See Habermas (1971).

"hermeneutic" interest of communication and improved self-understanding, and the critical social sciences with the "emancipatory" interest of liberating us from suppressing ideologies. In my old textbook mentioned above, I complained that here Habermas's approach is too instrumentalist, as he ignores the theoretical interest of knowledge for its own sake, or the "epistemic utility" of truth and information: basic research in the natural and human sciences has no place in his framework.

The notion of "interest" does not here refer to the motivations of individual researchers, but rather to the institutional standards for appraising the success of various branches of inquiry. In this sense, the three interests are certainly relevant for discussing the rationality of useful science. For example, historical human studies investigate the past development of culture, and thereby give valuable information about the origin of our ways of life, customs, languages, ideas, morality, religion, science, art, education, communication, economy, and social structures – and thereby promote the interpretation and transmission of cultural traditions as well as improve tolerance and reduce misunderstanding among people. Systematic human studies (e.g. anthropology, general linguistics, philosophy, theology, aesthetics, psychology, sociology, political science, and economics) investigate the nature of human beings in different cultural environments and the results of their intentional actions. Both types of inquiry provide hermeneutic understanding about the temporal aspects of human culture – helping us to know who we really are and what we potentially could be.

However, the dichotomies of Habermas are too rigid. Natural science can give results which are hermeneutic (e.g. paleontology and the theory of evolution tell significant facts about human origins) or emancipatory (e.g. research on environmental protection and climate change can change our valuations). Philosophy (both epistemology and ethics) can promote emancipation by teaching critical thinking, which can be practiced in the diagnosis of our times and the critique of social injustice (cf. Chapter 1 above). There are also many applied human and social sciences which serve the technical interest of influencing

human behavior (e.g. nursing science, heath care studies, applied psychology, pedagogic, language technology, administration, social policy studies, social work, media studies, business economics, urban studies, development studies, futures studies).[5] In treating situations and systems, which involve material and human components, they can be in fruitful co-operation with natural, engineering, and medical sciences. Their results may be important for our everyday activities and for political decision-makers who are struggling with "wicked problems" and "grand challenges" (e.g. climate warming, loss of biodiversity, crisis of democracy, poverty, inequality, ageing human populations). Further, commercialization of the results of the human sciences is today an increasingly important affair, and we hear a lot of talk about "content production", "culture industry", and "social innovations".

In my joint empirical survey with the medieval historian Tuomas Heikkilä, we found that 77% of the Finnish citizens and even 98% of political and economic decision makers think that human studies are useful in society.[6] So I conclude that we humanists should not be too shy in answering the question *Cui bono?*: we are good, important and useful in many ways!

Note. This chapter is based on a talk at the 10th Anniversary Symposium of The Helsinki Collegium for Advanced Studies, October 20, 2011. It was published in *Annual Report 2011-2012*, Helsinki Collegium for Advances Studies, Helsinki, 2012, 25-27.
www.helsinki.fi/collegium/english/about_the_collegium/Annual Report_2011-2012.pdf.

Literature

Boltzmann, Ludwig (1974). *Theoretical Physics and Philosophical Problems*. Dordrecht: D. Reidel.

[5] See Niiniluoto (2013).
[6] See Heikkilä and Niiniluoto (2017).

Habermas, Jürgen (1971). *Knowledge and Human Interests*. Boston: Beacon Press.

Heikkilä, Tuomas & Niiniluoto, Ilkka (2017). "The Value and Role of Humanities Research: Review of the Contemporary Status and Suggestions for Future", *Nordicom Information: Medie- och kommunikationsforskning i Norden* 39:1, 80-86.

Herschel, John (1831). *A Preliminary Discourse on the Study of Natural Philosophy*. London: Longman, Rees, Orme, Brown & Green.

Niiniluoto, Ilkka (1980). *Johdatus tieteenfilosofiaan: käsitteen ja teorianmuodostus*. Helsinki: Otava.

Niiniluoto, Ilkka (2013). "On the Philosophy of Applied Social Sciences". In Hanne Andersen, Dennis Dieks, Wenceslao Gonzales, Thomas Uebel & Gregory Wheeler (eds.), *New Challenges to Philosophy of Science*. Dordrecht: Springer, 265-274.

Nussbaum, Martha (2010). *Not for Profit: Why Democracy Needs the Humanities.* Princeton: Princeton University Press.

Chapter 11: Values in Design Sciences

Following Herbert Simon's idea of "the sciences of the artificial", one may contrast descriptive sciences and design sciences: the former are concerned with "how things *are*", the latter tell us "how things *ought to be* in order to attain goals, and to function". Typical results of design sciences are thus expressions about means – ends relations or technical norms in G. H. von Wright's sense. Theorizing and modeling are important methods of giving a value-free epistemic justification for such technical norms. The values of design sciences are not criteria for the acceptance of theories or models, but rather antecedents of conditional recommendations of actions. Design sciences are thus value-neutral and value-laden at the same time.

Design and Design Research

The concept of *design* is traditionally associated with the late 19th century movement of "applied arts" or the industrial production of artifacts intended for use as consumption commodities in everyday life. Design is the stage of shaping the desired object (e.g. a cup, a textile, a chair, or an interior) and conceiving the methods, tools, and procedures for its construction. Thus, industrial design is a process whose products are prototypes of concrete artifacts. The profession of designers and their systematic education was developed in Bauhaus in 1919-33 and Die Hochschule für Gestaltung in Ulm in 1955-68.[1] In Finland, the professional School of Industrial Arts was upgraded in 1973 to the University of Industrial Art Helsinki (later the University of Art and Design, today a faculty of the Aalto-University), and its doctoral programs were started in 1981.

---

[1] See Maldonado and Bonsiepe (1964).

By *design research* one could mean any scholarly study of design processes and products by such disciples as philosophy, aesthetics, semiotics, history, psychology, education, sociology, and economics. Another proposal is that design research could be a new kind of applied science, a "science of the artificial" in the sense of Herbert Simon (1969).[2] To motivate this proposal, it is useful to introduce *design in the broad sense* as the process of creating all kinds of artifacts or shaping and planning human-made objects and systems. While in the narrow sense design is identified with industrial design, in the broad sense it includes the fine arts (music, painting, sculpture), arts and crafts (preparing unique useful decorative goods by the hands of artisans), engineering (creating efficient material tools and machines), architecture (planning and creating houses and built environment), and economical and social planning (shaping social institutions). The sciences of the artificial are then studies which attempt to show how design in its different broad variations can best achieve its goals.

Design in this broad sense has essentially the same meaning as *techne* in Greek and *ars* in Latin.[3] For Aristotle, theoretical philosophy is concerned with *episteme* (knowledge, *scientia* in Latin) and practical philosophy with *phronesis* (deliberation), but "productive arts" involve *techne* in the sense of "stable habit of making", the tools and skills of such activity, and its products or "artifacts". Even symbolic creation like poetry was included in the domain of *techne*. While the Renaissance genius Leonardo da Vinci was still an engineer, architect, painter, and sculptor in one person, the range of technics was gradually split into separate directions. The distinction between the mechanical art of engineers (studied by *technology*) and the fine arts (studied "without interest" by *aesthetics*) was established only in the late 18th century by Immanuel Kant. The tradition of applied arts wished to reunite these two types of *techne* again.

The broad notion of design is in fashion today. As the *World Design Capital 2012*, the City of Helsinki advocated a value-laden notion of "embedding design in life", or user-friendly design with a global responsibility. This trend is also

[2] See Niiniluoto (1984).
[3] See Mitcham (1994).

witnessed also by the establishment of Aalto University in 2010 by merging the Helsinki University of Technology, Helsinki School of Business Economics, and the University of Art and Design. The unifying idea of this institution, which covers most areas of design in the broad sense, is that an engineer discovers a new tool, a designer shapes it as a product, and the business people sell it in the market as an innovation.[4] The University of Helsinki, a traditional research university, added to this idea of research-based design by participating in the WDC year with the slogan "designing society through thinking".

Science, Technology, and Utilities

The influential Frascati Handbook of the OECD (1962) consolidated terminology which became standard in science policy. The OECD distinction between *research* ("the pursuit of new knowledge") and *development* (the use of results of research "to develop new products, methods, and means of production") repeats the old Aristotelian distinction between *scientia* and *techne*. At least for a scientific realist, this R & D divide matches with the distinction between science and technology: development is science-based technology.[5] Pragmatists and instrumentalists, who see science as a problem-solving rather than a truth-seeking activity, instead tend to blur the difference between R and D.

The OECD Handbook makes a further distinction between *basic* and *applied* research. The former (fundamental, curiosity-driven, blue skies research) seeks knowledge for its own sake "without the aim of specific application", the latter pursues "knowledge with the aim of obtaining a specific goal". The historical background of this divide can be found in Francis Bacon's early 17th century program of the scientification of technology: "knowledge is power" in the sense that knowledge about causal laws helps us to produce desired things or to prevent undesired outcomes. The idea of systematically searching such practical knowledge was successfully realized – historically parallel to applied arts – in the late 19th century by the emerging engineering sciences and agricultural

[4] The name of this university is taken from Alvar Aalto, the Finnish architect and designer.

[5] See Niiniluoto (2016).

sciences. Thereby professional schools of technology (such as École Polytechnique established in Paris already in 1794) became Universities of Technology, working as new centers of applied research, while development was practiced in industrial laboratories.

The linear model of innovation assumed that there is a temporal process from basic research to applied research to development.[6] Recently this model of "innovation chain" has been replaced by a richer institutional description of innovation cycles or networks with interactive loops and feedbacks. But, as we shall see, it is still possible to make conceptual distinctions between R and D and between basic and applied research.

One way of looking at these conceptual distinctions is in terms of the related value goals. Pure science and applied science both seek knowledge, but for the former justified truths have *intrinsic value* while the latter is interested in their *instrumental value*. Similarly, for the fine arts beauty is an intrinsic value, while technology is interested in the instrumental value of mechanical artifacts.

A related distinction can be made between epistemic and practical utilities. *Epistemic utilities* are standards for assessing the quality of success in knowledge-seeking, such as truth, information, truthlikeness, confirmation, understanding, explanatory power, predictive power, and simplicity.[7] As Isaac Levi (1967) convincingly argued in his cognitive decision theory against Richard Rudner (1953), the tentative acceptance of scientific hypotheses can be based upon their epistemic utilities – without appealing to other value judgments. Levi's cognitivism is in harmony with the principle of *objective value neutrality*: the arguments for the acceptance or rejection of hypotheses are not allowed to appeal to such (assumed) facts that the truth or falsity of the hypothesis would benefit or harm us for political, religious, ethical, or economic reasons. If it happens that this norm is violated, as some sociologists of science have tried to illustrate by historical case studies, the conclusions or reasons have to be revised or corrected.

On the other hand, various kinds of *practical utilities* are relevant and appropriate in technology and design – and more

[6] See Godin (2006).
[7] Niiniluoto (2007, 2012).

generally in human decision-making in politics, economy, and everyday life. Each artifact has its intended use (or "final cause" in Aristotle's sense), and possibly some unintended side-effects. For example, the Roman Vitruvius in his classical *De Architectura* (c. 25 BCE) defined the basic values of architecture as *firmitas* (solidity), *utilitas* (usefulness), and *venustas* (beauty) (Vitruvius, 1914). Modern functionalist architects and designers have emphasized the principle "form follows function". Different branches of technology have their own specific goals, such as power, speed, and security for airplane engineering. Such values can be used as goals to be pursued by means of design methods. In *Technology Assessment* (TA), they serve as criteria of the success of technical tools and products. As a summary of such assessment, one can propose the formula TA = 6E + S (cf. Chapter 20 below). Here the six E's are

- *Effectiveness* (ability to achieve the intended use or function)
- *Economic* efficiency (cost-benefit success)
- *Ergonomics* (relations to the health of users)
- *Ecology* (relations to the health of the natural environment)
- *Esthetics* (beauty)
- *Ethics* (good or bad by moral standards).

Finally, S refers to the *social* impact of technical tools.

The criteria of this formula can be weighted in different ways in the various domains of design: esthetic values are dominant in the fine arts and crafts, effectiveness and economy in engineering. The ecological demand of sustainable development is today crucially important in all areas, and technology is inherently value-laden in the sense that the new possibilities opened by new tools can always be assessed from the ethical perspective.

The decision-theoretical framework is sufficiently flexible in its ability to handle situations where both epistemic and practical utilities are involved and interact. This can happen in cases, where some model is used as a basis of action or policy. An example is given by Helen Longino (1986).[8] Linear and

[8] Cf. Niiniluoto (2012).

quadratic models have been proposed for measuring the health risks of radiation. The loss (negative utility) of a mistaken model could be equated with its distance from truth, if the problem is purely theoretical and belongs to basic research. But if the safety standards are adopted by implementing the model in practice for the public and the workers in nuclear facilities, then it is safer to overestimate the health risks than to underestimate them. Hence, the practical interest of protecting people from radiation justifies a loss function which gives higher penalties for low risk estimates.

Design Sciences

Herbert Simon (1969) called attention to an important type of applied science: while descriptive basic sciences are concerned with "how things *are*", the "sciences of the artificial" tell us "how things *ought to be* in order to attain goals, and to function". They can be called *design sciences*, as they seek knowledge about design activities in the broad sense, i.e. knowledge that is useful in the science-based shaping and planning of artificial human-made systems. It is important that design sciences as a form of research should not be confused with design itself.[9] Together with predictive science (like predictive astronomy, meteorology, and econometrics), design science is the main form of applied science.[10]

Design sciences typically emerge by the "scientification" of productive arts and their professions.[11] First the practical skills are based on cumulative everyday experience and trial-and-error, then they are expressed by rules of thumb which are further developed into guide books. The next step is the scientific study of the rules by testing their efficacy and function with experiments. Examples include engineering sciences, agricultural sciences, evidence-based medicine, and evidence-based social policies. Thus, design sciences do not fit the linear model

---

[9] In the same way, we have to distinguish scientific research as knowledge-seeking from the applications of science in practical decision-making and problems-solving. Operations Research (OR) is a systematic tool of such decision-making (see Churchman, Ackoff, and Arnoff, 1957).

[10] See Niiniluoto (1993).

[11] See Niiniluoto (1995).

of innovation, but still we shall see that they can rely on theories and methods of basic research.

While basic sciences are assessed by epistemic utilities, applied research can be evaluated by a double standard. First, as applied sciences seek knowledge, their claims can be assessed by epistemic utilities: they should be true, informative, and justified. Secondly, they should be applicable or instrumentally relevant for some human activity. This kind of *relevance* requirement of practical utility cannot be directly demanded of basic research – even though the history of science shows that many important applications of science have been based upon fundamental research (cf. Chapter 10 above). Design sciences are thus value-laden in a sense which is different from value-free basic sciences (like physics, chemistry, geology, biology, history, ethnology, and sociology). The success of agricultural science partly depends on its ability to improve the productivity of fields and forests (without hazarding sustainable development), the success of medicine and pharmacology is shown by the improvement of human health (without fatal side-effects), and the success of social policy studies is shown by the well-being of citizens.

Following Simon's idea, design sciences are concerned with natural and social systems that can be *manipulated* by human actions. Astronomy can explain and predict eclipses, but not produce them. Meteorology is a predictive science, as human intervention in weather conditions is (still) largely beyond human capacities and technologies. But relative to each kind of artificial system, which can be designed and influenced by human actions, there is potentially a branch of design science.

Technical Norms

Simon (1969) already hinted that design sciences are special kinds of *normative* sciences which give us justified knowledge about means - ends relationships. As proposed by Niiniluoto (1993), this idea can be expressed by formulating the knowledge claims of applied design sciences by technical norms.

Basic descriptive sciences establish causal statements of the following form:

(C) X causes A in situation B.

Such general laws may be deterministic or probabilistic, and they can be used for explanation (answering why A occurred in situation B) and prediction (forecasting that A will occur after X in situation B).[12] If the factor X can be manipulated by human action, (C) can be expressed by a statement involving agent causality:

(AC) Doing X causes A in situation B.

(AC) justifies rules of the form "Do X in situation B in order to get A".[13] This rule tells that doing X is a *means* to the *end* A in circumstances B. By converting it to a conditional statement, where the end A is taken as the antecedent, we obtain a *technical norm* in the sense of G. H. von Wright (1963):

(TN) If you want A, and believe that you are in situation B, then you ought to do X.

According to von Wright, genuine norms or oughts arise from commands given by norm-authorities. In contrast, "technical oughts" are natural in the sense that they are based on causal or "anankastic" relationships (cf. (C) above). His paradigmatic example is "Unless the house is heated, it will not be habitable", which states that heating is a necessary condition or necessary cause of making the house habitable. The technical norm "If you want to make the house habitable, you ought to heat it", which is conditional on someone's wants, should be distinguished from the anankastic statement.[14]

As imperatives, unconditional recommendations of the form "You ought to do X!" or conditional recommendations "Given B, you ought to do X!" lack truth values. But do technical norms TN have truth values? Von Wright hesitated to make this conclusion. He pointed out that statements about necessary relations and wants may be true, but still he left the

---

[12] See Fetzer (1981).

[13] See Bunge (1963). Kuipers (2013) formulates design laws in terms of properties rather than actions. His counterpart to (AC) is "Imposing structural property X is context B causes functional property A".

[14] In his examples, von Wright usually does not mention beliefs about the relevant situation B, but his treatment of practical inference always includes two premises, one about intentions or wants, the other about beliefs concerning necessary conditions (see von Wright, 1983).

relation of technical norms to truth and falsehood as open.[15] Later he stated that the end and the necessary connection constitute a "foundation" or "justify" the normative conclusion of a technical norm.[16] So a person with appropriate wants and beliefs has a "technical ought" to do something. He also accepted the reduction of technical ought O_t to alethic modal logic by Alan Ross Anderson's formula $O_tp =_{df} N(S \rightarrow p)$, where N is the necessity operator and S is some good end (*ibid*., 154).

In my view, the technical norm TN can be taken to be true or false, depending on whether X is a necessary cause of A in situation B.[17] If X is a sufficient but non-necessary cause of A, where X does not bring about harmful side-effects, then the technical norm TN can be reformulated with the conclusion "it is profitable for you to do X". Similar reformulations can be given in cases where X is a probabilistic cause of A. More generally, the want A could be replaced by a description of the value preferences of an agent, and the beliefs about the situation B by her entire belief state, so that the conclusion could be restated as "it is rational for you to do X". The general form of such conditional recommendations is thus

(CR) Given that you have values A and beliefs B, it is rational for you do X.

TN is a relatively simple special case of such statements which can be justified in decision theory.

A further complication is that in decision theory there are different principles of rationality, like minimax and subjective expected utility (SEU).[18] One possible reaction is to include one's favorite decision criterion in the value profile of the agent, but this would make conditional recommendations quite complicated. The relativist proposal would be to say that the truth of CR is relative to a decision criterion. Perhaps the most straightforward alternative is to incorporate decision

---

[15] See von Wright (1963), 103.

[16] See von Wright (1983), 74.

[17] For an interesting attempt to develop formal semantics for means –end relations within dynamic logic, see Hughes, Kroes & Zwart (2007).

[18] See Luce and Raiffa (1957).

criteria in the consequent of CR, by letting CR conclude that it is minimax-rational for you to do X or SEU-rational to do Y.[19]

Justification of Technical Norms

As statements with a truth value, technical norms TN – and, more generally, conditional recommendations CR - can be results of scientific research.[20] For the idea of design science, it is therefore crucially important to consider different ways of justifying such recommendations of action.

There are three main ways of seeking knowledge about research domains and topics. One is directly by observation and experimentation. The second is indirectly by theorizing with the method of testing hypotheses. The third is modeling, where the study of models and simulations allows analogical inferences to the target systems.[21] Models have been characterized as "epistemic artifacts" by Tarja Knuuttila (2005). By the same token, theories as human constructions are artifacts as well, aimed at representing the domain of investigation (cf. Chapter 4 above). Thus, one could add the seventh E, *epistemic worth*, in the earlier formula 6E+S of assessing artifacts. In other words, the cognitive value of theories and models depends on their epistemic utility.

Observations, theories, and models may all play a role in the justification of technical norms. The derivation of the technical norm TN from the causal laws C and AC illustrates how TN can be justified "from above" by a basic theory. Causal laws may be parts of mathematical models with parameters which have to be estimated "from below" by empirical data. Ballistics

---

[19] For the difference between minimax and SEU in the issue of nuclear power policy, see Levi (1980), Appendix.

[20] Anna Alexandrova (2018) raises the question whether "mixed claims" like (C) can objective, if they contain a value term A like "well-being" or "health". She proposes that the value presupposition of A should be made explicit, which could be accomplished by giving an operational definition of the thick value concept A. For the "normative presuppositions" of mixed claims she recommends "deliberative polls" by concerned parties. However, her thesis that none of existing proposals to reconcile values with objectivity are suitable for mixed claims overlooks the possibility that e.g. the science of well-being can be a design science with technical norms.

[21] See Niiniluoto (2012).

is an example of applied science which combines theoretical information from Newton's mechanics with experimental information from "shooting tables".[22]

An important case of modeling is the use of the problem-solving methods of Operations Research to justify – instead of particular decisions – general conditional recommendations. A typical OR method, used in linear programming, is to construct a mathematical model which expresses the "effectiveness" E of a system under study as a function $E = f(x_i,y_j)$ of variables x_i subject to control by the decision-maker and variables y_j not under such control. The problem is then reduced to the choice of the values x_i with some constraints, which maximize the function E given the estimated values of y_j. Here the end is expressed by the optimal value of the function E, and situation is described by the values of y_j and the constraints on the permissible values of x_i.[23] The general policy result from such a model can thus be expressed as a technical norm.

Social constructivists may doubt the possibility of objective truths of the form TN, since for them the claims about the effectiveness of technological skills and procedures are relative to social interests.[24] I don't think that this is a valid objection to the idea of design science. If the causal factor X is manipulable in situation B, then the possibility of reaching the goal A depends on causal or nomic regularities in nature, not on our beliefs or interests. Even in the case of social systems there are regularities and tendencies and habits in human behavior. In the implementation of social technologies (e.g. lowering interest rates in economy or distributing contraceptives for birth control), the effects may be sensitive to many kinds of situational factors.

A more interesting challenge to design sciences comes from the argument of Hubert L. Dreyfus against the scientification of human arts.[25] He claims that the skills of true experts cannot be formulated by explicit rules in language. Simple rules are needed only by novices. However, it can be argued that the increasing power of technological sciences and evidence-based

---

[22] See Niiniluoto (1994).
[23] See Churchman et al. (1957), 13.
[24] See Bijker, Hughes & Pinch (1987).
[25] See Dreyfus and Dreyfus (1984).

medicine in guiding human action indicate that these fields have been able to find at least truthlike technical norms.[26]

The Place of Values in Design Sciences

We have seen above that the justification of technical norms can be *value-neutral* in the sense that it relies on theoretical and empirical information assessed by epistemic utilities - and commitment of the researcher to the conditional value A is not needed. Technical oughts are binding only for those who accept the antecedent value premise of the technical norm. For example, a pacifist may accept the results of military studies, and a militarist may approve the conclusions of peace research – even though they will disagree in the social and political *relevance* of their respective fields. Still, the conditional norm TN is indispensably *value-laden* in the sense that it essentially involves a value premise as its antecedent. Hence, design sciences do not satisfy any principle of value-freedom which would require that all value terms are excluded from their language.[27]

Our account of design sciences shows that any social value could take the place of the value A of a technical norm. Such values are not criteria for accepting the theories or models which are used to justify the claims of design sciences, but rather they are antecedents of technical norms or conditional recommendations of action.

Design sciences can be used for rational planning and decision-making, when the end A is accepted as a basis of action. Practical inference starts from the want A and the belief B and leads, via (TN), to the action X. Design sciences thus give instrumentally rational tools for promoting the end A. In this

---

[26] Cf. Niiniluoto (2007).

[27] There is an influential trend in the social studies of science to reject the traditional internal – external dichotomy: "science *is* society, inside and out" (Cozzens and Gieryn, 1990, 1). We have maintained this dichotomy in the context of basic research, which uses epistemic utilities in the internal evaluation of cognitive hypotheses. Technical norms of design sciences are also grounded relative to epistemic utilities, but their formulation may include as goals A any external social values. This is a straightforward deconstruction of the external –internal dichotomy.

sense we may say that a scientist working in a design science is morally responsible for the results of this activity.

The relevant value goal A may be characteristic to a special design science. For example, engineering sciences have their own technological utilities. Other examples include health for medicine and nursing science, profit for business economics, welfare for social policy studies and social work, and peace for peace research. The case of medicine shows that for many design sciences the choice and specification of the relevant value goal may be a matter of philosophical, legal, ethical, and political dispute. The sources of values of technical norms can be sought in philosophical arguments, general morality and ethics, situational preferences, empirical value studies, value profiles of institutions and funding bodies of research, and political debates.

Especially in the social sciences, the end A may be conservative (preservation of status quo), reformist (Popperian piecemeal social engineering) or emancipatory (radical changes in the social order). But it is also important to observe that the division between the situation B and manipulable factors X in TN is not only relative to our abilities, but sometimes also to our value-laden decision to keep B constant and to let X change. A Marxist political economist would be ready to change the prevailing structures of capitalism.

Finally, as noted by Niiniluoto (1993), the notion of technical norm illustrates the nature of policy conflicts. It is often the case that design scientists e.g. in the fields of environmental studies or economics give conflicting unconditional recommendations with hidden value premises. For example, should energy problems be solved by nuclear power plants, fossil fuels (coal oil, natural gas) or renewable sources (solar power, wind)? Such disagreement about the best policies X may be due to differences in knowledge about the situation B, decision to keep B stable or change it, knowledge about the law $X\&B \rightarrow A$, the valuation of the goal A, and the assessment of the risks of alternative actions. Technocrats, who are blind to value issues, think that such conflicts always can be resolved by empirical research. Anti-science positions, which reject the knowledge provided by research, may be inclined to radical political power play. The concept of design science shows that successful approaches to policy conflicts may need more

research together with respect for enlightened democratic discussion about values.

Note. This chapter was first presented on the conference VaNiM (Values and Norms in Modeling) at the Eindhoven University of Technology, The Netherlands, June 26, 2012. It was published in *Studies in History and Philosophy of Science* 46 (2014), 11-15. It is published here by the permission of Elsevier.

Literature

Alexandrova, Anna (2018). "Can the Science of Well-Being be Objective?", *The British Journal for the Philosophy of Science* 69, 421-445.

Bijker, Wiebe, Hughes, Thomas & Pinch, Trevor (eds.) (1987). *The Social Construction of Technological Systems*. Cambridge, MA: The MIT Press.

Bunge, Mario (1966). "Technology as Applied Science". *Technology and Culture*, 7, 329-349. Reprinted in Friedrich Rapp (ed.) (1974), *Contributions to a Philosophy of Technology*. D. Reidel: Dordrecht, 19-36.

Churchman, C. W., Ackoff, R. L. & Arnoff, E. L. (1957). *Introduction to Operations Research*. New York: John Wiley.

Cozzens, Susan E. & Gieryn, Thomas F. (eds.) (1990). *Theories of Science in Society*. Bloomington: Indiana University Press.

Dreyfus, Hubert & Dreyfus, Stuart (1984). *Mind Over Machine*. Oxford: Blackwell.

Fetzer, James (1981). *Scientific Knowledge: Causation, Explanation, and Corroboration*. Dordrecht: D. Reidel.

Godin, Benoit (2006). "The Linear Model of Innovation: The Historical Construction of an Analytical Framework", *Science, Technology, and Human Values* 31, 639-667.

Hughes, Jesse, Kroes, Peter & Zwart, Sjoerd (2007). "A Semantics for Means-End Relations", *Synthese* 158, 207-231.

Knuuttila, Tarja (2005). "Models, Representation, and Mediation", *Philosophy of Science* 72, 1260-1271.

Kuipers, Theo (2013). "Philosophy of Design Research". In H. Andersen, D. Dieks, W. Gonzalez, T. Uebel & G. Wheeler (eds.), *New Challenges to Philosophy of Science*. Dordrecht: Springer, 457-466.

Levi, Isaac (1967). *Gambling With Truth: An Essay on Induction and the Aims of Science*. New York: Harper & Row.

Levi, Isaac (1980). *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*. Cambridge, MA: The MIT Press.

Longino, Helen (1986). "Biological Effects of Low-Level Radiation: Values, Dose-Response Models, Risk Estimates", *Synthese* 81, 391-404.

Luce, R. Duncan & Raiffa, Howard (1957). *Games and Decisions: Introduction and Critical Survey*. New York: John Wiley.

Maldonado, Thomas and Bonsiepe, Guy (1964). "Science and Design", *Ulm*, 10/11, 10-29.

Mitcham, Carl (1994). *Thinking through Technology: The Path between Engineering and Philosophy*. Chicago: The University of Chicago Press.

Niiniluoto, Ilkka (1984). "Philosophical Perspectives on Design". In Susann Vihma (ed.), *Designforschung – Design Research*. Helsinki: University of Industrial Arts, 13-31.

Niiniluoto, Ilkka (1993). "The Aim and Structure of Applied Research", *Erkenntnis* 38, 1-21.

Niiniluoto, Ilkka (1994). "Approximation in Applied Science". In Martti Kuokkanen (ed.), *Idealization VII: Structuralism, Idealization and Approximation*. Amsterdam: Rodopi, 127-139.

Niiniluoto, Ilkka (1995). "The Emergence of Scientific Specialties: Six Models". In W. E. Herfel, W. Krajewski, I. Niiniluoto & R. Wojcicki (eds.), *Theories and Models of Scientific Processes*. Amsterdam: Rodopi, 211-223.

Niiniluoto, Ilkka (2007). "Evaluation of Theories". In Theo Kuipers (ed.), *Handbook of the Philosophy of Science: General Philosophy of Science - Focal Issues*. Amsterdam: Elsevier, 175-217.

Niiniluoto, Ilkka (2012). "The Foundations of Statistics: Inference vs. Decision". In D. Dieks, W. Gonzalez, S. Hartmann, M. Stöltzner & M. Weber (eds.), *Probabilities, Laws, and Structures*. Dordrecht: Springer, 29-41.

Niiniluoto, Ilkka (2013). "Models, Simulations, and Analogical Inference." In Vassilios Karakostas & Dennis Dieks (eds.), *EPSA11: Perspectives and Foundational Problems in Philosophy of Science*. Cham: Springer, 19-27.

Niiniluoto, Ilkka (2016). "Science vs. Technology: Difference or Identity?". In Maarten Franssen, Pieter Vermaas, Peter Kroes & Anthonie Meijers (eds.), *Philosophy of Technology after the Empirical Turn*. Cham: Springer, 93-106.

OECD (1962). *The Measurement of Scientific and Technological Activities: Proposed Standard Practice for Research and Development*. Paris: OECD.

Rudner, Richard (1953). "The Scientist *qua* Scientist Makes Value Judgments". *Philosophy of Science* 20, 1-6.

Simon, Herbert (1969). *The Sciences of the Artificial*. Cambridge, MA: The MIT Press. (2nd ed. 1981.)

Vitruvius, Pollio (1914). *The Ten Books on Architecture*. Cambridge, MA: Harvard University Press.

von Wright, G. H. (1963). *Norm and Action*. London: Routledge and Kegan Paul.
von Wright, G. H. (1983). *Practical Reason*. Oxford: Blackwell.

Chapter 12: Futures Studies: Science or Art?

Introduction

Futures studies is still seeking its identity as a new academic discipline. Work in this area ranges from the development of highly sophisticated mathematical methods and models to political activities in "making the world a better place in which to live".[1] Its self-understanding has reflected the conceptions of science among the researchers and practitioners in this field. It is thus no wonder that the discussion about the nature of futures studies has borrowed ideas from many traditions in the philosophy of science. But it has turned out to be difficult to place this branch of studies within the conceptual framework of such trends as positivism, hermeneutics, and critical theory. Therefore, I will argue, it is more instructive to look at the pattern of the emergence of new scientific disciplines.[2]

In this chapter, I make an attempt to approach futures studies with the distinction between descriptive science and design science (cf. Chapter 11 above). Instead of the more traditional dichotomy between basic and applied research, this distinction seems to be helpful in the task of understanding the nature of many new "practical" and "professional" disciplines and applied social sciences. It also serves to clarify the question raised in my title: the opposition between "science" and "art" is not here the one between natural science and the humanities, but rather the Latin distinction between *scientia* (as a form of knowledge) and *ars* (as a form of skill). As *ars* is a translation of the Greek term *techne*, the question can be reformulated as the query whether futures studies is a science or a form of social technology.

---

[1] Bell (1997a), 33.

[2] See Niiniluoto (1995).

Descriptive Knowledge about the Future

Perhaps the most common model of scientific research is *descriptive*: at least the so-called scientific realists claim that the basic aim of science is to give true or truthlike information about reality, where reality includes nature, mind, culture, and society (cf. Chapter 11 above). In other words, scientific knowledge consists of statements, typically in the indicative mood, which attempt to truthfully describe some singular or general facts. Singular facts may be states of affairs that obtain in the present, past, or future. General facts are described by laws of nature or other lawlike statements. Knowledge about such laws helps us in giving scientific explanations as answers to why-questions (Why did this event happen?), making predictions about future events, and postdictions about past events.[3] Thus, the tasks of descriptive science include systematic studies and explanations of the present state of reality and its lawlike regularities, historical or postdictive studies about the past, and predictive studies about the future.

Many "ordinary" scientific disciplines - like physics, astronomy, psychology, and economics - have futuristic relevance in the sense that their theories, together with initial conditions about the present, yield predictions about observable events in the future. Without such ability the theories would not satisfy the requirement of empirical testability. As it does not make sense to explain future events, the descriptive model implies that futures studies should be a *predictive* science. In this view, the futurist's role is to use the scientific method to make forecasts and projections - and thereby to replace various kinds of unscientific "prophecies" and "prognoses" proposed by religious thinkers, philosophers, novelists, and fortune tellers.

But is such "foresight" possible as a form of *knowledge*? This question has been discussed and debated by philosophers already for more than 2000 years, starting from Aristotle's example of the sea battle tomorrow.[4] Recall that, according to Plato's classical definition, knowledge means the same as justified true belief. So are there propositions that we already now know to be true e.g. in the year 2050? As trivial answers we may mention mathematical statements (like "2 + 3 = 5") which

[3] See Hempel (1965).
[4] See Hintikka (1963).

are eternally true. It is also plausible that the basic laws of nature do not change, so that, for example, the statement expressing the law of gravitation will be true in 2050. Further, the truth of some future statement is determined with mathematical necessity by past facts: for example, in the year 2050 the city of Helsinki will have its 500th anniversary.

True statements of this kind are not very interesting for future studies, however. It would be more relevant to seek predictions about contingent events or states of affairs in the future – such as the number of mobile phones or electric cars in Finland in 2050. The obvious difficulty is that these events have not yet been realized, and in this sense the future does not exist. Hence, the statements about future contingents do not have now truth values, and cannot be known now in Plato's sense.[5] It thus seems that at best we can now *guess* what will happen in the future.

As noted by Jan Lukasiewicz (1970), one way out from this problem would be to assume that the world is deterministic, so that the present state of the world completely determines its future states. If this is the case, all statements about future have already now a fixed but usually unknown truth value. But such a metaphysical assumption about the world is not very plausible.[6] At best there are some relatively isolated systems which behave according to deterministic laws. For example, classical mechanics allows us to derive predictions about the future positions of the planets. But these predictions are valid only on the *ceteris paribus* assumption that our solar system remains closed and no external disturbance brings about unexpected changes.

Recent work in chaos theory shows, moreover, that a deterministic system may be unstable in the sense that it is highly sensitive to small variations in its initial conditions.[7] In this case, statements about the future of the system would have determinate truth values, but no finite human being or computer

---

[5] See Knuuttila (2020).

[6] See e.g. Suppes (1984).

[7] See Earman (1986) about this "butterfly effect". For chaotic systems and the possibility of their management, see Eriksson (2017). For prediction in general, and in economics in particular, see Gonzalez (2015).

could have sufficiently precise knowledge of its initial conditions and, thereby, of its future behaviour.

Future as an Open Tree

The arguments in the preceding section suggest that futures studies have some, but quite limited, prospects as a descriptive science. An alternative view might be suggested by claiming that the object of futures studies is not the future but the present. In the same way it has been claimed that the object of historical research is constituted by the present traces of past events. I find this misleading, although the present is the common "empirical basis" of historical and futurological research. In my view, it is important to make a clear distinction between the *object* and the *evidence* of research: knowledge about the present is evidence for statements about the past in historical studies and about the future in futures studies.[8]

A more promising approach is to view the future as a branching *tree* with alternative possibilities.[9] The future is still open to some extent, as its features will depend on chance events (e.g. natural and technological accidents) and human choices (e.g. the terrorist attack 9/11, Russian invasion of Ukraine in 2022). The task of the futurist is to investigate this tree as a whole. Its branches define alternative scenarios, which we may wish to realize or avoid. More precisely, the received view is that the futurist should (i) construct alternative possible futures, (ii) assess the probability of alternative futures, and (iii) evaluate the preferability or desirability of alternative futures.[10]

Various graphical, statistical, and quantitative methods have been developed for the tasks (i) and (ii). The task (ii) is in fact a weakening of the descriptive predictive model: instead of knowing in advance the truth value of a statement about a future event we try to estimate its *probability*.[11] Improbable

---

[8] Niiniluoto (2014).

[9] See von Wright (1971), Ch. II. Malaska and Virtanen (2017) follow de Jouvenel in speaking about "futuribles" (a fan of future possibilities) – and "future maps", "future manifolds", and "future spaces".

[10] See Cornish (1969), Bell (1997a).

[11] Probability theory allows us to investigate some general properties of indeterministic systems. A well-known example, studied in detail by

events with great impact are called "black swans" by N. N. Talib. The reliable estimation of objective risk probabilities for natural, human, and social systems is in principle a scientific endeavour, but in many cases a difficult task with incomplete information. As Bertrand de Jouvenel (1967) has said, the futurist often has to apply here "the art of conjecture".

Edward Cornish (1969), the founder of the World Future Society, states that the futurist needs "artistic" imagination in the construction of alternative scenarios. This is true, but it should be remembered that creative imagination is needed in the discovery of scientific theories and new technological artefact as well.[12] The most important systematic difficulty for the task (i) arises from the fact that the future - and the alternatives to be depicted in the future tree - depend on new scientific and technological discoveries. As Karl Popper (1957) argued, such novel discoveries cannot be known in advance: if we knew them already now, they would not be future discoveries. It can be pointed out against Popper that sometimes it may be possible to know or guess that something *will be* technologically possible, even though we do not yet know *how* (e.g. Leonardo da Vinci's visions of aeroplanes and parachutes). But still it is relevant to see that the realm of possibilities may change in time in an unpredictable way.

Preferable Futures

Already Ossip Flechtheim in 1943 demanded that futurology should be committed to improving the freedom and welfare of humankind. Futures study has an emancipatory function: instead of passive prediction, it should engage in political dialogue and thereby transform society. Following Mika Mannermaa's (1986) project of alternative futures, Wendell Bell in his *Foundations of Futures Studies* states that proposing action is "part of futurist's job *qua* futurist".[13] These commitments are built into the third task (iii) of assessing preferable

---

mathematicians, is *random walk*, where alternative steps are equally probable. A memorable illustration was given by Carl Barks in 1953 in a Donald Duck strip, where a nutty Professor Batty persuades Donald to make decisions by flipping a coin.

[12] See Verschraegen et al. (2019).

[13] Bell (1997a), 97.

futures, which is clearly different from the traditional ideal of value free descriptive science.

I personally share with many futurists the humanistic ideals of individual freedom, social responsibility, democracy, and sustainable development. However, it is important to ask in what sense the acceptance of these values could be a part of the scientific study of future. Can futures studies fulfil the evaluative task (iii) and maintain its status as a scientific discipline? Is it possible to justify *value assertions* by using the scientific method?

By empirical research we can at best identify what is actually regarded as good or bad by different persons or cultures. Such a relativity of preferences does not imply that value judgments are arbitrary, purely subjective opinions, or merely emotional expressions. It is the task of philosophy - ethics and politics as Aristotelian "practical sciences" - to find arguments for axiological and normative views. Some important philosophers from Plato and Aristotle to Kant and Habermas have claimed that there are rational or discursive methods for justifying value statements and prescriptions, but these different forms of value objectivism are controversial, since they violate Hume's guillotine: from facts, or how things *are*, one cannot derive how they *ought to be* (cf. Chapter 10 above). Thus, while Bell's (1997b, 87) battery of tests for value assertions is in many ways useful, it is at least problematic to suggest that it allows us to develop "a consistent, coherent, rational, and objective morality".

Bell (2009) appeals to Keekok Lee's method of giving "serious, relevant, and empirically true" reasons for value statements. For example, people ought not to smoke tobacco, because smoking probably will increase their chances of dying of lung cancer at some future time. However, this argument does not avoid Hume's guillotine, since it presupposes the value premise that dying of cancer is a bad or undesirable thing.

Even granting that futurists could successfully act as practical philosophers, their putative knowledge of objective morality would not change the need to distinguish between the value commitments of the futurists themselves, their employers, and the actual and possible persons occupying positions in the future tree. It is not only a fact of life that such values may differ from each other - assuming the principles of liberal democracy, it is also desirable that people are allowed to make their own morally and politically relevant choices.

Designing the Future

If futures studies does not quite fit into the traditional model of descriptive science, is it perhaps a new form of a "critical theory", a combination of scientific inquiry and social criticism, whose "emancipatory" interest is not reducible to the "technical" interest of the natural sciences? (cf. Chapter 10 above). I think this is doubtful, too. In my view, the key to futures studies should be sought in those research areas that Herbert Simon (1981) has called "design sciences" or "sciences of the artificial".

By *design* or *planning* in the broad sense we may mean any activity where optimal means are used in a systematic way for reaching accepted ends. The goal may be the construction of a material artefact (such as field, landscape, work of art, craft, engine, or building), social organization, solution of a problem, or a decision to act. Such an activity is value-laden in the sense that the goal is taken to be desirable or valuable. It is rule-governed to the extent that the process is guided by general rules of action. But conceptually such constructive planning and problem-solving should be distinguished from scientific research which aims at new knowledge.

It has been suggested that futurology is a new form of planning:[14] the future can be regarded as an artefact that is created by human actions. In this respect, futures studies would not be a knowledge-seeking activity, but rather a form of social technology, "future planning", comparable to the more restricted field of urban planning.

I think this is an important insight, and explains why many futurists have emphasized the importance and relevance of the value-laden task (iii). But it still fails to analyse the role of scientific research within the study of future - a feature which distinguishes this field from ordinary political activities.

For many practical activities (such as farming, engineering, and nursing), and for the associated professions (farmer, engineer, nurse), there are corresponding scientific disciplines (agricultural science, engineering science, nursing science) which seek scientific information that is intended to make this practice or art more effective. Simon (1981) states that such *design sciences* do not tell how things *are*, but how they *ought to be* in order to attain some ends. In other words, such sciences do not fit into the descriptive model of research. Instead, their results

---

[14] See Julien, Lamonde & Latouche (1975).

typically express relations between means and ends. For example, a practising medical doctor accepts the assumption that health is a value. To maintain and improve the health of her patients, on the basis of so-called evidence-based medicine, she needs rules which tell what she ought to do in different types of circumstances to achieve this goal.

Statements about means - ends relationships are called *technical norms* by G. H. von Wright (1963). The general form of such norms is

(1) If you want A, and believe that you are in situation B, you ought to do X.

Unlike categorical or non-conditional norms (You ought to do X!), technical norms have truth values. Statement (1) is *true*, if doing X in situation B is a necessary condition of reaching the goal A. (In weaker variants, X may be a probable sufficient condition of A, and the conclusion states that it is rational for you to do X.) To justify (1), we should be able to show that

(2) X causes A in situation B.

Such a general causal law may be derivable from more general scientific theories; this is a typical case in the applied sciences which use information taken from basic research. But its justification often has to be based upon approximate mathematical models together with empirical or experimental data (cf. Chapter 11 above). These methods respect the scientific principle of value neutrality, but still they serve to justify via (2) conditional statements (1) with a value premise in their antecedent.

To accept (1) as true, one need not personally be committed to the goal A. But if we accept the goal A, and have good scientific evidence for the fact that we are in a situation of type B, we can derive from these valuations and factual premises the recommendation of doing X.

The main point is that the instrumental statement (1) may be true and well-justified, and thus satisfy the classical conditions of knowledge. As (1) does not describe how the world is, but rather how it ought to be if a goal is desired, (1) cannot be a result of science in the descriptive model. But it may a part of what we may call *design science*.

Assuming that we are designing the future, in the relevant technical norms the situation B may our present state of nature and society (including the needs and hopes of real people) or any possible state of affairs in the future tree. The goal A may

be the realization of some preferable future state (e.g. increasing the number of jobs or public services) or the avoidance of some undesirable threat (e.g. ecological catastrophe or nuclear war). As the choice of A may a matter of ethical and political controversy, the futurist can study technical norms with alternative goals A. Some goals may be conservative (the preservation of status quo), some emancipatory (a radical change or alternative trend). For a utopian goal, there is no action X and no technical norm available, but the situation may change if the underlying state B is itself transformed. The recommended actions X (relative to B and A) express means for achieving the desired goals. Their sequences thus constitute alternative scenarios for the future.

The scientific method of planning the future resembles the *abductive* or *retroductive* reasoning from effects to preceding causes.[15] This inference is familiar from historical research. What is sometimes called "backcasting" (in contrast to forward-looking forecasting) proceeds from aims to means: a desirable goal or vision in short or long run is set, and then rational means for reaching it are systematically searched in a carefully charted but changing environment.[16]

The notion of technical norm shows how the planning of future can combine knowledge and values. Futures studies can indeed help in making the world a better place, but its model should be design science instead of descriptive science. But I do not wish to reduce futures studies merely to design science, since its practitioners will always include philosophers who are creative in thinking about alternative valuable goals and activists who implement such goals in their action. It is, indeed, important that futures studies provide for the future (cf. Chapter 1 above) in the sense that its results (i.e. justified technical norms) can be taught to citizens, who by this "futures education" obtain "future literacy" and "futures competence" – and thereby are able to engage in the activity of transforming the society.[17]

I conclude that futures studies, when it combines the tasks of exploring probable and preferable futures, is a mixture of theoretical and empirical research, methodology, philosophy, and political action. But in its core we find a design science which attempts to help the rational planning of our future.

---

[15] For Peirce's account of abduction, see Niiniluoto (2018).

[16] See Neuvonen (2022).

[17] See the articles in Heinonen et al. (2017).

Note. Earlier versions of this article have been published in Finnish in *Futura* 6:1 (1987), 42-47, and in English in *Futura* 28:1 (2009), 59-64. See also Heinonen et al. (2017), 22-27.

Literature

Bell, Wendell (1997a). *Foundations of Futures Studies, vol. 1: History, Purposes, and Knowledge*. New Brunswick and London: Transaction Publishers.

Bell, Wendell (1997b). *Foundations of Futures Studies, vol. 2: Values, Objectivity, and the Good Society*. New Brunswick and London: Transaction Publishers.

Bell, Wendell (2009). "Moral Discourse, Objectivity, and the Future". *Futura* 28 (1), 43-58.

Cornish, Edward (1969). "The Professional Futurist". In Robert Jungk & Johan Galtung (eds.), *Mankind 2000*. Oslo: Universitetsförlaget.

Earman, John (1986). *A Primer on Determinism*. Dordrecht: D. Reidel.

Eriksson, Jarl-Thure (2017). "Chaos Theory and the Manageability of Complex Systems." In Heinonen et al., 64-91.

Gonzalez, Wenceslao (2015). *Philosophico-Methodological Analysis of Prediction and its Role in Economics*. Cham: Springer.

Heinonen, Sirkka, Kuusi, Osmo & Salminen, Hazel (eds.) (2017). *How Do We Explore Our Futures? Methods of Futures Research*. Helsinki: The Finnish Society for Futures Studies.

Hintikka, Jaakko (1973). *Time and Necessity: Studies in Aristotle's Theory of Modality*. Oxford: Oxford University Press.

de Jouvenel, Bertrand (1967). *The Art of Conjecture*. New York: Basic Books.

Julien, Pierre-André, Lamonde, Pierre & Latouche, Daniel (1975). *La Méthode des Scénarios. Paris:* La Documentation Francaise.

Knuuttila, Simo (2020). "Medieval Theories of Future Contingents." In Edward Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), https://plato.stanford.edu/archives/sum2020/entries/medieval-futcont/

Lukasiewicz, Jan (1970). *Selected Works.* Amsterdam: North-Holland.

Malaska, Pentti & Virtanen, Ilkka (2017). "Theory of Futuribles and Historibles". In Heinonen et al., 121-138.

Mannermaa, Mika (1986). "Futures Research and Social Decision Making: Alternative Futures as a Case Study". *Futures* 18 (5), 658-670.

Neuvonen, Aleksi (2022). *Refocusing on the Future: Backcasting Carbon Neutral Cities*. Tampere: University of Tampere.

Niiniluoto, Ilkka (1995). "'The Emergence of Scientific Specialties: Six Models." In W. E. Herfel, W. Krajewski, I. Niiniluoto & J. Wolenski (eds.), *Theories and Models of Scientific Processes.* Amsterdam: Rodopi, 211-223.

Niiniluoto, Ilkka (2014). "Scientific Realism: Independence, Causation, and Abduction." In K. R. Westphal (ed.), *Realism, Science, and Pragmatism*. New York: Routledge, 159-172.

Niiniluoto, Ilkka (2018). *Truth-Seeking by Abduction*. Cham: Springer.

Popper, Karl R. (1957). *The Poverty of Historicism*. London: Routledge and Kegan Paul.

Simon, Herbert (1981). *The Sciences of the Artificial*, 2nd ed. Cambridge, MA: The MIT Press.

Suppes, Patrick (1984). *Probabilistic Metaphysics.* Oxford: Blackwell.

Verschraegen, Gert, Vandermoere, Frederic, Braeckmans, Luc & Segaert, Barbara (eds.) (2019). *Imagined Futures in Science, Technology and Society*. Abingdon: Routledge.

von Wright, G. H. (1963). *Norm and Action*. London: Routledge and Kegan Paul.

von Wright, G. H. (1971). *Explanation and Understanding.* Ithaca: NY: Cornell University Press.

Chapter 13: The Idea of a University: Humboldt vs. Newman

Universities have a long history, with roots in the ancient philosophical schools in Greece. The first medieval universities based professional education in higher faculties (theology, law, medicine) upon propaedeutic studies in the faculty of philosophy. After the Eighteenth century Enlightenment, which promoted education for all, universities as centers of learning wished to separate themselves from vocational schools and new polytechnic institutes of technology. Two leading figures in this movement were John Henry Newman in Dublin, Ireland, and Wilhelm von Humboldt in Berlin, Prussia. For Newman, who published his classic *The Idea of a University* in 1852, the university is a place of teaching universal knowledge and the real cultivation of mind, but not of extending knowledge by research. For Humboldt, who founded the University of Berlin in 1810, university educates its students by teaching them critical methods of scientific inquiry. The unity of research and teaching is realized in research-based teaching. An eloquent formation of this German model of a university was given by the Finnish philosopher Johan Vilhelm Snellman at the University of Helsinki in 1840. The Humboldtian model placed philosophy and the humanities in the center of academic life. Today Newman's tradition can be found in the liberal arts colleges in the United States, while Humboldt's ideal is dominant in the concept of a research university. Research universities score high in international university rankings, but they have adopted additional tasks and administrative practices from entrepreneurial universities.

Universities have a Long History

The long history in critical thinking and cultivation of learning has its ancient roots in Greece. Some philosophers of nature, who distinguished their thinking from religious groups, already founded their own schools after 600 BCE. The philosopher Socrates (ca. 469 – 399 BCE) acted as a "mental midwife", who challenged people in dialogues in parties (*symposia*) and the *Agora* of Athens. He contrasted the genuine philosophers, friends (*philo*) of wisdom (*sophia*) as truth-seekers, with the sophists, who gave paid lessons for political power and economic prosperity. He was sentenced to death for "the corruption of the youth", but – via Plato's dialogues – succeeded in formulating a still practiced model for knowledge-seeking and learning by questioning.[1]

Socrates' student Plato (ca. 426 - 347 BCE), wrote philosophical dialogues with Socrates as the main hero – on the definition of love, knowledge, justice, and other important themes. He founded a school, the *Academy*, in Athens in 387 BCE. Plato's classical conception of knowledge (Gr. *episteme*) distinguishes justified true beliefs from false errors and unfounded opinions. The Academy promoted the cultivation of the rational powers of thinking by the study of geometry and philosophy (in addition to some gymnastics). The students were young men, but in his ideal state Plato allowed education for females as well. With some interruptions, the Academy survived in Athens for 900 hundred years (ruled mainly by skeptics and neo-Platonists) until it was closed as a pagan institution by the East-Roman emperor Justinian I in 529 CE.

Aristotle (384-322 BCE) studied in Plato's Academy for twenty years, and then founded his own school, the *Lyceum*, in Athens. He wrote classical philosophical texts in logic, scientific method, dialectic, rhetoric, poetic, metaphysics, physics, psychology, ethics, and politics. Aristotle's distinction between theoretical and practical philosophy is still followed in many academic institutions today.

The Greek educational tradition distinguished manual skills (mostly practiced by slaves, later by farmers and craftsmen) from the skills needed by free men, who had leisure time

[1] See Hintikka's (2007) "Socratic epistemology".

for leaning (Gr. *schole*; Lat. *otium*). Skill in Greek is *techne*, in Latin *art*, in plural *artes*. Later Hellenistic time consolidated the system of *artes liberales* as the basis of education: grammar (rules of language), dialectic (art of disputation), rhetoric (art of speech), geometry, arithmetic, astronomy, and music.

The ancient philosophical and scientific tradition was preserved by the Arabic culture, whose oldest institutes of learning were established in Fez (in Morocco) in 859 CE and Cairo (in Egypt) in 970 CE.[2] The Greek classics were transmitted to the West by Latin translations in the 11th century onwards. Medieval scholasticism emphasized the importance of logic as a tool of thinking. A synthesis of Aristotle's philosophy and Christian theology by achieved by Thomas Aquinas (1225-74).

The First Universities

In the medieval Christian tradition, learning and education was the province of cathedral and monastic schools. The first institutes of higher education were established by students (*studium generale*), kings, and cities. *Universitas* was understood as *alma mater* for students and as a place where teachers from several faculties meet each other as a whole. The first such universities were founded in Bologna (1088), Oxford (1096), Salamanca (1134), Paris (1150), Cambridge (1209), Padua (1222), and Naples (1224).[3] Many universities were connected to the churches (e.g. Catholic and Pontifical universities). This was true of the first colonial universities in Hispanic America (San Marcos in Lima, Peru, 1551; Santo Tomas de Aquino in Dominican, 1558) and the first university in Asia (Santo Tomas in Manila, Philippines, 1611). The German university of Wittenberg became in the 16th century the center of Lutheran reformation. The oldest university in North America was Harvard (1636). Latecomers among continents are sub-Saharan Africa (Sierra Leone, 1827) and Australia (Sydney, 1850).

In Finland, the Queen Christina of Sweden signed in 1640 the charter for The Royal Academy of Turku.[4] Following the

---

[2] Some historians classify al-Qarawiyyin in Fez, founded by a female scholar Fatima al-Fihri, as the oldest university in the world.

[3] For medieval science and education, see Lindberg (1978).

[4] See Klinge (2010).

model of Paris and Uppsala, it gave professional education in three higher faculties: theology (priests), jurisprudence (lawyers), and medicine (physicians). Propaedeutic studies were given in the Faculty of Arts, including theoretical philosophy (metaphysics and the art of thinking), practical philosophy (ethics, politics, and history), classical languages, eloquence, mathematics, and physics. The task of the university was to compile and transmit existing knowledge, so that its teaching emphasized critical thinking: in disputations, the professor formulated the thesis and the student defended it.

The distribution of learning was enhanced by Johann Gutenberg's printing press (ca. 1450). The Renaissance stimulated work in classical humanities and arts. The universities were mainly conservative in the slow acceptance of the new theories of the scientific revolution (Copernicus, Kepler, Galileo, Newton). It was only the early 17th century that Francis Bacon and René Descartes convinced their contemporaries that the main task of science is to seek new knowledge by research. The scientific community was born in the 1665 with the establishment of the Royal Society in London, with scientific journals using peer review as a quality control.

The 18th century Enlightenment demanded education for all (Comenius),[5] and the collection of knowledge to encyclopedias (Diderot). Philosophers as free intellectuals (Voltaire, Rousseau) campaigned for freedom of thinking, against religious and political authorities. Immanuel Kant, professor at the University of Konigsberg, defined the spirit of Enlightenment by the advice *sapere aude* ("dare to think for yourself").

A tension within the university system was created when the technological revolution in the late 18th century made visible the demand of developing higher education also for civil engineers. Among the first such institutes were École Polytechique in Paris (1794) and the Massachusetts Institute of

[5] Comenius belonged to the progressive thinkers who planned schools both for boys and girls. When elementary and secondary schools started to open for girls in the latter part of the 19th century, many universities accepted female students by special permission. In the University of Helsinki, Emma Irene Åström graduated as the first female Master of Arts in 1882. Today the majority of students and graduates up to the doctor's level are female.

Technology (M.I.T.) in Boston (1856). Many of them were upgraded to universities in the early 20th century.

In the French model, supported by emperor Napoleon around 1800, emphasis was on professional education, so that the universities would become vocational schools with strict and narrow disciplines, while research would be concentrated in separate institutes.[6] Two important university ideals against the French model were developed in Germany (Humboldt, 1809) and Great Britain (Newman, 1852).

Cardinal Newman's Idea of a University

John Henry Newman (1801-90) was a British educationalist. After studies at Oxford he became a Catholic priest in Rome in 1847. He was the first Rector of the Catholic University in Dublin, Ireland, in 1854-58 (today University College Dublin). In 1879 Newman was appointed as Cardinal of the Catholic Church.

Newman gave in 1852 his first ten lectures *Discourses on the Scope and Nature of University Education*, and in 1858 lectures on *University Subjects.* In 1873 his classical work *The Idea of a University: Defined and Illustrated* was published as a single volume with two parts.[7] It considers the essence of a university "independently of its relation to the Church", but still Christianity has an important status among university subjects, as "the Church is necessary for its *integrity*" (p. ix). In a university all branches of knowledge are connected (unity of the subject matter).

Newman defends Cicero's ideal of liberal knowledge (for its own sake) against Francis Bacon's conception of knowledge useful for professional skills (p. 151-153). He summarizes the thesis about "universal teaching" in the following way:

---

[6] This plan was realized in the late 20th century, when professional schools in many countries were upgraded to "universities of applied science" (without the right to educate doctors). In England, "polytechnics" were upgraded to the university status in 1992.

[7] See Newman (2009). I am grateful to Professor Dermot Moran who invited me to Dublin in 2010 and gave me a copy of Newman's work, published by the University College of Dublin International Centre for Newman Studies.

> The view taken of a University in these Discourses is the following: - That it is a place of *teaching* universal knowledge. This implies that its object is, on the one hand, intellectual, not moral; and, on the other, that it is the diffusion and extension of knowledge rather than its advancement. If its object were scientific and philosophical discovery, I do not see why a University should have students; if religious training, I do not see how it can be the seat of literature and science. (p. ix)

The aim of the university is "real cultivation of mind" - not merely the moral and intellectual manners and habits of English gentlemen (p. xi), but "the force, the steadiness, the comprehensiveness and the versatility of intellect" (p. xvi). The basic values include "the good sense, sobriety of thought, reasonableness, candor, self-command, and steadiness of view", and the mind is "brought into form" by "the influence of Truth" (p. xviii). Subjects like grammar, mathematics, chronology, geography, history, and poetry make the students "more intelligent, capable, active members of society" (p. xix).

About the relations of teaching and research Newman's position is clear: "there are other institutions far more suited to act as instruments of stimulating philosophical inquiry, and extending the boundaries of our knowledge, than a University" (p. xii). Such institutions include Academies and the Royal Society. So there is reason

> … to recommend us this division of intellectual labour between Academies and Universities. To discover and to teach are distinct functions; they are also distinct gifts, and are not commonly found united in the same person. He, too, who spends his days in dispensing his existing knowledge to all comers is unlikely to have either leisure or energy to acquire new. (p. xii)

Humboldt's Model of a Bildungsuniversität

Baron Wilhelm von Humboldt (1767-1837) was a German philosopher, linguist, and diplomat. Besides his early studies in philology and aesthetics (as a friend of Schiller and Goethe), he participated in discussion about educational reforms in a treatise on "the limits of state action" in 1790. From a diplomatic position in Rome, he was invited by the king of Prussia Friedrich Wilhelm to serve in the minister of education in 1809-10.

Prussia and its capital Berlin were recovering from Napoleon's wars, which posed a challenge to Humboldt to prepare in 1809 a memorandum on "the internal and external organization of higher scientific institutes". This plan led to the founding of the *Friedrich-Wilhelms-Universität zu Berlin* in 1810 (since 1949 *Humboldt University* in Berlin).

Humboldt's model of higher education highlights the difference between schools and universities: in the latter knowledge is searched for its own sake, which presupposes "solitude and freedom". The basic principle is *Bildung durch Wissenschaft:* the education, edification, or cultivation of individual subjective mind by means of objective science.[8] Such general education realizes the unity of research and teaching: in contrast to Newman's later view, the Humboldtian ideal formulates the principle of research-based education. In contrast to the French system of *grandes ecolés* with severe disciplinary curriculum, self-education should proceed without professional goals, so that the students will be educated as autonomous individuals and world citizens. The state should guarantee the freedom of academic work in teaching and learning (i.e. *Lehrfreiheit und Lernfreiheit*) and take care of the salaries, but otherwise a university has a corporate autonomic governance.

In practice, the statutes of the Berlin University 1816 stated that it gives "general and specific scientific education", which allowed some elements vocational training. Freedom from censorship was guaranteed, but the state appointed professors, and the internal governance was organized by ordinary professors.

In sum, both Newman and Humboldt advocated the intrinsic value of knowledge and opposed the French model of transforming universities to institutes of vocational training. But while Humboldt's model in 1810 praises research-based education, Newman (without reference to Humboldt) in 1873 proposed to delegate inquiry to other institutions than

---

[8] The original medieval connotation of the German term *Bildung* refers to the Christian idea of *Imago Dei* (*Genesis* 1:27)*:* human beings become more and more images of God through education. There is no good English translation, but alternatives include education in the broad sense, edification, cultivation, and civilization.

universities. Temporally between them, in 1840, is the account of academic studies by the Finnish philosopher J. V. Snellman.[9] It is still largely unknown in international discussions about the idea of a university, but in my view offers a definite improvement of Humboldt's model.

Snellman on Academic Freedom

In 1809 Finland was separated from Sweden to become an autonomous Grand Duchy of the Russian empire (until its independence in 1917). In 1828 the Academy of Turku was moved to the new capital Helsinki as the University of Helsinki. It was directly subordinated to the Emperor's "special protection", but the new statutes defined the task of the university in the Humboldtian spirit as "to promote the development of sciences and the liberal arts in Finland, and to educate the youth of the country in the service of the Emperor and the fatherland".[10] The Faculty of Philosophy, now allowed to grant doctor's degrees, moves to the center of the University, which in the mid-19th century helped to build Finland as a nation with its own history, culture, language, folklore, literature, and music. The key figure of this national movement was Johan Vilhelm Snellman (1806-1883), Hegelian philosopher, professor of philosophy, and senator.

In 1837 young docent J. V. Snellman proposed to give lectures on "the true meaning and essence of academic freedom". The Rector feared that the controversial topic might encourage student unrest and declined Snellman's right to give these lectures. After a clash with conservative university authorities the contentious philosopher started a journey, via Sweden, to Germany, where he wrote in Tübingen his main work on "the idea of personality". In the fall of 1840, Snellman published in Stockholm his essay *Om det akademiska stadium*, a revised version of his forbidden 1837 lectures.

The essay "On Academic Studies" is powerful defense of the Humboldtian ideal. Snellman starts by distinguishing *memory knowledge*, where a content is passively preserved without thinking, and *grasping knowledge*, where the subject actively adopts and understands the content by his own concepts

[9] See Manninen et al. (2020), 79-84, 90-91.
[10] Klinge (2010), 290.

and abilities. Memory knowledge, or knowledge contents as such, constitutes the tradition which is independent of the conscious subject, while its opposite is *creative knowledge* where the new content is freely formed by a self-conscious subject.

According to Snellman, *school* is an institution where an individual is educated to self-consciousness, to a thinking and willing subject. Its teaching is wholly based on memory knowledge and learning exercises by heart, and the teacher is an authority to his students. *University*, on the other hand, is an institution where a thinking and willing subject is educated to "a reconciliation of self-consciousness and tradition". The aim of a university as a civilizing institution is *Bildung*, which includes education both in knowledge and morality.[11] When a self-conscious young man enters the university, he wants to decide for himself all matters of knowledge and conduct. This abstract self-consciousness can be reconciled with tradition only by recognizing its rights. Academic freedom consists in this recognition of the right and obligation of self-consciousness. Its basic principle can be expressed by the Biblical words: "Try all things and keep that which is good."

Snellman's solution to the reconciliation of self-consciousness and tradition – or his model of critical thinking and learning - is based upon the Humboldtian idea of research-based teaching and learning:

> The university teacher in his lectures should always prove that his knowledge is based upon his own research. He should not present only ready results, but lead his students to the road that he self has wandered. Thereby the student sees the recognition of self-consciousness in the independence of the teacher, since from the teacher's right to inquire and judge he can infer that he himself has the same right.

In this way, the academic student replaces memory-based authority by true knowledge and understands what is rational in tradition. He is not only a schoolboy any more, but has searched for a conviction by solving independently some problem of science or life. Parallel recognition of the academic

---

[11] Snellman's Swedish word for *Bildung* is *bildning* (in Finnish, "sivistys"). For the concept of morality or ethical life, Snellman used the term *sedlighet*, which is a translation of Hegel's *Sittlichkeit*.

freedom of students in the area of their conduct is based upon their autonomous regional associations like the nations where they learn to relate their own subjective will to the equal rights of their fellows.

While students are still juvenile in the state, they are declared as "moral citizens" when they graduate from the university. By learning the method of creating new knowledge the graduates have gained permanent love for the truth and willingness to use their capacities for the benefit of society and the state as a rational and ethical order. In this way, by generating the process of civilization by means of academic freedom, the university has achieved its cognitive and moral goal:

> Education does not end with school or some specific exam, but the whole life is a school, where an individual is transformed into a human being, and this process of *Bildung* itself constitutes what is humane in a human being.

With these words Snellman formulated his conception of humanity and what is later known as the idea of lifelong learning.

Snellman improves Humboldt by showing in detail how education can be based on research: the student has to learn the method of scientific thinking from his teachers. In comparison to Humboldt, Snellman was more outspoken about the graduated student's responsibilities for the service of the state and humanity. Using later metaphors, university is not an "Ivory Tower" but rather a "light house" influencing the surrounding society.

Snellman returned to the nature of academic studies in his lectures as professor of philosophy in 1856. This time he cites Fichte's statement that university is "the most important and holy institution of mankind". Snellman defends the freedom of teaching and learning. He repeats his own thesis that the task of the university is to reconcile tradition and self-consciousness, but emphasizes that the development and creative renewal of the tradition is a function of science. Academic freedom of teaching and learning should be respected so that choices are not "sinking to tradition". Indeed, tradition should be presented to new generations in novel forms, by learning from "great independent minds" instead of those who just "lag behind". In the same way, morality consists in voluntary

action in agreement with the rational state, but "the prevailing custom is not rational just by its existence but by its development". Thus, academic studies invite students "not only as persons who preserve customs but also develop them".

Snellman's version of the Humboltdtian university is in harmony with new theories of learning in psychology and education, which have changed the focus from teaching "from above" (i.e. pedagogic) to active and creative learning by the student. Examples of such approaches include John Dewey's (1916) "learning by doing", constructivist learning theories by psychologists and cognitive scientists, and "problem-based learning" (PBL). The problematic relativism of some constructivist theories of learning is avoided, since Snellman's account includes a respect for truth as given by the "tradition". This is in harmony with the contemporary idea of critical thinking skills and "learning to learn": principles of logical and scientific reasoning are more important than constantly changing factual knowledge.[12] In particular, Jaakko Hintikka's (2009) interrogative model of inquiry returns to the Socratic method of systematic questioning.

Research Universities

Newman's idea of a university lives in the American tradition of liberal arts colleges. The Humboldtdian tradition is continued in the concept of a *research university* (or research-intensive university), which takes scientific research to be its basic function and teaching is based upon research. Especially the modern research universities are central places for critical thinking among the staff and students - today giving equal opportunities to both genders.

Today there are about 17 000 universities and institutes of higher learning – with a heavy competition. Research universities with excellent doctoral training have a good chance of being at the top of the university rankings. The "Shanghai" ranking ARWU, designed by Jiao Tong University, measures the success of universities by their research activities: Nobel prizes and Fields medals, publications in *Nature* and *Science*,

[12] See Salmon (2013).

number of indexed publications, and highly-citied researchers. The ten best universities in the world in 2021 are

1. Harvard, 2. Stanford, 3. Cambridge, 4. MIT, 5. UC Berkeley, 6. Princeton, 7. Oxford, 8. Columbia (New York), 9. Caltech, 10. Chicago.

The best Asian universities are Tokyo (24th) and Tsinghua (28th). Helsinki is in the 82th place, so that it ranks among the top of 0.5 per cent of the universities.

The World University Ranking by Times Higher Education (THE) takes into account research, teaching, and impact (including technology transfer, international outlook, and reputation). Top ten includes Oxford, CalTech, Harvard, Stanford, Cambridge, MIT, Princeton, Berkeley, Yale, and Chicago. With one exception, these are the same as in the ARWU Ranking.

In the USA, eight top research universities constitute the Ivy League. In 2002, twelve leading multi-faculty universities established a network LERU (the League of European Research Universities), among them Oxford, Cambridge, Edinburgh, Milan, Geneva, Leiden, Leuwen, Heidelberg, Munich, Strasbourg, Karolinska institutet, and Helsinki. Today LERU has 23 members.

The main rival to the research universities is the model of *entrepreneurial university*, which may refer to an institution using practices from business corporations in its governance or producing useful knowledge and innovations to the market by knowledge and technology transfer. Some of these institutions are private, but still not-for-profit.

This distinction is not absolute, since to some extent the models of research university and entrepreneurial university can be combined or reconciled. Even state-funded universities may charge tuition fees, and gain external funding from public and private sources. A research university can be effective, and a commercial university can use research-based teaching in the education of innovative talents.

A more flexible four-fold classification can be based on two axes:

- TEMPLE: traditional university with strict internal hierarchies (power of professors) and autonomy with respect to the church, state, and economy

- OASIS: autonomous university with academic freedom and democratic internal structures (student power)
- FACTORY: effective disciplined production of useful knowledge and academic degrees (masters and doctors)
- BAZAAR: a lively market-place of commercial activities and functions (adult education, technology villages, startup companies).

Universities have survived for 900 years by adapting to different demands and circumstances. Even though they may dynamically adopt features from all the four models, a research university (à la Humboldt and Snellman) has passed the test of time: in a good successful university, research and education are in mutual interaction with other.

Note. I have written about the Humboldtian university in several essays of my collection *Dynaaminen sivistysyliopisto: Sata puhetta ja kirjoitusta vuosilta 1987-2010*, Helsinki: Gaudeamus, 2011. This unpublished chapter is based on a lecture in Sichuan University, Chengdu, China, on November 21, 2016.

Literature

Dewey, John (1916). *Democracy and Education*. New York: Macmillan.

Hintikka, Jaakko (1999). *Inquiry as Inquiry: Toward a Logic of Scientific Discovery* (Selected Papers V). Dordrecht: Kluwer.

Hintikka, Jaakko (2007). *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning*. Cambridge: Cambridge University Press.

Lindberg, David (1978). *Science in the Middle Ages*. Chicago: The University of Chicago Press.

Manninen, Juha, Ihanus, Juhani, Jalava, Marja & Niiniluoto, Ilkka (2020). *A History of Philosophy and Psychology in Finland, 1809-1917*. Helsinki: Societas Scientiarum Fennica.

Newman, John Henry (2009). *The Idea of a University*. Dublin: Ashfield.

Salmon, Merrilee (2013). *Introduction to Logic and Critical Thinking*. Boston: Cengage Learning.

Chapter 14: Europe and Scientific Education

What is Europe?

There is a well-known story about a group of Europeans who went to a zoo and saw there an elephant. The Englishman asked: "How does this thing function?". The German inquired: "How much does it cost?". The Italian queried: "What does it eat?". The Frenchman thought: "How does this creature make love?". And, finally, the Finn wondered: "What does it think of me?".

The point of this story is evident. Among all the European people, there are important and interesting differences, or at least stereotypical beliefs and expectations about differences, which influence the forms and prospects of our co-operation. For example, substitute the European Union for the elephant in the story, and you gain some insight about the attitudes towards the EU among the practical British, economic Germans, culinarist Italians, cultural French, and shy Finns.

More seriously, we have to recognize that there are enormous differences among all European nations, their historical background, cultural traditions, languages, religions, political systems, economies, and values. For example, think of the split between the Roman Catholic and Greek Orthodox churches, and the struggle between Catholic and Protestant churches. Or think of the diversity in democratic traditions, liberal market economy, social welfare systems, folklore, art, and relations between sexes. Remember the hundreds of armed conflicts between the European nations, culminating in two World Wars in the twentieth century. Such differences seem to divide Europe in many (perhaps interrelated) ways. So is Europe simply a conventional geographical category,[1] a continent surrounded by three big seas and the Ural mountains? Does the "real"

[1] This was bluntly claimed by Otto von Bismarck in 1879.

Europe extend beyond the EU?[2] Is there any unity or identity in Europe?[3] Or is Europe only an artificial and arbitrary social construction? Is the attempt to find some identity an expression of the malady of "eurocentrism"?[4]

In my view, if we try to find some important common aspects of Europe, we should not search for existing things on the surface of our culture. Rather, we ought to penetrate deeper to the forces, ideals, and values that have transformed our history and shaped us and our present condition.

Scientific Education as the "Essence" of Europe

It is usually said that the European culture has three roots: the Judaeo-Christian religion, Greek philosophy, and Roman law. Our values of human worth, love, charity, truth, goodness, beauty, and justice go back to this ancient Mediterranean tradition (partly transmitted to us by the medieval Islamic culture). While the present European nations still carry memories from their prehistorical past (e.g. oral epic tradition, "pagan" religions), there was not yet Europe in the pre-modern time.[5] When the building of Europe was carried on by integrating the southern and the northern areas, the philosophical and political ideals of education, etiquette, moral virtues, bravery, industriousness, liberty, fraternity, equality, democracy, and human rights were developed and mixed with each other.[6]

The European culture was transported to all over the world, especially to America. The moral ideals failed miserably in the greedy colonial exploitation, but also bore fruit in the North America.[7] In the old continent, we have also frequently failed

---

[2] So, for example, is Putin's Russia excluding itself from Europe by its military attack against Ukraine and hostility to the EU? Is United Kingdom still a part of Europe after the Brexit?

[3] See Niiniluoto & Löppönen (1996).

[4] For a critical account of euro-universalism and eurocentrism, see Weller (2021).

[5] In his history of modern culture, Egon Friedell placed the birth of Europe to the Black Plague in the mid-fourteenth century.

[6] For the historical emergence of Europe, see Mikkeli (1998) and Pagden (2002).

[7] The French postmodernist Jean Baudrillard, in his book America (1989), ironically claims that the inability of the Americans to understand history

to live up to our ideals, but they have never lost their importance for our enlightened conception of good human life.

Hope for peace after the catastrophic World War II inspired new attempts at integration: the Council of Europe in 1949, the European Coal and Steel Community in 1952, the European Economic Community (EEC) in 1957, and the Maastricht Agreement of the European Union (EU) in 1992. Even though a common free economic market was seen as a catalyst of peaceful relations, the EU was clearly intended to be a value community which promotes welfare, democracy, human rights, and social justice. As a NGO, established in 1954 by active citizens to add "cultural and human dimensions" to more narrow economic interests, one may mention the European Cultural Foundation.

The European (or, more broadly, Western) ideal of a human being can be expressed by a number of pairs where the first member has a preferential position: rationality vs. emotionality, free vs. constrained, activity vs. passivity, optimism vs. pessimism, learning vs. ignorance, individualism vs. collectivism, organization vs. chaos.[8]

This list suggests that the "essence" of Europe lies in its tradition of philosophical and scientific education.[9] Here I am using the term "science" in the broad sense of *Wissenschaft* which ranges from the natural sciences to the social sciences and the humanities. I am putting "essence" in quotation marks, since I am not an essentialist. As cultures are in mutual interaction by the exchange of goods and ideas (i.e. cultural appropriation in a positive non-pejorative sense), no continent can be expected to have an "essence" which would be a *differentia specifica* in

---

is due to the fact that America is a "realized utopia", with all dreams come true.

[8] This somewhat schematic list the Enlightenment values was given in 1926 by Eino Kaila, who was then attracted by the Paneuropean movement of count Coudenhove-Kalergi. Kaila was aware that Arthur Schopenhauer's philosophical pessimism had affinities with Indian religions.

[9] The phenomenologist Edmund Husserl characterized in 1935 "the concept of Europe as the historical teleology of the infinite goals of reason" (see Miettinen, 2013). The philosopher Karl Jaspers suggested in 1947 that the most essential features of Europe are freedom, history, and science (see Mikkeli, 1998). For the conception of Europe within philosophy, see Meachen & de Warren (2021).

the Aristotelian meaning. In particular, given the early and contemporary traditions of science in other continents, it would be wrong to maintain that science is exclusively a European phenomenon. Yet, our attempt to define Europe must refer to philosophy and science in some way or another.

In the early medieval schools and universities (the first of them established in Bologna in 1089), the Latin education combined Greek learning (especially logic and philosophy) with Christian religion. It achieved the first real integration of Europe: even the Finnish students from the ultimate North travelled to the University of Paris in the early 14th century. The program of scientific education was further continued by the "modern project" with the scientific revolution, academies and scientific societies, philosophical enlightenment, combination of research and teaching in the universities ("*Bildung durch Wissenschaft*"), public elementary and secondary schools, professional polytechnics, laboratories of applied science, national programs of science policy, new tools of mass communication, open university and adult education. The new networks of education and research and the exchange programs for university students (Erasmus, Leonardo, Socrates) are clearly extensions of this European ideology of scientific education.

The idea of a university was successfully "exported" to other continents as well (cf. Chapter 13 above). So the Europeans are not the only people who have access to scientific education. But it is significant for the identity of European citizens that the curriculum of elementary and secondary school is primarily secular - and therefore updated by the progressing results of science.

It is important to see that we all have multi-layered identities: besides European, I am from Helsinki, a white Finn male, a football fan, a philosopher as well as a world citizen. As a member of the scientific community, I am committed to the universality of truth, since the truth of indicative assertions like laws of nature is not relative to cultural differences (see Chapter 16 below). In this respect, scientists share cosmopolitan values.[10] As a modest ethical relativist, I do not believe that

[10] The first program of cosmopolitanism with universal values like human dignity was developed by Stoic philosophers (see Sihvola, 2004). The European Enlightenment can be viewed as a continuation of this tradition.

moral values can be justified in the same universal ways as scientific truth, but still I am committed to the ideals of human dignity, justice, equality, and freedom. Richard Rorty (1991, 207) argues in his "anti-anti-ethnocentrism" that such commitments are entirely contingent on the fact that we happen to be Western liberals who believe in such values. But here Rorty ignores the fact these liberal values are expressed by the Universal Declaration of Human Rights, approved by the member states of the United Nations in 1948, and moreover codified by agreement and principles of international law.[11] Thus, Rorty's (1989, xvi) disjunction between "contingency and solidarity", or between "an endless proliferating realization of Freedom, rather than a convergence toward an already existing Truth" is a false opposition: freedom and truth do not exclude each other, and we need them both! This also means that instead of local competition and power play, international co-operation is needed so that social constructions like the EU and the United Nations are stepping-stones to global governance with cosmopolitan values.

Science and the Modern Project

It is clear that scientific research and education can flourish only in a cultural and social environment which pays respect to knowledge and civilization, individual rights, freedom of thinking, educational equality among citizens, creativity and talent, peace, democracy, and human rights. These are values that have been shared in principle, even though not always in practice, in Europe and the North America.

On the other hand, science will not have the necessary intellectual autonomy in totalitarian societies (the Lysenko affair in Stalin's Soviet Union; persecution of Jewish scientists and their exile from the Third Reich; theocratic states). Cognitive progress is fostered by the strong demand for dynamic change and reliable new information in the society.

If there is a difference between Europe and other continents, it could arise primarily from the European sense of history and

[11] Building on the UN Declaration of Human Rights (1948), the International Covenant of Economic, Social and Cultural Rights defended in 1966 the universal "right to enjoy the benefits of scientific progress and its applications".

its long tradition in linking progress with scientific education. This key idea of the Enlightenment (cf. Chapter 15 below) should also give us the basis for understanding the intrinsic value of knowledge (i.e. the worth of basic research which aims at informative truth independently of useful applications).

Economic and scientific progress have supported each other in a positive cycle. The link between science and protestant utilitarian commitment to work was strong already in the 17th century Europe, as Robert Merton (1970) has argued in detail, and the exploitation of research as a source of technological and social progress was strongly emphasized both by capitalism and socialism. The utilitarian approach to science as a tool of economic competition has so far been stressed almost exclusively by the research activities of the EU – with the exception of ERC grants for high-quality basic research. Even though material resources are necessary for science, other cultural activities, and human welfare, one-sided emphasis on short-term applicability of knowledge and knowledge-based technologies has also led to harmful environmental and social consequences. As the "dialectic of enlightenment" has tended to truncate the domain of human reason to instrumental rationality (cf. Chapter 8 above), the legitimacy of all science and the whole modern project have been questioned.

These observations raise serious problems concerning the future of science. Today it is often claimed that the modern project has come to its end, optimism and progress have collapsed, the value of novelty and excess of old limits has withered away, communication has split into incommensurable language games, authentic human subjects have died, and the world has become a place for competition between disillusioned players. In such a postmodern condition, it is claimed, there is not room any more for the old European ideals and for philosophy, science, art, and morality in the sense we have interpreted them.

So is there still a future for philosophy and science? Given the historical connections emphasized above, this leads immediately to another question: Is there a future for Europe?

My short answer to these pressing questions is the following. I think it is correct to say that we have come to the end of an era - the future historians will probably state that the

Modern Age lasted to the end of the 20th century, but we don't yet understand this change or know the title and nature of the next historical period. It is crucially important that the Western civilization learns to adapt to the conditions posed by the health of our natural environment and by sustainable social development. Economy should serve good life – instead of being the supreme goal. This process will desperately need scientific knowledge - especially from biological, medical, human, and social research. The urge for scientific inquiry and education is by no means diminished in the future. Moreover, the prospects for finding new important knowledge have not disappeared. The alleged crisis of human rationality is not primarily a problem concerning cognitive rationality and the objective methods of truth-seeking in science, but rather a crisis in the narrowly egoistic and instrumental conception of human values that has spread over almost all areas of our culture.

The Relation between Philosophy and Science

The academies and universities were created to promote philosophy and science. In the medieval conception of education, philosophy (especially logic) had a central methodical or propaedeutic function of teaching the art of clear thinking and thereby serving the "higher" professional disciplines (theology, medicine, law). But in the nineteenth century Humboldtian university, philosophy became the "highest" of all subjects (cf. Chapter 13 above).

Today we often hear the claim that science does not need philosophy as its foundation. This thesis has been supported in two radically different ways. The "positivist" view urges that science may be a child of philosophy, but has then grown completely independent of her mother, i.e. mature science has happily got rid of metaphysics and epistemology. The "postmodern" view of Richard Rorty (1980) asserts against the "Kantians" that nothing has foundations; hence, science in particular has no foundations, either. Both views seem to imply that there is no special task for a philosophy of science: science studies simply collapse into historical and sociological description. For the positivist, this is motivated by the belief that science, as it is, is the paradigm of human rationality. For the

postmodern thinker, on the other hand, there is no interesting account of rationality to be found anywhere.

I think both of these extremes are wrong. Science as a rational cognitive enterprise is not yet complete: its tentative results are always corrigible and in need of analysis and interpretation, and its methods can still be improved in their reliability and effectivity. Also the ethics of science has to be developed as a part of the philosophical conversation about the social role of scientific practices. Philosophy of science cannot give any absolute and final foundation for science, but it cannot leave science as it is. There is a legitimate need to raise normative questions about scientific inquiry and knowledge, to set up standards, and (if necessary) also to criticize the activities of science. To be sure, such pronouncements are fallible and cannot be exclaimed from the armchair: philosophy of science and special sciences have to be able to engage themselves in a mutual dialogue.

The Faustian Western Culture

Oswald Spengler's *Der Untergang des Abendlandes* (vol. I, 1918; vol. II, 1922) presented a grandiose account of world cultures as organisms, whose development from birth, bloom, and withering follows fateful regularities. According to the main thesis, the "Faustian" Western culture, which around the 11th century followed the "Apollonian" Ancient (Greek and Roman) and the "Magian" Arabian cultures, has already come to its winter and is destined to decline.[12] The "ursymbol" of the West is clock, which measures linearly progressing time, and its spirit yearns towards freedom and infinite space – as evidenced in Gothic cathedrals, crusades and voyages of discovery, monotheistic religion, Bach's music, advances in technology, and exact science. The "sunset" or decay of this cultural form can be seen in skepticism, materialism, democracy, socialism, metropolises, and sport.[13]

Even though Spengler's romantic philosophy of history is laden with many methodological problems – e.g. the historicist

[12] As an update of Spenglerian themes, the cultural historian Joseph Tainter (1988) presents a pessimistic view of complex cultures, which have collapsed when their problem-solving fails with diminishing returns.

[13] See Spengler (1991).

assumption of "morphological" laws with "fate" replacing causality – it is interesting to compare his anti-modernism with our characterization of Europe.

One of Spengler's claims is that "Faustian" natural science will eventually be exhausted and returns as "tired" to its "spiritual home": "in this very century, I prophecy, … exact science must presently fall upon its own keen sword". In our view, this would be fatal to Europe. But as a prediction, Spengler's forecast seems to fail badly. Today, after exponential growth in modern times, there are more scientific disciplines and researchers than ever in the history. It can be argued that the natural and social worlds are structurally infinitely deep and complex, so that the number of potential research problems about this "endless frontier" is limitless – and solutions of cognitive problems always generate new problems for research.[14] This has been well expressed by the slogan: "The larger the sea of knowledge, the longer the coastline of ignorance". Nicholas Rescher (1978), however, points out scientific progress may slowly decelerate, since new problems become harder than earlier ones (Planck's Principle of Increasing Effort) and the economic cost of reaching first-rate and important results may be too high.[15]

It may seem odd that Spengler named the Western culture with a man, a necromancer and astrologer, who was so eager about magic that the legend told about his pact with the devil to gain knowledge. Even in Christopher Marlowe's play in 1604 Dr. Faustus loses his soul and has to face eternal damnation. The enlightened author Lessing in 1780 allowed salvation for Faust, as pursuit of knowledge is a noble activity. The most sophisticated version of Faust – the one behind Spengler's view – was Johann Wolfgang Goethe's drama *Faust* (part I, 1808; part II, 1832). It has a happy ending with Faust's

---

[14] For arguments why science can grow forever, see Niiniluoto (1984), Ch. 5.

[15] An example of rapidly rising costs is the Large Hadron Collider (LHC) which was used in 2012 to solve the hypothesis about the existence of the Higgs boson. But the construction of this huge and expensive particle accelerator, with a 27-kilometer ring of superconductive magnets, in the European Organization of Nuclear Research (CERN) in Switzerland, is a prime example of successful co-operation of 27 European countries in basic research.

redemption, but the crux of the story is the wager between God and Mephistopheles: as complete knowledge is not attainable, Faust will perish if he ever ceases his continuous striving in "a hollow conceit". Faust seals this convention with the Devil by the words:

> If to the fleeting hour I say
>
> ´Remain, so fair thou art, remain!´
>
> Then bind me with your fatal chain,
>
> For I will perish in that day.

According to von Wright (1993, 150), Goethe's lesson is that the power provided by human knowledge is not as such evil – but it becomes an evil, if "man in his delight stops to enjoy the fruits of his work without realizing its incompleteness or feeling the yearning for something better".

The Faustian culture is thus an expression of "the dialectic of the unattainable".[16] The fallibilist Charles S. Peirce argued that self-corrective science indefinitely approaches the truth in the long run – and added that science "does not consist so much in knowing, nor even in organized knowledge", but rather in the pursuit of finding out.[17] So for a critical scientific realist, science is indeed a "Faustian" enterprise in truth approximation.

If scientific research and education are "essential" to Europe, as we have suggested, Spengler's characterization of the soul of the Western culture as Faustian is quite fitting. Otto Neurath, later an activist in the Vienna Circle and a champion the scientific world outlook, suggested in his *Anti-Spengler* in 1921 a famous metaphor: "we are like sailors who on the open sea must reconstruct their ship but are never able to start afresh from the bottom". As an anti-foundational supporter of the coherence theory of truth (cf. Chapter 16 below), Neurath asserted that the boat can be repaired only by its own beams and driftwood, so that visiting a harbor or dock is forbidden, but a realist – continuing the metaphor - would stress the importance of interaction and investigation of the sea and climate

---

[16] This term has been used by Jari Ehrnrooth.
[17] See Niiniluoto (1984), 42.

in the boat's environment. In any case, for us the important point of this simile is never ending movement by a boat (or even spaceship): *navigare necesse est* – keep moving, don't stop![18] In the same way as science, Europe as a Faustian project is always unfinished.[19] But sailing is not random drifting in the wind, as it may have a direction, insofar as navigation is based on enlightened values and research-based education.

Note. This chapter is based on a paper presented in the workshop on science, philosophy, and history of science in Europe in Paris, December 9-10, 1994. It has been published as "Europe, Scientific Education, and Modernity", in Dominique Lecourt (ed.) (1998), *Science, philosophie et histoire des sciences en Europe*, European Science and Technology Forum, Luxembourg: European Commission, 97-101. The last section is based on my article "Mitä on olla eurooppalainen?", in Niiniluoto & Löppönen (1996), 19-67.

Literature

Meachen, Darian & de Warren, Nicolas (eds.) (2021). *The Routledge Handbook of Philosophy and Europe*. Abingdon: Routledge.

Merton, Robert (1970). *Science, Technology and Society in Seventeenth-Century England*. New York: Harper & Row.

Miettinen, Timo (2013). *The Idea of Europe in Husserl's Phenomenology: A Study in Generativity and Historicity*. Helsinki: University of Helsinki.

Mikkeli, Heikki (1998). *Europe as an Idea and an Identity*. New York: St. Martin's.

Niiniluoto, Ilkka (1984). *Is Science Progressive?*, Dordrecht: D. Reidel.

Niiniluoto, Ilkka & Löppönen, Paavo (eds.) (1996). *Euroopan henkinen tila ja tulevaisuus*. Helsinki: WSOY.

---

[18] This demand should be understood as a directive for the scientific community. It does not imply that each scholar must at every moment maximize her efforts in the style of F. W. Taylor's sweaty system of scientific management. Progress in creative work needs a balance of activity and rest – or what the Stoic philosopher Seneca called otium sapientis (leisure of the wise).

[19] In contrast to the regular skyscrapers in New York and Shanghai, a moving symbol of Europe is the still unfinished church Sagrada Familia started by Antoni Gaudi a hundred years ago in Barcelona.

Pagden, Anthony (ed.) (2002). *The Idea of Europe: From Antiquity to the European Union*. Cambridge: Cambridge University Press.
Rescher, Nicholas (1978). *Scientific Progress: A Philosophical Essay on the Economics of Research in Natural Science*. Oxford: Blackwell.
Rorty, Richard (1980). *Philosophy and the Mirror of Nature*. Oxford: Blackwell.
Rorty, Richard (1989). *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.
Rorty, Richard (1991). *Objectivity, Relativism, and Truth*. Cambridge: Cambridge University Press.
Sihvola, Juha (2004). *Maailmankansalaisen etiikka*. Helsinki: Otava.
Spengler, Oswald (1991). *The Decline of the West*. New York: Oxford University Press.
Tainter, Joseph A. (1988). *The Collapse of Complex Societies*. Cambridge: Cambridge University Press.
Weller, Shane (2021). *The Idea of Europe: A Critical History*. Cambridge: Cambridge University Press.
von Wright, G. H. **(**1993). *The Tree of Knowledge and Other Essays*. Leiden: Brill.

Chapter 15: Progress by Enlightenment: Fact or Fiction?

The public use of one's reason must always be free, and it alone can bring about enlightenment among mankind.
– Immanuel Kant 1784

Enlightenment is a powerful intellectual and cultural movement which has vigorously promoted the idea of progress. However, the value-laden notion of progress is in many ways ambiguous even when it is restricted to some special area like science, art, technology, or politics. Some critics, who otherwise favor the Enlightenment commitment to education and research, have argued that progress is a myth. Some others regard the Enlightenment as a transitory period of modernity which has already been surpassed by disillusioned postmodernism. Considering the philosophical aspects of these debates, my melioristic conclusion in this chapter is that progress is still a viable possibility or prospect: even though there is no general historical law which would make progress of the humanity necessary, it is up to us to hold on the Enlightenment values in our strife for a better world.

History of Progress

In his classical study of the idea of progress, J. B. Bury (1932) argued that the conception of progress in human history was established only by the optimist thinkers of the Enlightenment. In the narrow sense, defended by Gay (1973), the Enlightenment is defined as an intellectual trend from the mid-17th century to the late 18th century. Using light as a metaphor,[1] its

[1] The terms *Aufklärung* (in German), *les Lumières* (in French), *upplysning* (in Swedish), and *valistus* (in Finnish) have a similar etymology.

advocates argued that education enlightens the dark world by giving freedom from superstition and religious dogmatism, if its contents are provided by scientific research. Freedom of thinking and political rights of citizens belonged to their progressive agenda.

A broader perspective on the history of progress is given Robert Nisbet (1980). The early modern age already knew many discoveries from the 15th and 16th century: printing, voyages to new continents, Copernican astronomy, Renaissance art, and reformation as a fight against clerical authority. They were continued in the early 17th century with the scientific revolution. These developments inspired many philosophers. Francis Bacon in *Novum Organum* in 1620 formulated a threefold optimism upon "the advancement of learning": the proper inductive method guarantees new scientific discoveries, the growth of knowledge gives us power to control nature, and the new powers "help to subdue the necessities and miseries of human life" (Bacon, 1960). Bacon became the hero of the Royal Society which founded the basic principles of the scientific community in the 1660s. Comenius wrote in 1650 his "pansophism" which demanded the collection of all human knowledge and thereby education for all. This program was continued by the French Encyclopedists in the next century (Diderot, Voltaire). G. W. Leibniz assured that, in spite of apparently bad events and evil things, we live in "the best possible world". This optimism culminated in Condorcet's vision of human progress as the completion of civilization, cognitive and moral perfection of man, while at the same time the French revolution championed the political slogan *liberté, eqalité, fraternité*.

The roots of the idea of progress and regress go back to prehistoric myths about paradises and golden ages. As Juha Sihvola (1989) shows, in the ancient Greece these themes can be found in Hesiod's poetry and Protagoras's philosophy of human culture. The Roman poet Lucretius in his epic *De rerum natura* gave a moderately optimistic account of the advance of humanity from primitive origins. An important addition to the early history of the idea of progress is the Judeo-Christian eschatology which replaced the cyclic conception of time with linear time. An influential version of this doctrine of divine providence, plan of creation and approach to the last times (the

Savior's Second Coming) was given by Augustine in his *De Civitate Dei* in the early 5th century CE.

Ancient philosophy, Plato's definition of knowledge (Gr. *episteme,* Lat. *scientia*), Aristotle's logic, the Greek ideal of *paideia* (education), classical drama and architecture, and Roman justice were important steps in the development of human culture. Transmitted to the medieval Latin scholastics by the Arabs, this heritage was unified with the Christian religion. In spite of attitudes favoring stagnation and ruling conservatism, the view of science as a collective and cumulative enterprise started gradually to emerge in the medieval universities.[2]

An illuminating episode is told in Dante's *Divina Commedia* in 1310: when the poet meets Odysseus in the Inferno, the Greek hero tells him that after returning from Troy he directed his crew thought the Gibraltar toward the unknown, since man is not created to the life of the brutes but to "the way of virtues and cognition" (*per sequir virtute et conoscenza*).[33] A picture of the same pillars of Hercules was printed on the cover of Bacon's *Novum Organum* three hundred years later, with a quotation from *Daniel* 12:4: "Many will go back and forth, and knowledge will increase" (*Multi pertransibunt & augebitur scientia*). Virtues and knowledge are the two ingredients of practical and theoretical wisdom (Gr. *sophia*, Lat. *sapientia*) which were common ideals of ancient philosophers and their followers during the modern age. It is no accident that the German philosopher Immanuel Kant in 1784 defined Enlightenment by the recommendation *Sapere aude*!, i.e. "dare to know", or "have courage to use your own understanding!". According to Kant, Enlightenment is "man's emergence from his self-imposed immaturity".[4] In this sense, ancient philosophers can be regarded as the first representatives of the Enlightenment tradition.

Kant's three critiques made the important distinction between theoretical reason, practical reason, and judgment (taste). The three Platonist values of truth, goodness, and beauty, which were intertwined in the Renaissance practices of

---

[2] See Molland (1978).

[3] See the essay "Dante between Ulysses and Faust" in von Wright (2003), 193-201. Horkheimer and Adorno (1947) also use Odysseus as the starting point of their discussion of the early interrelations of myth and reason.

[4] See Gay (1973), 384.

scholars, thinkers, and artisans, were now separated from each other as science, ethics, and art. Pure science, the province of universities working with the Humboldtian 1809 program of *Bildung durch Wissenschaft*, was thereby distinguished from technology (cf. Chapter 13 above). Similarly, pure art, studied by the new discipline of aesthetics, was also distinguished from useful technologies. After Kant, questions about progress could be discussed separately in the domains of science, art, and politics.[5]

Rationalization, secularization, and disenchantment are often mentioned as characteristics of enlightened modernity.[6] Kant agreed with skeptical philosophers like Voltaire and David Hume about the impossibility of rational proofs of God's existence, but he acknowledged religion as a postulate of practical reason.

In the meantime, industrial revolution had started in England in the 1760s with the invention of steam engines and weaving machines, opening the way to labor in factories, urbanization, and new means of traffic and communication. First these advances were independent of science, but at the end of the 19th century engineering and agriculture learned to build their activities upon the Baconian ideal of scientification of technology and applied science (cf. Chapter 11 above).

Happy End or Cultural Pessimism?

Enlightenment was continued in the 19th century in many different ways. Auguste Comte's positivism in 1830 argued that humanity has progressed from theology and metaphysics to science. Positive observational knowledge about the laws of phenomena can be rationally applied by the schema: from science comes prevision, from prevision comes action (Comte, 1970). Positivist social philosophers directly influenced progressive politics in some countries like Mexico. Similar emphasis on practical action was typical of the American school of pragmatism.

---

[5] Bruno Latour (1993) sees the distinctions between nature and culture, or science and politics, as constitutive of modernity. He concludes, not entirely convincingly, that "we have never been modern".

[6] See Habermas (1987), von Wright (1993).

Hegel's objective idealism was based on extreme rationalism, but it formulated an influential vision of the dynamic development of mind and culture. Hegel's dialectics as a general theory of change inspired many process metaphysicians, among them the pragmatist Charles S. Peirce. For Hegel the history will come to an end in a finite time, when the objective spirit becomes conscious of itself. Variations of this "endism" were given by Karl Marx's doctrine of the communist society, revived by Georg Lukács and Alexandre Kojève in the early 20th century, and more recently by Francis Fukuyama (1992) as the claim that history has come to its happy end with the victory of the liberal market economy.[7]

Another model of progress was based upon the 19th century idea of evolution which some philosophers like Herbert Spencer and Peirce interpreted as a process toward a predestined limit (such as perfect harmony). While Charles Darwin's evolution by natural selection is a way in which organisms adapt to a changing environment, and thus not directed to some predetermined goal, the models of directed evolution were usually conceived as consisting of indefinite approach which may continue forever. Peirce even defined truth as the ideal limit toward which the opinion of the scientific community with sufficient investigation would lead. A counterpart in politics was Eduard Bernstein's socialist revisionism which takes the journey to be more important than the destination.

The humanist opposition to science started already with the early Renaissance poet Petrarch who in the 14th century raised the question whether knowledge about nature helps us toward a happy life.[8] The same claim was made by Jean-Jacques Rousseau in the heyday of Enlightenment in 1750. The 19th century romantic thinkers emphasized emotion and imagination over reason, unique and particular aspects of history over general laws, individual genius over social communities, art and religion over science. Friedrich Nietzsche suggested that in the reevaluation of all values we have to go "beyond good and evil".

---

[7] Bacon presented in 1620 a theological argument for the progress of science: man's control of nature, lost in Adam's fall, can be gradually regained by science, and the history will end with the completion of knowledge.

[8] See Randall (1940), 213.

The humanist skepticism about technology was expressed by Oswald Spengler in the 1920s in his work on the decline of the Western civilization (cf. Chapter 14 above). Warnings about the dangers of technology were given by Martin Heidegger and Herbert Marcuse. Max Horkheimer and Theodor Adorno (1947) turned this criticism into a questioning of the "dialectics of Enlightenment". In their work, written in the aftermath of the barbarism of the Holocaust, the leading figures of the Frankfurt School claimed that Enlightenment has betrayed its mission by leaving too much power to "technological reason". Jacques Ellul (1964) argued that we have lost our control of technology which now follows its own internal logic with "technological imperatives" (cf. Chapter 20 above).

It is no wonder that C. P. Snow in his thesis about "the two cultures" put pessimistic literary intellectuals in opposition to optimist natural scientists (cf. Chapter 2 above).

The romantic critique of Enlightenment took a new form in the 1980s with the "postmodern" revolt against modernity. Jean-Francois Lyotard (1984) argued that in the postmodern condition "grand stories" or "metanarratives" have lost their credibility, and human mind and culture have been split into incommensurable language games and unsolvable value conflicts. Philosophy should be reduced to deconstruction (Jacques Derrida) or conversation (Richard Rorty), and the notions of truth, goodness, and beauty should be treated as ideological constructions. The modernist urge for novelty in science and art should be abandoned in favor of imitations and repetitions of old pastiches.

Postmodernism is no doubt a reflection of the disillusioned multi-cultural media society: an enormous agora for superstition, old myths, and New Age propaganda. The world has turned out to be much more complex and "messy" than envisaged in modern science and politics. As Elkana (2000) argues, there are good reasons to rethink modernity and Enlightenment in the post-industrial information society.[9] But in spite of

[9] Two limitations of typical Enlightenment thinkers of the 18th century can be mentioned. Newton's paradigm of classical mechanics was enormously successful, but it underestimated the complexity and holism of natural systems, emphasized in a speculative way by romantic philosophers of nature. Cartesian dualism made a sharp division between the rational soul and

healthy self-criticism and some joyful ideas, postmodernists often sound as reactionary advocates of a return to a pre-modern culture and society.

It is therefore significant that there are also cultural pessimists who do not accept the lure of postmodernism and keep up the faith in human reason as out last hope. The Finnish philosopher Georg Henrik von Wright (1916-2003), a champion of logic and analytic philosophy, is a good example of this intellectual quest. In his essay in 1988, he argued that progress is a myth.[10] This challenging thesis provoked a lot of discussion and protests, as it was made by a respected Academician against the optimist trend supported by the projects of science-based economic growth, welfare state and information society. But the Club of Rome, in its 1972 publication *The Limits of Growth*, already had convincingly shown the impossibility of progress as continuous exponential material growth. As further evidence for von Wright's evaluation one can allude to the dark history of the 20th century: world wars, violations of human rights, concentration camps, holocaust, genocide, poverty, hunger, diseases, violence, pollution, and environmental crisis.

Von Wright's argument refutes the notion of unlimited, continuous, and necessary progress of the humanity. Similar arguments, with reference to global terrorism and financial crises in the early 21st century, serve to refute Fukuyama's thesis about the happy end of history. A more general argument against "historicism", or the assumption that there are general laws of human history, was presented already by Karl Popper (1957).

Von Wright seems to infer from the premise that progress is not necessary to the conclusion that progress is not a fact either. But as an expert on modal logic, he certainly saw that this

---

irrational passions, failing to recognize the rationality of many emotions and the support that they may give to cognitive actions (see Chapter 8 above). However, these limitations can be overcome by relying on Enlightenment values, i.e. by new research concerning quantum theory, field theories, non-predictable chaotic and non-linear dynamical systems, mind - body interactions, neurophysiology and cognitive psychology.

[10] See the essay "The Myth of Progress" in von Wright (2003), 202-228.

inference would not be warranted.[11] Rather, his point is that progress is a value notion. To see what this means, it is useful to briefly consider the prospects of progress in the fields of science, art, and society.

Progress in Science

The historian of science George Sarton (1936) argued that "the acquisition and systematization of positive knowledge are the only human activities which are truly cumulative and progressive", and "progress has no definite and unquestionable meaning in other fields than the field of science". However, when the positivist cumulative account of scientific progress was challenged by new models of scientific change in the 1960s, it turned out that there is no consensus among philosophers of science about the definition or characterization of advancement in science.[12]

The accumulation-of-truths view does not pay attention to the uncertainty of all results of scientific inquiry. Even the best theories may be refuted by new observations and experiments or corrected by new theoretical frameworks. Popper's (1972) fallibilism suggests that science progresses by the succession of theories which are false but approach the truth. Progress in this sense means increasing truthlikeness or verisimilitude.[13] Thomas Kuhn (1962) argued that during periods of "normal science" the scientific community agrees on a paradigm which determines the relevant problems and criteria for their solution. Normal science eventually leads via anomalies to a crisis and a revolutionary paradigm shift. While Popper defended the idea of permanent revolution in science (but not in politics), Kuhn and Paul Feyerabend claimed that new theories are incommensurable with the old ones: new conceptual frameworks change the interpretation of observations and the formulation of facts, so that there are no theory-independent truths. Imre Lakatos proposed a compromise where the development of science is seen as a battle between rival "scientific

[11] A non-necessary proposition may be contingently true. As we shall see, this is the melioristic view about human progress.

[12] See Niiniluoto (2019).

[13] See Niiniluoto (1984).

research programmes".[14] Kuhn and Larry Laudan (1977) suggested that progress could be understood in terms of the problem-solving capacity of research traditions, while the methodological pragmatist Nicholas Rescher (1978) identified progress with new useful technological applications.

The main lesson from these lively debates is that progress in science is a normative value concept, relative to the aims of scientific inquiry, such as truth, accuracy, information, explanatory power, simplicity and even beauty. Definition of "real" progress in terms of such "epistemic utilities" has to be distinguished from indicators of "epistemic" progress which give us reasons for thinking that some actual steps in the development of science in fact are progressive. This distinction between real and apparent progress is important to the realist view which takes seriously the idea that science attempts to describe an external reality (nature, mind, or society), but it is not easy and straightforward to know that progress in this task has actually been achieved. But on the whole the increasing success of scientific theories in predictions and the pragmatic guidance of action is an indicator of the increased truthlikeness of these theories.[15]

Progress in Art

In science there is a fairly general consensus that Einstein's theory of relativity is in some sense an improvement of Newton's classical mechanics. But is Picasso better than Rembrandt or Stockhausen better than Bach? At the end of the 19th century, the Austrian art historian Alois Riegl concluded that the history of art (painting, sculpture, architecture, music) exhibits different styles (*Stilformen*) which cannot be compared with each other. Hence, the notions of progress and regress have no applications in art. The epistemological anarchist Paul Feyerabend in *Wissenschaft als Kunst* (1984) joyfully agreed and urged that Riegl's account is valid in science as well.

The other extreme position would be that everything in art is progressive: all creative works, which involve something

[14] See Lakatos & Musgrave (1970).
[15] See Niiniluoto (1999).

personal originality and novelty, are examples of artistic progress.[16]

When scientists formulate new theories, their success is assessed by additional epistemic criteria like truth and explanatory power. For art, one might think that beauty could serve as a similar overarching criterion: the artifactual works of creative imagination should be beautiful by some objective standard or by personal aesthetic experiences. But beauty is an even more difficult concept to analyze than the notion of truth (cf. Chapter 3 above). The fact that old and classical forms of art retain their popularity shows that it is not always easy for radical modernists to educate their audience, who assess works of art by their own taste.

Many theories of art have proposed general principles for the evaluation and criticism of works of art. Realism sees art as representation of external or internal reality (cf. Chapter 4 above). Emotivism defines art as expression or transmission of emotions. Functionalism treats works of art as artefacts whose form should follow their practical function. Political art is assessed by its impact on social progress, and religious art by its impact on spiritual life. However, each of these approaches, in any case, seems to favor a particular perspective on artistic activities.

Riegl's treatment of art resembles the Kuhn - Lakatos picture of scientific change as a competition between rival paradigms and programs - such as realism, symbolism, impressionism, expressionism, and cubism. This allows to speak about progress and regress within the conventions of an artistic program: it is widely agreed that Johann Sebastian Bach was the greatest master of baroque music. But it is always possible to be an artistic revolutionary or "modernist" and propose changes in the rules of criticism and evaluation in art. Even postmodern art, in spite of its denial of genuine progress, can be viewed as an attempt to create something new.[17]

---

[16] Among common features of progress in science and art one can mention creative imagination and the construction of new instruments.

[17] See Wallis (1984).

Social Progress

According to von Wright (2003, 211), "no facts about diminishing illiteracy, improved sanitary conditions or increased per capita income" can prove the "modern Myth of Progress", which claims that "men and their societies will thrive better if they are free to follow Kant's maxim to trust reason rather than authority". Such facts do not measure progress, since progress is a value concept, so that the sole measure of progress is "the way people thrive in the circumstances under which they live" (*ibid*., 223).

Here von Wright seems to forget the lesson from Aldous Huxley's novel *Brave New World* in 1932: satisfaction can be artificially brought about by *soma* or happiness drug. Experienced "thriving" is a relation between goals and their realization, so that satisfaction can be too easily guaranteed by lowering the level of aspiration. As von Wright himself argued in *The Varieties of Goodness* (1963), the value standards of welfare or quality of life should refer both to subjective experiences (emotions, happiness) and objective criteria (social indicators, measurable states of affairs or social facts) (cf. Chapter 18 below).

I agree with von Wright about the significance of the fact - value distinction (cf. Chapter 8 above). From the facts about human history we cannot logically derive the value criteria of good human life. But values as human constructions are not fictions but real factors which influence and guide our behavior. So after all fact - fiction is not the right contrast in the discussion about progress. We have also seen in the context of science and art that factual developments can be assessed as progressive or regressive relative to some value standards: a step from conditions A to B is progressive if B is better than A by standards S.

In a democratic society it is up to the citizens themselves to decide about their standards S of social progress. This does not mean that S should describe some perfect ideal state: in his defense of "piecemeal social engineering", Popper (1957) argued that utopias have been harmful in human history, and a better model of political progress is the reformist removal of some concrete defects like human suffering.

An important alternative to the traditional method of measuring progress by means of economic growth, the United Nations Development Project (UNDP) has since 1990 compared the UN member states by means of the Human Development Index (HDI), calculated as the mean of three factors: health (life expectancy at birth), education (adult literacy, years of schooling), and living standards (wealth measured by GDP per capita). In 2020 the top countries by HDI are Norway, Ireland, and Switzerland. Finland is number 11, and the last one in the list is Niger (189).

The Genuine Progress Index (GPI), proposed by Redefining Progress, adds to GDP other economic factors like income distribution, services outside the market, and costs of negative effects (crime, resource depletion, pollution, loss of wetland). In Finland GPI per capita has decreased since 1989.

The Prosperity Index, published by Legatum Institute, includes as its dimensions economy, entrepreneurship and opportunity, governance, education, health, safety and security, personal freedom, and social capital. In 2021 Denmark, Norway, Sweden, and Finland are leading, and the last is South Sudan (167).

The Happy Planet Index (HPI), published by the New Economic Foundation since 2006, takes seriously the value of environmental protection and sustainable development. It uses the formula: life satisfaction x life expectancy per ecological footprint. Satisfaction is assessed by tests of happiness (the best results are achieved by Finland, Denmark, and Iceland). The worst countries by their ecological footprint are Qatar, Luxembourg, and Bahrain. In the overall HPI ranking the highest scores in 2021 are reached by Costa Rica, Vanuatu, and Colombia, while Finland (33) and Norway (39) are in the upper middle range, and Qatar on the bottom (152).

These new indicators have been criticized for difficulties in the reliable measurement of the components and their arbitrary combinations. However, they are all variations of the theme of social progress: data from successive years show whether a nation has made progress relative to chosen criteria.[18] The governments of many countries – including France, United Kingdom, and Finland – have seriously discussed the

[18] See Rosling (2018).

development and use of these indicators. They have also been used to monitor the advances of the UN Millennium goals (e.g. education of girls, reduction of serious poverty) and the Agenda2030 program.

An interesting example of a complex future oriented social indicator is the State of the Future Index (SOFI), produced by the Washington University Millennium Project. By using the Delphi method of expert evaluation, it gives predictions about 29 variables, forecasting global progress e.g. in life expectancy, adult literacy, GDP per capita, internet users, and reduction of conflicts and child mortality, and regress in carbon dioxide waste, terrorism, corruption, global warming, and losses of suffrage and employment. SOFI is calculated as a weighted average of these factors, with variable weights of importance assigned by the panelists.

Conclusion

The clash between optimism and pessimism, or modernism and conservatism, is a pervasive feature of our cultures. Recent illustrations are the books by Matt Ridley (2010) and Roger Scruton (2010): the former is cheerful about economic trade and co-operation as means to progress,[19] the latter argues that constraints and boundaries of human nature and custom make impossible any rational transformation of society. In my view, this clash should be overcome by the melioristic position which does not assume any lawlike development of progress or regress.[20] This view is in fact presupposed in the fashionable notion of sustainable development, defined by Brundtland's commission in 1987 as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs" (cf. Chapter 17 below). In spite

[19] A strong statement for optimism is given by the cognitive scientists Steven Pinker's bestseller *Enlightenment Now* (2018). The physician and statistician Hans Rosling (2018) presents facts to show that "things are better than you think". It remains to be seen how the optimist stance can recover from the blow of the Russian attack on Ukraine in 2022.

[20] Meliorism (from Latin *melior* = better) was advanced as a via media between optimism and pessimism by the pragmatists William James and John Dewey (see Pihlström, 2021, 83). Rosling (2018) advocates this view with his "possibilism".

of the ambiguity of the value concept of need, the notion of environmental, economic, and social sustainability is remarkable, as it appears to acknowledge that progress is not a necessity but rather an ethical demand for all of us. The critical von Wright (2003, 227) agrees with this melioristic conclusion when he states that we should not "abandon work for progress as a task".

Thus, in spite of many drawbacks and threats, the Enlightenment idea of progress is still alive. While we do no more believe that progress is a lawlike necessity in any sector of human life, it is still a possibility which partly depends on our own attitudes and activities. In futures studies, methods have been developed for dealing with complex unpredictable systems where the goals are also local, multidimensional, and revisable in new situations. Typically, the future is seen as an open tree of possibilities, but it can be at least partly influenced and designed by assessing the probability and desirability of alternative scenarios (cf. Chapter 12 above).

Even in a chaotic and unpredictable world we have all reason to work together in planning and realizing a better future.

Note. This chapter is based on a lecture in the Aboagora symposium in Turku on August 16, 2011. It has been published in *Temenos: Nordic Journal of Comparative Religion* 47 (19) (2011), 97-110.

Literature

Bacon, Francis (1960). *The New Organon*. Indianapolis: Bobbs-Merrill.

Bury, J. B. (1932). *The Idea of Progress*. New York: Macmillan. (1955 New York: Dover.)

Comte, August (1970). *Introduction to Positive Philosophy*. New York: Bobbs-Merrill.

Elkana, Yehuda **(**2000). "Rethinking - Not Unthinking – Enlightenment". In Wilhelm Krull (ed.), *Debates in Issues of Our Common Future*. Weilerwist: Velbruck Wissenschaft.

Ellul, Jacques **(**1964). *The Technological Society*. New York: Alfred A. Knopf.

Feyerabend, Paul **(**1984). *Wissenschaft als Kunst*. Frankfurt am Main: Suhrkamp.

Fukuyama, Francis **(**1992). *The End of History and the Last Man*. New York: The Free Press.

Gay, Patrick (ed.) (1973). *Enlightenment: A Comprehensive Anthology*. New York: Simon and Schuster.

Habermas, Jürgen (1987). *The Philosophical Discourse on Modernity*. Cambridge, MA: The MIT Press.

Horkheimer, Max & Adorno, Theodor (1947). *Dialektik der Aufklärung*. Amsterdam: Querido. (English transl. 1972 *Dialectics of Enlightenment*. New York: Herder and Herder.)

Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press. (2nd enlarged ed. 1970.)

Lakatos, Imre & Musgrave, Alan (eds.) (1970). *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.

Latour, Bruno (1993). *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.

Laudan, Larry (1977). *Progress and Its Problems: Toward a Theory of Scientific Growth*. London: Routledge and Kegan Paul.

Lyotard, Jean-Francois (1984). *The Postmodern Condition: A Report on Knowledge*. Manchester: Manchester University Press.

Molland, A. G. (1978). "Medieval Ideas of Scientific Progress". *Journal of the History of Ideas* 39, 561-578.

Niiniluoto, Ilkka (1984). *Is Science Progressive?*. Dordrecht: D. Reidel.

Niiniluoto, Ilkka (1999). *Critical Scientific Realism*. Oxford: Oxford University Press.

Niiniluoto, Ilkka (2019). "Scientific Progress". In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Winter 2019 edition). https://plato.stanford.edu/entries/scientific-progress/

Nisbet, Robert (1980). *History of the Idea of Progress*. London: Heinemann.

Pihlström, Sami (2021). *Pragmatist Truth in the Post-Truth Age*. Cambridge: Cambridge University Press.

Pinker, Steven (2018). *Enlightenment Now: The Case of Reason, Science, Humanism, and Progress*. Harmondsworth: Penguin Books.

Popper, Karl (1957). *The Poverty of Historicism*. London: Routledge.

Popper, Karl (1972). *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press. (2nd enlarged ed. 1979.)

Randall, John H. Jr. (1940). *The Making of the Modern Mind*. New York: Houghton Mifflin.

Rescher, Nicholas (1978). *Scientific Progress: A Philosophical Essay on the Economics of Research in Natural Science*. Oxford: Blackwell.

Ridley, Matt (2010). *The Rational Optimist: How Prosperity Evolves*. London: Harper Collins.

Rosling, Hans (with Anna Rönnlund & Ola Rosling) (2018). *Factfulness: Ten Reasons We're Wrong about the World – and Why Things are Better than You Think*. New York: Flatimer Books.

Sarton, George (1936). *The Study of the History of Science.* Cambridge, MA: Harvard University Press.

Scruton, Roger (2010). *The Uses of Pessimism and the Dangers of False Hope*. London: Atlantic Books.

Sihvola, Juha (1989). *Decay, Progress, the Good Life? Hesiod and Protagoras on the Development of Culture*. Helsinki: Societas Scientiarum Fennica.

Wallis, B. (ed.) (1984). *Art After Modernism: Rethinking Representation*. New York: The New Museum of Contemporary Art.

von Wright, G. H. (1963). *The Varieties of Goodness*. London: Routledge and Kegan Paul.

von Wright, G. H. (1993). *The Tree of Knowledge and Other Essays*. Leiden: Brill.

Chapter 16: Defending Truth in the Post-Truth Era

The two most fundamental notions in theoretical philosophy are "truth" (Gr. *aletheia*, Lat. *veritas*) and "knowledge" (Gr. *episteme*, Lat. *scientia*). When Plato in the dialogue *Theaetetus* defined knowledge as true belief with justification or explanation (201d), he implied that truth is even more basic than knowledge. In the first two sections of this chapter, I review the ongoing debate among philosophers about theories of truth, and defend Alfred Tarski's semantical definition as a form of the correspondence theory of truth. This objective non-relativist concept, amended by the notion of truthlikeness, is appropriate for the project of combining fallibilist epistemology with critical scientific realism. The next two sections comment on the recent discussion about the "post-truth era" in media and politics, asking what kind of responsibility various kinds of alethic relativists and anti-realists may have for this unfortunate "postmodern condition".

Tarskian Truth as Correspondence

The concept of objective truth is part of the legacy of Plato's and Aristotle's metaphysics and epistemology. In this view, truth is *correspondence* between beliefs and reality: the truth of a belief or statement p is constituted by the correlation of p with some fact or facts obtaining in the world W. Thus, p is true if it describes an actually existing state of affairs, i.e. expresses a fact; otherwise it is false. Aristotle argued for the objectivity of truth by his statement in *Metaphysics* (1011b25): "to say of what is that it is not, or what is not that it is, is false, while to say of what is that it is, and of what is not that it is not,

is true".[1] According to the medieval formula, supported e.g. by Thomas Aquinas, *veritas est adequatio rei et intellectus*. In modern semantical terms, employed by Ludwig Wittgenstein's *Tractatus* in 1922 and Alfred Tarski's definition in 1935, truth is a relation between language and reality – so that it can be treated as a relational property of a representation (cf. Chapter 4 above). According to *semantical realism*, whether the truth-relation obtains between p and W is independent of our beliefs and wishes. Thus, factual or material truth in this objective sense has to be distinguished from the epistemological and methodological indicators (if any) that help us to recognize truths.[2]

Plato and Aristotle assumed that *episteme* can be given complete justification by evidence or proof which shows that it could not be false. Plato's paradigm for such justification was geometrical proof, and Aristotle relied on syllogistic logic. This kind of infallibilism was advocated by modern rationalists (e.g. René Descartes) but doubted by most empiricists. Indeed, it is not plausible outside formal sciences (logic, mathematics). Later mainstream epistemology, called *fallibilism* by Charles Sanders Peirce in the late 19th century, states that no justification of factual truth claims guarantees full certainty: objective truth exists, but even our best attempts to find these truths in science are to some extent uncertain and corrigible by further evidence (cf. Chapter 7 above).

According to the correspondence theory, truth is a *relation* between truth-bearers and truth-makers. *Truth-bearers* are linguistic entities, like interpreted sentences, statements, propositions or beliefs. *Truth-makers* are extra-linguistic entities, like existing states of affairs, facts, Russellian propositions or situations.[3] Such truth-makers are parts or fragments of reality, but sometimes the actual world as a whole is chosen as the truth-maker of factually true statements. On the basis of the truth relation, truth as a property can be defined by

---

1 Cf. Tarski (1944), 342.

2 Factual truth in the actual world has to be distinguished also from logical truth (i.e. truth in all possible worlds).

3 See Armstrong (2004).

(1) Truth-bearer p is true iff there is a truth-maker W such that p is true in W.

The classical objection to the correspondence theory is that notions like "correspondence", "correlation" or "fit" are not well defined. The rival *coherence* theory argues that truth means "coherence" or "consistency" between propositions. This view was supported by some British neo-Hegelians (F. H. Bradley), whose idealist system did not allow to make a difference between thought and reality. Another version was advocated in the early 1930s by some logical empiricists (Otto Neurath, Carl G. Hempel), who claimed that it is meaningless to speak about language – world relations. Therefore, they thought, the definition of truth has to be expressed by intra-linguistic relations between statements. This leads to the unresolved problem that "truth" is not uniquely fixed, since there are several mutually incompatible but internally coherent classes of propositions. Hempel (1935) suggested that we should pick out the system which is "actually adopted by the scientists of our cultural circle", but this is a characterization of knowledge rather than truth.

The Polish logician Alfred Tarski published in 1933 in Polish a new *semantical definition* of truth for formal languages. When he presented this theory in a congress in Paris in 1935, this was a revelation to many of his fallibilist contemporaries: Rudolf Carnap learned to distinguish objective truth from time-dependent and evidence-relative confirmation, and Karl Popper rejoiced about the "rehabilitation of the correspondence theory of truth".[4] Tarski's starting point was an interpreted language L which has a well-defined syntax, and a metalanguage ML which speaks about the linguistic elements of L (e.g. if p is a sentence in L, 'p' is its name in ML). He required that an adequate definition of truth for L should entail all instances of the T-equivalence of the form

(T) x is true in L iff p,

where x is a name of a sentence in the object language L and p is the translation of this sentence in the metalanguage ML. In a

4 See Niiniluoto (1999c).

special case, where L is English and included in ML, (T) could be e.g. of the form

(2) 'Snow is white' is true iff snow is white.

The mistaken assumption that (2) alone would be a definition of truth has led to the charge that Tarski's account is merely "disquotational" or leads to the redundancy theory of truth.[5] In contrast, Tarski's main idea was to reduce the truth conditions for complex sentences with quantifiers (all, some) to the satisfaction of atomic formulas by sequences of elements from the "class of all objects".[6] This formalism was made explicit in the 1950s in *model theory*. Even though Tarski followed his logical teachers from the Lvov-Warsaw school[7] in avoiding to speak about "facts", his later set-theoretical account of relational systems and models (i.e. domains of objects with subsets and relations) can be understood so that these constructions serve as truth-makers of true sentences, and the truth conditions specify what is meant by the "correspondence" between sentences and models.[8]

In Tarski's mature model theory, the truth of a sentence s in language L in model or L-structure W is mediated by a third factor, viz. the interpretation function I from L to W.[9] Thus, truth and falsity in L are relative to the model W and the interpretation I. Tarski's early account did not make the interpretation of a language explicit. In this respect, Rudolf Carnap's version of logical semantics in the late 1930s and in his *Introduction to Semantics* (1942) was an advance in comparison to Tarski's papers in 1936 and 1944: like in later model theory, Carnap took a *semantical system* S to consist of a language as a

---

5 See Horwich (1990).

6 For example, the universal sentence $\forall x F(x)$ is true in W iff every object a from the domain of W satisfies the open formula F(x). See Tarski (1944, 1956).

7 See Wolenski (1989), Wolenski & Simons (1989).

8 Niiniluoto (1999b, 2004).

9 For example, function I maps individual constants of L to elements of W, one-place predicates to subsets of W, two-place relations to relations in W etc. The treatment of quantifiers is objectual in the sense that open formulas are satisfied by objects from the domain of W. The interpretation function is one way of explicating the semantical concept of meaning (see Niiniluoto, 2001).

set of uninterpreted syntactical signs and a designation function Des such that names designate objects, predicates designate properties and relations of objects, and sentences designate propositions. Then

(C) Sentence s is true in S iff there is a proposition p in S such that s designates p and p.[10]

This general definition is Carnap's counterpart to Tarski's T-schema which can be written in the explicit form as follows:

(T´) If p in metalanguage ML is the translation of sentence s in L, then s in L is true iff p.

According to Kirkham (1992), when propositions are replaced by states of affairs, schema (C) expresses the "essence" of the correspondence theory of truth.

As the actual world W is one among the possible worlds, the notion of actual truth can be defined within model theory.[11] Sometimes it is objected that the actual world is not a set-theoretical structure in the sense required by model theory: the world is not "ready-made", as "metaphysical realists" erroneously assume.[12] But let the pair K = (L,I) be an interpreted linguistic framework. If the interpretation function I is semantically determinate, its values for the extra-logical terms of L (i.e. individual constants, predicates, and relations) within the domain of actual objects in W constitute a unique L-structure W(K). This structure or "word version" consists of the facts of the world W from the point of view of the framework K. Here W(K) is not an epistemic notion or a Kantian "world for us" - it is not what we believe about the world W, but what the language L is able to tell about the world W if W were investigated via the framework K.[13] Factual truth about W in language L interpreted by I (or with meanings specified by I) is

---

10 See Niiniluoto (2003b). In his later work Carnap developed the idea of propositions as functions from possible worlds to truth values, which led to the discovery of the possible world semantics by Stig Kanger and Jaakko Hintikka in 1957.

11 See LePore (1983), Niiniluoto (1999a), 220-226.

12 See Putnam (1981).

13 In terms of analytic metaphysics, the non-causal dependence relation between W and W(K) is called "grounding".

defined by Tarski's model-theoretical definition of truth in structure W(K). We need not assume with metaphysical realists that there is an ideal framework K_{id} which covers all of the variety of the actual world W so that $W = W(K_{id})$. Instead, each conceptual framework K captures only a partial fragment of the inexhaustible reality W. But for each framework K truth is objective in the following sense: we may choose L and I, but the world W decides the referents of the terms of L and the relevant truth values of sentences of L. W(K) thus includes the truth-makers for the sentences of L, and truth in W(K) is truth about the actual world. For other frameworks K' we have other truths in W(K'), but as descriptions of the same world W they cannot logically contradict the truths of K. This means that "genuine relativism" is avoided. Thus, the position outlined here is not relativistic. Rather, it expresses *conceptual pluralism*: the world can be described in alternative linguistic frameworks, and all of these frameworks may have interesting objective truths for us to offer.

For the semantical approach, it is important that truth and falsity are properties of complete sentences. For example, the sentence "The sun is shining" is sometimes true and sometimes false. But such an incomplete or elliptical sentence can be completed with spatio-temporal indicators, e.g. "The sun is shining in Helsinki on the 10st of March 2022". This complete judgment as a proper truth-bearer is absolutely true or false, depending on facts about the city of Helsinki on that day. If true then, it is eternally true or true at all later times. This treatment is today a standard way of assigning truth values to statements or utterances with *indexicals* (like "we", "here", "now"). In this extension of Tarskian truth-conditional semantics, truth is determined by interpretation (meaning), world (model), and context (including agent, location, and time).

Other similar cases have recently been discussed by a new school of "truth relativists".[14] Examples include statements about taste ("Marilyn is prettier than Jane"), epistemic possibility ("the murderer might have been the butler"), knowledge attributions ("Charles knows that he has hands"), value statements ("honesty is good"), normative statements ("one ought

14 See Carcia-Carpintero & Kölbel (2008). For discussion, see Cappelen & Hawthorne (2009) and Niiniluoto (2013).

not to kill"), and future contingents ("Spain is the World Champion of football in 2022", stated before 2022). The relativist strategy, largely inspired by David Kaplan, is to treat these statements as special kinds of propositions, comparable to temporally indefinite sentences, with a truth value changing relative to extra factors - such as standards of taste, state of knowledge, standards of justification, systems of morality, and time of utterance. But, in my view, it is more natural to treat them as incomplete statements which have Tarskian truth values only after a relevant completion.

For example, categorical value statements (e.g. "killing is wrong") lack truth values, but they can be completed by reference to some axiological system (e.g. "killing is wrong by the ethical principles of Christianity") or a code accepted by some community (e.g. "killing is a punishable crime by the legal order valid in Finland"). With such a completion, which is generally not uniquely determined by the context of utterance, these statements have objective truth values. I have called this view *modest moral relativism*,[15] but it need not and should not be construed as an instance of alethic relativism about truth.

Against Relative Truth

The idea of objective truth was challenged in the ancient Greece by the sceptics and the relativists. The sceptics went as far as denying the possibility of any interesting representations of reality, so that there is no difference between knowledge and error. They also claimed against Plato's and Aristotle's strong notion of justification that all human beliefs are mere opinions (Gr. *doxa*), so that the best advice is to suspend judgment (Gr. *epochê*). The relativists suggested that truths may be relative to our own species or its individual members: the famous slogan of the sophist Protagoras asserted that "man is the measure of all things" (*homo mensura*). In his dialogue *Theaetetus* (152a), Plato presents Protagoras as arguing that any given thing "is to me such as it appears to me, and is to you such as it appears to you".

---

15 See Niiniluoto (1999a), Ch. 8.2. This view differs from those scientific realists who are also moral realists (e.g. Richard Boyd), i.e. think that both factual and value statements have objective truth conditions.

Alethic relativism about truth claims that truths and falsities are relative to persons, communities, cultures, beliefs, paradigms, or perspectives. In its global version, this relativity holds for all propositions. In its radical form, all truth claims are claimed to be equally good (or equally bad), while modest relativism allows some rational comparisons between them. But there are good arguments to show that such views of are mistaken.[16] A straightforward formalization of subjective *Protagorean* relativism would try to define "p is true for person a" by the condition that agent a believes that b.[17] But this concept would not behave in the same way as truth – in particular, as beliefs may be false, it would not satisfy Tarski's T-equivalence (T). Besides such logical worries, which show why it would be at least misleading to conflate the notions of truth and belief, there are epistemological objections to the idea of personal truth. If truth is identified with personal belief, I could not admit that there are some truths unknown to me or that some of my beliefs are false.[18] To such observations one can add difficulties related to classical incoherence arguments concerning relativism.[19] What does it mean to a relativist that she believes that p? If this statement about belief has non-relative truth conditions, then at least the global relativist has lost the battle. Alternatively, the claim $B_a p$ (a believes that p) should be understood relatively as the thesis that $B_a p$ is true for person a, i.e. $B_a B_a p$. But this iteration of belief-operators or relative truth-operators can be continued without limit. For example, Putnam (1981, 120) argues that according to Protagorean relativism my utterance "Snow is white" has to be understood as

(3) I think that I think that I think that I ... (with infinitely many 'I thinks') that snow is white.

But this makes it difficult for the relativists to communicate their position and for others to understand it.

Similar critical remarks apply to *cultural relativism*. In the latter half of the 19th century, it became fashionable as a

16 See Niiniluoto (2006, 2013).

17 Cf. Swoyer (1982). For the generalization "p is true-for-V", see Meiland (1979).

18 Krausz & Meiland (1982), 82.

19 See Siegel (1987, 2004) and replies by Kusch (1991), 200-206.

philosophical interpretation of the emerging historical, ethnological, and social studies of the amazing *de facto* variation in human cultures. Even though these cultures include different cognitive systems, which may function in their local environments, we need not say that their beliefs or world views are "truths". This way to relative the notion of truth is simply a case of sloppy terminology. Sometimes the relativists formulate their position by stating that persons with different belief systems do not share a common "world", where this notion is used in a non-objective sense. For example, Thomas Kuhn (1970) asserts that scientists with different paradigms "live in different worlds", so that there is no theory-neutral concept of truth. Claims of this sort have also been made by some sociologists of science and social constructivists, who take science to be only one of the alternative cultures.[20] The "strong programme" of David Bloor's (1976/1991) Edinburgh School defends relativism by the doctrine of "finitism", which claims that all applications of descriptive terms in language depend on social factors and negotiations. The "empirical programme of relativism" by Harry Collins (1991) urges that "the natural world has a small or non-existent role in the construction of scientific knowledge". However, what these claims reveal, instead of relativism about truth and reality, are only instances of the relativity of beliefs.[21] As social constructivists, Bruno Latour and Steve Woolgar (1979/1986) claim that facts and theoretical entities are not causes but consequences of scientific work, so that they are "socially constructed" by group negotiations in laboratories. More generally, nature and society emerge from "nature/society making".[22] Among philosophers of science, a radical form of relativism has been advocated by Paul Feyerabend's (1975, 1987) "epistemological anarchism", which urges that "anything goes" and treats all beliefs – science and voodoo – on a par.

---

20 See Barnes & Bloor (1982).

21 For a critical evaluation of Bloor and Collins, see Niiniluoto (1999a), Ch. 9.1 and 9.2.

22 For criticism, see Niiniluoto (1999a), Ch. 9.3.

Similar critical questions can be raised about the thesis that all truths are relative to perspectives or points of view.[23] *Perspectivism* is a view that is often attributed to Friedrich Nietzsche: "there are no facts, only interpretations".[24] If an alethic relativist proposes that the notion of truth should be relativized to perspectives or viewpoints, then we may ask about the status of statements of the form "Proposition p is true-from-perspective A". If this statement is only true from a perspective, we end up with an infinite number of iterations of viewpoints.

An alternative approach would be to define truth as *assertibility*, characterized by the condition that a proposition p is *provable* in an axiomatic system S. It is known that this treatment leads to truth-systems satisfying the principles of intuitionistic logic,[25] but the attempt to extend this notion of proof from mathematics to empirical statements faces serious difficulties.[26] Further, Gödel's incompleteness theorem shows that for all axiomatic systems of arithmetic there are true but non-provable statements, so that truth and provability cannot be identified even in mathematics.

The school of American pragmatism developed fallibilist epistemology with an *epistemic concept of truth*: notions like "verification" (William James) or "warranted assertability" (John Dewey) serve as surrogates of objective truth. For Putnam (1981), truth is an epistemic notion. For Richard Rorty, "truth is what your contemporaries let you get away with". Such epistemic definitions of truth easily lead to alethic relativism, since the quality and amount of warrant or evidence varies with time – and therefore they are comparable to epistemic probability and confirmation rather than truth.[27] Instead of defining knowledge by truth (and justification), they reverse Plato's classical definition by trying to give the best description of human knowledge seeking, e.g. by coherence, empirical verification, proof, acceptance, acceptability, warranted

---

23 Hautamäki (2020) develops a precise and useful account of "points of view", but his version of alethic relativism is compatible with critical realism which accepts conceptual pluralism.

24 Danto (1973), 37.

25 Dummett (1978).

26 See Niiniluoto (2001).

27 For criticism, see Niiniluoto (1999a), Ch. 4.6.

assertability, or community consensus. Then the results satisfying such conditions are deemed to be instances of "knowledge" and "truth". This strategy has been characteristic to pragmatist and neo-pragmatist thinkers who think that truths are not waiting out there to be discovered, as the realist assumes, but they have to "made" by our acts of verification. These approaches do not exclude the possibility that some conflicting propositions are true at the same time. Moreover, for this view there are no unknown truths, and all truths up to now have to be true for someone. It is also clear that warranted assertability does not satisfy Tarski's T-equivalence: a statement may still be false in spite of its warrant, and conversely some unknown truths cannot be confirmed. For example, facts about the year 1949 determine the unknown truth value of a statement about the number of sneezes in 1949 by Winston Churchill, but no evidence is any more available as a warrant for this assertion.[28] To avoid these conclusions, one could try to appeal to ideal conditions, such as "the limit of inquiry" (Peirce), the "ultimate consensus" of a discourse community (Habermas) or "ideal acceptability" (Putnam),[29] but it is not clear that this can be done in a non-circular way without presupposing some sort of semantical realism. Further, the epistemic project of defining such ideal conditions has hardly so far been as successful as the semantic account of objective truth.

To wish to avoid skepticism by restricting meaningful sentences to those which can be verified or decided by empirical means has been a reason for Dummett's (1978) *semantical anti-realism* - and even anti-realism about the past (cf. Chapter 1 above). In the philosophy of science, a similar motive has been used to defend instrumentalism, which take scientific theories to be conceptual tools without truth values (Pierre Duhem). Truth has been under attack also from various kinds of theoretical anti-realists: as alternative versions to traditional instrumentalism, theoretical statements are assumed to have truth values but they are irrelevant (Bas van Fraassen), utopian (Larry Laudan) or beyond our grasp (Kyle Stanford). Such views clash with scientific realism in their skepticism about

---

28 This is Bertrand Russell's (1940) argument against Dewey.
29 Putnam (1981).

theories, since realism allows the scientists to postulate theoretical entities to explain phenomena, but otherwise they are not intended to undermine the value of science.[30]

Another popular *pragmatist* notion of truth, inspired by some of James' statements about truth as a species of value (e.g. "truth pays"), identifies "being true" and "being useful for human purposes".[31] But again this notion fails to satisfy the T-equivalence: sometimes false hopes and even lies may be helpful for our lives, and successful hypotheses may be only truthlike. Still, as the Marxists rightly observed, its core idea is valid in the sense that "practice is the criterion of truth".

The critical remarks on subjective relativism apply also to attempts to link truth and various kinds of group beliefs. What is usually called "scientific knowledge" is constituted by the shared opinion of the scientific community.[32] These tentatively accepted results of scientific inquiry are historically changing. As we are accustomed to thinking with Plato that knowledge implies truth, the change from Newton's theory to relativity theory and quantum theory has been taken to indicate that scientific "truth" has changed. Such considerations gave motivation to Hegel's dynamic account of "truth as a process", the British neo-Hegelian doctrines of "degrees of truth", and the Marxist discussion about the "dialectics of absolute and relative truth".[33] But a more natural interpretation of scientific change is to acknowledge that truth does not change or become relative with the variable collective beliefs and theories of the scientists: the time-dependent system of "scientific knowledge" does not generally satisfy Plato's definition of *episteme* with the success condition, but by the self-corrective method it approaches the truth.[34]

The lure of gender relativism – e.g. distinctions between men's and women's knowledge or between masculine and feminine methods of inquiry – has been a common feature of

---

30 See Niiniluoto (1999a), Ch. 5.

31 It may be questioned whether James intended such characterizations as a definition of truth. See Pihlström (2021), 15-22. But Hasok Chang (2022) emphatically presents his neo-Jamesian account of "operational coherence" as a definition of truth.

32 See Niiniluoto (2003a).

33 See Niiniluoto (1987), 164-179.

34 See Niiniluoto (1987, 1999a).

feminist epistemology.[35] Sandra Harding's (1987) feminist standpoint theory appeals to the difference between women's and men's experiences, claiming that the former as "oppressed" give better grounds for knowledge than the former. This is a modification of the Marxist doctrine that morality is relative to social class, but the morality of the "progressive" proletariat in a capitalist society is right or epistemically privileged in relation to the bourgeoisie class. Linda Nicholson's (1990) postmodernist feminism deconstructs the standpoint theory by denying the legitimacy of speaking in the name of abstract "woman". But tighter and tighter intersectional divisions of the relevant classes of women eventually end with singletons with one member only, so that as an account of "truth" this position is reduced to Protagorean subjectivism.

The sociologists of science have been interested in investigating the background conditions and circumstances where scientific knowledge is produced (e.g. concepts, cultures, social interests, and power). This as such useful program is shared by Michel Foucault, who has studied how political power produces knowledge – or how truth subdues our freedom. His work on the discourses and "regimes of truth" has inspired many political scientists. But Foucault's genealogy, which is influenced by his reading of Nietzsche, is one-sided in the sense that it largely overlooks the other dimension in the double determination of scientific knowledge: the interaction with the external world by means of observations and experiments. It is also misleading that truth is not clearly distinguished from beliefs (i.e. holding to be true). Martin Kusch (1991) acknowledges that Foucault's approach leads to perspectivism, epistemic relativism, and irrealism – even though as a relativist himself Kusch does not find this problematic. The political fear that objective truth would be detrimental to democracy is expressed by the Italian hermeneutic philosopher Gianni Vattimo's (2011) "farewell to truth".

Indifference and even hostility to science have been characteristic to the broad trend of "postmodernism", which flourished in the late twentieth century philosophy in France and in the US departments of literature and cultural studies. Jean-Francois Lyotard made this term fashionable in 1979 by

---

35 See Niiniluoto (1996), (1999a), 242-251.

arguing that the end of the Enlightenment's "grand narratives" has opened a new "postmodern condition". Some of its proponents were disappointed idealists and Marxists, who reacted to the apparent decline of modern values (such as truth, beauty, and justice). Some of them came from the "Continental" schools of post-structuralism, semiology, deconstruction, and social constructionism. These thinkers did not conduct systematic studies about the notion of truth in the style of analytic philosophers, but their outlook was influenced by the conception of language as the universal medium, which denies the possibility of looking one's language from outside – so that semantics is ineffable. In contrast, Tarski's semantical approach accepted the conception of language as a re-interpretable sign-system, shared also by Boole, Peirce, Hilbert, Husserl, later Carnap, and Hintikka, while the universal medium conception was assumed by Heidegger and Gadamer.[36] Examples of the influence of the universal medium view include Lyotard's incommensurable language games and impossibility of metalanguage, Derrida's statement that there is nothing outside language, Baudrillard's sign's without referents, and Rorty's anti-representationalism (cf. Chapter 4 above). From such standpoints, even the formulation of the correspondence theory of truth becomes impossible, and one ends up with linguistic and alethic relativism.

Even before French postmodernism, there was a trend in Western Marxism where nature as an objective ontological category disappeared and society as "second nature" was regarded as merely apparent "pseudoconcrete totality". This view was influenced by Hegel's idealist doctrine of nature as the otherness of the spirit and the 1923 work of Georg Lukács, who in self-criticism in 1967 points out that his mistake was to confuse alienation and reification.[37] The so called "capital logicians" (Hans-Jørgen Schanz) argued against objective truth by claiming that money as capital deforms truth (in the same way as matter bends light waves in Einstein's general relativity). In Finland, this discussion led to a hot debate whether objective

---

36 According to Hintikka (1997), the universal medium view was also supported by Frege, Russell, Wittgenstein, and Quine. For Husserl, Heidegger, and Gadamer, see Kusch (1989).

37 See Lukács (1972), xxiii.

journalism is possible at all.[38] The realist solution to social ontology is to treat society as a part of Popper's world 3: it is human-made but real, so that statements about social structures and phenomena may be true or false in the sense of the correspondence theory.[39]

The Post-Truth Era

Oxford University Press and *The Economist* chose *post-truth* as the word of the year 2016. This was motivated by the Brexit referendum in Britain and the election of Donald Trump as the President of the United States. In both cases these elections were influenced – perhaps decided – by lies e.g. about the UK payments to the EU, the alien origin of Barack Obama, and alleged crimes of Hillary Clinton. This misinformation was distributed to the social media by Boris Johnson, Trump's campaign office, and Russian trollies. The same reckless practice was followed by President Vladimir Putin, when he annexed Crimea to Russia in 2014. European right wing populists have strengthened their position by defame campaigns and conspiracy theories (cf. Chapter 21 below). *The Washington Post* counted that the busy tweeter Trump expressed false or misleading claims 30,573 times during his four years presidency – before his Twitter account was closed – thus averaging about 21 erroneous claims per day. An EU committee counted 15.000 public lies from Russia between 2014 and 2021. A number of web journals are transmitting disinformation, but Trump himself used the term *fake news* for negative information about him distributed by C*NN* and *The New York Times*. Another terminological novelty was *alternative fact*, used by the White House counsellor Kellyanne Convay in her defense of the flatly misleading description of the attendance numbers in Trump's inauguration. German linguists chose this this Orwellian term as the "pseudoword of the year 2017". In 2022, after a year from his defeat to Joe Biden, Trump still falsely insists that the elections were dishonest and he really won them. Russia's attack against Ukraine on February 24, 2022, proved once again the old truth that "the first casualty of war is truth". But lying is not restricted to politics and media: many customers were

38 See my contribution in Niiniluoto (1990).
39 See Niiniluoto (1984).

disappointed, when they learned that big car factories had cheated them about the carbon emissions.

According to *Oxford Dictionary*, the term *post-truth* refers to "circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief". This definition is not entirely successful, as it does not use the term "truth" and does not deny the existence of truth. The real opposites of truth are not emotion and opinion, which may be accidentally correct, but falsity, error, mistake, and lie. The intentional presentation of claims, which are false or known to be mistaken, is what call lying. The aim of this kind of lying is usually to mislead or deceive another person. The philosopher Harry Frankfurt has distinguished lies from *bullshit*, where the speaker is not interested in the difference between truth and falsity. The aim of such random and meaningless jargon is not always deception, but sometimes only play with words; yet its goal may be bluffing or confusing the adversary.[40]

Writing fictive novels is not lying, since the aim of art is not to deceive the reader – at least those who are able to distinguish fact and fiction. In fact, good entertainment - like documentary news – may be informative and refreshing. Skillful fictive novels and films may tell a lot about the real world and people (cf. Chapter 5 above) – at least more than the so-called "reality television", which has almost nothing to do with truth.

By the post-truth era we may mean circumstances, where intentionally spread lies and disinformation have received a significant role in human communication.[41] But, as many commentators have asked, has not this been the case always in human history? The cultural historian Yuval Noah Harari (2014) argues that the success of the human species is explained its capacity of imagination to create attractive fictive narratives, such as the holy books of great world religions. In politics, the history of propaganda, indoctrination, and cunning exercise of

---

40 See Spicer (2017) for bullshit in business organizations.

41 In a book published in Finnish in 1989, I suggested that the proportion of true messages in the communication channels of the noisy information society is decreasing (even though the number of scientific publications is exponentially growing). But I did not guess that this would happen by the increase of fake news and other falsities.

power extends from the antiquity and Machiavelli via Adolf Hitler to contemporary populist leaders. But a dramatic new feature is the unprecedented ability of the social media to share fake news and bullshit.

Lying is not a novel discovery. Self-conscious human beings have this ability, which probably is lacking in other animals.[42] Already the Law of Moses in the *Old Testament* includes the eight commandment: "You shall not bear false witness against your neighbor". Truth was one of the three main values (with goodness and beauty) for Plato, and honesty was a noble virtue for Aristotle. The sophist Gorgias argued that truth is not accessible, but Thomas of Aquino regarded error as a sin. A witness in court vows "to tell the truth, the whole truth, and nothing but the truth", and perjury is a crime. The most rigid general view was advocated in the Lutheran Immanuel Kant's deontological ethics, where lying is absolutely forbidden in all circumstances - even when a lie could save a friend from a chasing murderer. In contrast, the Catholic church developed more flexible principles, which excuse helpful and playful "small white lies", while malicious lies for harming another person belong to mortal sins. In Jesuit handbooks, justified lies included silence or partial truths as answers to too personal or intimate questions.

Lying in the media is typical of the post-truth era. Traditional press and news in radio and television have been committed to responsible freedom of speech. In many countries the journalists have accepted ethical principles and guidelines, where truth and honesty have a crucial place. In Finland such an ethical code was first approved in 1957. This demand of honesty is parallel to the research ethics in science, which helps to guarantee the truthfulness or truthlikeness of scientific publications (cf. Chapter 20 above). During the last couple of decades the situation has changed, when the internet provides unlimited access to information without quality control. In contrast to great expectations, social media has become a channel of anonymous hate mail where false messages are shared more rapidly than true ones. Facebook, Instagram, and Twitter

---

42 For the fascinating history of lying, see Phillips (2019). But my worry about the present post-truth era is not diminished by the fact that already our ancestors were speaking rubbish.

transmit misleading trolls and in the future also fake videos, which influence the behavior of readers and voters. This development is strengthened by AI algorithms which influence the users of Google and Facebook. The political opportunity is taken over by political-commercial channels and fake media, which do not commit themselves to journalistic ethics (e.g. Breitbart).

It is surprising how receptive the audience is to misleading information – not only in totalitarian societies with oppressed citizens but also in liberal democracies. Fake media would not survive without readers, who have lost their trust in scientific experts. One of the influencing factors is the human propensity to biased thinking. Another is the so called "false balance", where research-based facts and purely subjective opinions are presented in television as if having equal weight. Sometimes the anti-science attacks are mobilized by companies with economic interests: tobacco companies tried to dement the research results about the causal connection between smoking and cancer. Climate skeptics and anti-vaxxers seek support from pseudo-publications. This kind of confusion about truth and reliability is reflected in the political life: to be caught of lying does not have consequences to the unscrupulous and power-seeking political leaders, when enthusiastic supporters are willing to tolerate untruths and deceptions which are believed to be beneficial for their own purposes. This has been seen in the behavior of pro-Trump followers who were ready to violate American democracy and to occupy the Capitol building in Washington on January 6, 2021.

Who Lost the Truth?

An objection to the notion of "post-truth era" questions whether there ever existed a golden age of truth-telling. As we have noted, this would indeed be a naïve interpretation of the actual course of world history. Therefore, it is better to relate this notion to *truth as a normative ideal*, typical of the Enlightenment period starting in the 18th century but roots in the ancient Greece. It is the withering of this ideal which we can now observe as marking "post-truth". But this also suggests that the mentioned period did not start with Trump and Putin: truth has been a casualty of the battles that are metaphorically called

cold war, information war, and trade war. The legitimacy of objective truth and the epistemic credibility of science in relation to other belief systems were the main issues of the Science Wars in the 1990s.

The erosion of truth of course has many interacting causes or contributing factors, like populist politics as a reaction to globalization, digitalization, internet, social media, and distrust in authorities and scientific expertise.[43] But, from the general perspective, it is also important that we philosophers look in the mirror and ask whether philosophy bears some responsibility for the gradual loss of truth. Such influence would not be an accident, since there are many examples of the cultural and political influences of philosophy. The principles of Aristotle's poetics are still followed by the Hollywood scriptwriters. Philosophers of the Enlightenment inspired the Declaration of Independence of the United States in 1776 and the French revolution in 1789. It may be disputed whether Nietzsche's superman really had an impact on Hitler. Karl Marx raised the political movement of the working class, but he cannot be blamed for Stalin's terror.

As we have seen in the first section of this chapter, philosophers have debated lively about theories of truth for more than two millenniums without reaching any consensus. But, with Tarski's achievement, for a realist there is a satisfactory correspondence notion of truth which is worth fighting for. The need of certified medical information and vaccines for the corona pandemic has strengthened the public trust in research – in spite of the laud anti-vaccine conspiracy theories. The public consternation for the talk about "alternative facts" and obvious lies by Trump's team indicates that the ideal of objective truth is still alive among educated people around the world. On the other hand, these relativist phrases sound like quotations from Nietzsche's perspectivism, and it is no wonder that Nietzsche has been regarded as a precursor of the post-truth phenomenon.[44] As Nietzsche was one of the main influences of the French postmodernism, philosophers like Daniel Dennett and Lee McIntyre (2018) have argued that the postmodernist

---

43 Putin's war propaganda does not have philosophical premises.

44 Spencer (2018) argues that Nietzsche and Heidegger are dangerous philosophers because of their influence on the far right.

challenge of the modernity and Enlightenment has prepared the way to the disregard of truth.

In my view, the link between postmodernism and post-truth is real but indirect. No one beliefs in a direct causal relation: most probably Trump has never read Foucault, Derrida, Rorty, or Feyerabend. It has been claimed that in reality these thinkers were not against truth, but in their own ways supported the rational pursuit of objective knowledge in science.[45] But in fact the impact of philosophical ideas is normally mediated by their reception and interpretation among different audiences: in the case of postmodernism, this included the popularity of social constructionism among social scientists and the public message in journals that truth is an outdated ideal.[46] In cultural history, the "memes" are transmitted – besides linguistic messages - as habits and attitudes. The typical attempts to dissolve distinctions between truth and falsity, knowledge and opinion, expertise and subjective experiences, fact and fiction, documents and entertainment, facts and values can be seen especially in television and social media. Donald Trump, with a background as a ruthless businessman, was no doubt infected by the "spirit of the age" in his role as the stale host of the reality tv-series *The Apprentice*.

These remarks apply also to American pragmatism, which is largely known in the public by the vulgar identification of true and useful. This may be known even to Trump – at least he applies it in an extremely subjective form: everything which improves his political goals is true, while everything which is against his interests is just false or "fake news". Sami Pihlström (2021), Ch. 1, known for his careful scholarship on pragmatism, has taken this problem seriously by asking whether the pragmatist notion of truth is too soft so that it allows a slippery slope from James to Rorty and Trump. He challenges Richard Rorty's account truth in George Orwell's 1949 dystopic novel *Nineteen Eighty-Four*, where the party defines what is true (e.g.

---

45 See Nora Hämäläinen's (2019) as such interesting defense of postmodernism. Bruno Latour, in an interview in the New York Times in 2018, has regretted (perhaps sincerely) that he never intended to undermine the public authority of science.

46 In Finland, in a review of Robert Pirsig's *Zen and the Art of Motorcycle Maintenance* in 1987, a young postmodernist writer Markku Eskelinen amusingly suggested that I should get rid of my "truth addition".

war is peace, freedom is slavery, ignorance is power, and 2 + 2 = 5). With reservations against Rorty's statement ""if we take care of freedom, truth can take care of itself",[47] Pihlström defends a responsible form of pragmatic realism, which gives space for a sincere search for truth.

The hope that the post-truth era would be a short period of disturbance is threatened by the continuation of war in Ukraine and the horror picture of Trump's wish to be re-elected in 2024. More generally, we should fight against resistance to knowledge, were public opinion is uncritically shaped without good and proper epistemic reasons.[48] What is needed as a remedy is continuous support of responsible journalism and scientific research as strongholds of honesty and truthfulness, education of critical thinking and media literacy, good practices in traditional and social media, source criticism and fact checking, research integrity and trust in scientifically validated knowledge. The citizens should appoint by democratic elections wise and responsible political leaders. We philosophers should provide a clear analysis of the objective notion of truth and a restoration of its valuation.

Note. The first two sections of this chapter use some material from Niiniluoto (2006) and (2013). The last section is based on my articles "Kuka hukkasi totuuden?", *Tieteessä tapahtuu* 37 (2) (2019), 9-15, and "Pitääkö olla huolissaan totuudesta?", *Tieteessä tapahtuu* 38 (3) (2020), 56-64.

Literature

Armstrong, David M. (2004). *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Barnes, Barry & Bloor, David (1982). "Relativism, Rationalism and the Sociology of Knowledge". In Martin Hollis & Steven Lukes (eds.), *Rationality and Relativism*. Oxford: Blackwell, 21-47.

Bloor, David (1976). *Knowledge and Social Imagery*. Chicago: University of Chicago Press. (2nd ed. 1991)

47 Rorty (1989), 176.

48 See Wikforss (2019) for biased human thinking.

Cappelen, Herman & Hawthorne, John (2009). *Relativism and Monadic Truth*. Oxford: Oxford University Press.

Carnap, Rudolf (1942). *Introduction to Semantics.* Cambridge, MA: Harvard University Press.

Chang, Hasok (2022). *Realism for Realistic People: A New Pragmatist Philosophy of Science*. Cambridge: Cambridge University Press.

Collins, Harry (1981). "Stages in the Empirical Programme of Relativism", *Social Studies in Science* 11, 3-10.

Danto, Arthur (1973). "Nietzsche's Perspectivism". In Robert Solomon (ed.), *Nietzsche: A Collection of Critical Essays*. New York: Anchor Books, 29-57.

Dummett, Michael (1978). *Truth and Other Enigmas*. London: Duckworth.

Feyerabend, Paul (1975). *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: New Left Books.

Feyerabend, Paul (1987). *Farewell to Reason*. London: Verso.

Garcia-Carpintero, Manuel & Kölbel, Max (eds.) (2008). *Relative Truth*. Oxford: Oxford University Press.

Hämäläinen, Nora (2019). *Är Trump postmodern?* Helsingfors: Förlaget.

Harari, Yuval Noah (2014). *Sapiens: A Brief History of Mankind*. London: Harvill Secker.

Harding, Sandra (ed.) (1987). *Feminism and Methodology*. Bloomington, IN: Indiana University Press.

Hautamäki, Antti (2020). *Viewpoint Relativism: A New Approach to Epistemological Relativism based on the Concept of Points of View*. Cam: Springer.

Hempel, Carl G. (1935). "On the Logical Positivist Theory of Truth", *Analysis* 11, 49-59.

Hintikka, Jaakko (1997). *Lingua Universalis vs. Calculus Ratiocinator: An Ultimate Presupposition of Twentieth-Century Philosophy*. Dordrecht: Kluwer.

Horwich, Paul (1990). *Truth*. Oxford: Blackwell.

Kirkham, Richard L. (1992). *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press.

Krausz, Michael & Meiland, Jack (eds.) (1982). *Relativism: Cognitive and Moral*. Notre Dame, IN: University of Notre Dame Press.

Kuhn, Thomas (1970). *The Structure of Scientific Revolutions*, 2nd ed. Chicago: The University of Chicago Press.

Kusch, Martin (1989). *Language as Calculus vs. Language as Universal Medium: A Study in Husserl, Heidegger and Gadamer*. Dordrecht: Kluwer.

Kusch, Martin (1991). *Foucault's Strata and Fields*. Dordrecht: Kluwer.

Latour, Bruno & Woolgar, Steve (1979). *Laboratory Life: The Social Construction of Scientific Facts*. London: SAGE.

LePore, Ernest (1983). "What Model Theoretic Semantics Cannot Do?", *Synthese* 54, 167-188.

Lukács, Georg (1972). *History and Class-Consciousness*. London: Merlin Press.

McIntyre, Lee (2018). *Post-Truth*. Cambridge, MA: MIT Press.

Meiland, Jack (1977). "Concepts of Relative Truth", *The Monist* 60, 568-582.

Nicholson, Linda (ed.) (1990). *Feminism/Postmodernism*. London: Routledge.

Niiniluoto, Ilkka (1984). "Realism, World-Making, and the Social Sciences". In *Is Science Progressive?* Dordrecht: D. Reidel, 211-225.

Niiniluoto, Ilkka (1987). *Truthlikeness*. Dordrecht: D. Reidel.

Niiniluoto, Ilkka (1990). "Realismi, praksismi ja relativismi". In *Maailma, minä ja kulttuuri*. Helsinki: Otava, 43-87.

Niiniluoto, Ilkka (1996). "The Relativism Question in Feminist Epistemology". In Lynn Hankinson Nelson & J. Nelson (eds.), *Feminism, Science, and the Philosophy of Science*. Dordrecht: Kluwer, 139-157.

Niiniluoto, Ilkka (1999a). *Critical Scientific Realism*. Oxford: Oxford University Press.

Niiniluoto, Ilkka (1999b). "Tarskian Truth as Correspondence - Replies to Some Objections". In Jaroslaw Peregrin (ed.), *The Nature of Truth - If Any*. Dordrecht: Kluwer, 91-104.

Niiniluoto, Ilkka (1999c). "Theories of Truth: Vienna, Berlin, and Warsaw". In Jan Wolenski & Eckehat Köhler (eds.), *Alfred Tarski and the Vienna Circle*. Dordrecht: Kluwer, 17-26.

Niiniluoto, Ilkka (2001). "Information, Meaning, and Understanding". In Lars Lundsten, Arto Siitonen & Bernt Österman (eds.), *Communication and Intelligibility*. Helsinki: Societas Philosophica Fennica, 43-54.

Niiniluoto, Ilkka (2003a). "Science as Collective Knowledge". In Matti Sintonen, Petri Ylikoski & Kaarle Miller (eds.), *Realism in Action*. Dordrecht: Kluwer, 269-278.

Niiniluoto, Ilkka (2003b). "Carnap on Truth". In Thomas Bonk (ed.), *Language, Truth, and Knowledge: Contributions to the Philosophy of Rudolf Carnap*. Dordrecht: Kluwer, 1-25.

Niiniluoto, Ilkka (2004). "Tarski's Definition and Truth-Makers", *Annals of Pure and Applied Logic* 126, 57-76.

Niiniluoto, Ilkka (2006). "The Poverty of Relative Truth". In Tuomo Aho & Ahti-Veikko Pietarinen (eds.), *Truth and Games: Essays in Honour of Gabriel Sandu*. Helsinki: Societas Philosophica Fennica, 165-174.

Niiniluoto, Ilkka (2013). "Against Relative Truth". In Kevin Mulligan, Katarzyna Kijania-Placek & Thomas Placek (eds.), *The History and Philosophy of Logic: Essays in Honour of Jan Wolenski*. Houndmills: Palgrave Macmillan.

Niiniluoto, Ilkka, Sintonen, Matti & Wolenski, Jan (eds.) (2004). *Handbook of Epistemology*. Kluwer: Dordrecht.

Phillips, Tom (2019). *Truth: A Brief History of Total Bullsh*t*. London: Wildfire.

Pihlström, Sami (2021). *Pragmatist Truth in the Post-Truth Age: Sincerity, Normativity, and Humanism*. Cambridge: Cambridge University Press.

Putnam, Hilary (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.

Rorty, Richard (1989). *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.

Russell, Bertrand (1940). *An Inquiry into Meaning and Truth*. London: Allen & Unwin. (new ed. 1962)

Siegel, Harvey (1987). *Relativism Refuted: A Critique of Contemporary Epistemological Relativism*. Dordrecht: D. Reidel.

Siegel, Harvey (2004). "Relativism". In Niiniluoto, Sintonen & Wolenski, 747-780.

Spencer, Richard (2018). *Dangerous Minds: Nietzsche, Heidegger and the Return of the Far Right*. Philadelphia: The University of Pennsylvania Press.

Spicer, André (2017). *Business Bullshit*. London: Routledge.

Swoyer, Chris (1982). "True For". In Krausz & Meiland, 84-108.

Tarski, Alfred (1944). "The Semantic Conception of Truth and the Foundations of Semantics", *Philosophy and Phenomenological Research* 4, 341-375.

Tarski, Alfred (1956). "The Concept of Truth in Formalized Languages". In *Logic, Semantics, Metamathematics.* Oxford: Oxford University Press, 152-278.

Vattimo, Gianni (2011). *A Farewell to Truth*. New York: Columbia University Press.

Wikforss, Åsa (2019). *Alternative Facts: On Knowledge and its Enemies*. Stockholm: Grand Agency.

Wolenski, Jan (1989). *Logic and Philosophy in the Lvov-Warsaw School.* Dordrecht: Kluwer.

Wolenski, Jan & Simons, Peter (1989). "De Veritate: Austro-Polish Contributions to Truth from Brentano to Tarski". In Klemens Szaniawski (ed.), *The Vienna Circle and the Lvov-Warsaw School*. Dordrecht: Kluwer, 391-442.

Society

Chapter 17: Promoting Sustainable Development

Human Beings and Nature

The ascendants of *Homo sapiens* have used wooden and stone tools in their adaptation to the environment for more than 2 million years: as Benjamin Franklin and Karl Marx stated, "man is tool-making animal". As cultural animals evolved from the material world, human beings are dependent on nature - they have feared and worshipped nature and facilitated interaction with nature by means of technical tools. About 10 000 years ago a leap in the evolution of the humanity occurred, when our ancestors started to actively influence the course of nature in agriculture, so that "pure" nature was transformed into a cultivated environment with material artefacts. The rise of new human cultures increased man's positive freedom by bringing about novel tools and professions, villages, and towns. With the advent of the skill of writing, the ancient world created philosophy and science, which started to flourish in the scientific revolution of the early modern age. Simultaneously with the 18th century Enlightenment, the "age of utility" resulted in a technological revolution, whose later stages can be seen in the 20th century industrial society and in the 21st century post-industrial information society.

We should not reify and personify technology as a monster and blame it for our problems (cf. Chapter 20 below). This was the irrational strategy of the Luddites of the 18th century, who attacked machines without realizing that it was men who were running the machines and factories. Today it is *we* who are maintaining and advancing technology for our own purposes. Environmental damage and catastrophes are caused both by the rich and the poor people, agriculture as well as industry, ways of life of whole nations and patterns of behavior of individual consumers.

In the cultures of the East and the ancient Greece, the ideal relation between man and nature was harmonic and peaceful co-existence. In contrast, the Jewish and Christian traditions understood man as "an image of God" which is higher than other creatures with a license to rule or master them. Modern "Baconian" science also emphasized that human beings are not any more in the mercy of natural forces, but with the help of science-based technology men can command nature to fulfill their own purposes. In idealistic metaphysical systems, nature is a secondary category beside or below the subjective or objective mind. In Cartesian dualism, animals are mechanical machines without consciousness. The tradition of romanticism thought that the main task of human beings, as gardeners of nature, is to protect and admire God's creations and their beauty. In naturalism, human beings are conceived as parts of nature, developed from other organisms by biological and cultural evolution.

The influences of human actions on the natural environment were visible already during the old age in the destruction of Mediterranean forests. Statutes against the urban pollution of air were issued already in the 15th century. Industrialization motivated demands about the protection and conservation of nature since the 18th century, and the Yellowstone National Park in the United States was founded in 1872. However, it was only after World War II that a world-wide awareness emerged: men have intentionally and non-intentionally burdened nature, ruthlessly exploited non-renewable natural resources, and polluted air and seas. The leaders of the Frankfurt School, Max Horkheimer and Theodor Adorno argued in 1947 that the urge to dominate nature by instrumental reason has led the Enlightenment to its self-destruction (cf. Chapter 8 above). The World Wildlife Fund (WWF), which was to become the world's largest environmental organization, was founded in 1961.[1] UNESCO published in 1962 its recommendation concerning "the safeguarding of beauty and character of landscapes and sites". Rachel Carson's *Silent Spring* in 1962 was a warning cry against the deaths of birds by toxic agrochemicals (DDT). In 1967 the medieval historian Lynn White

[1] See Schwarzenbach (2002) for the first 50 years of WWF in its fight to stop "the degradation of our natural environment".

Jr. published his famous article, blaming the Western religious tradition for the "ecological crisis". Greenpeace, favoring direct action, was founded in 1971, and the green political movement started to emerge in the 1970s. The responsible futurists in the Club of Rome published in 1972 their work *The Limits of Growth*, which showed the mistake in thinking that nature is an unlimited resource to be exploited for human purposes. Thus, in spite of its glorious cognitive progress, and its multiple applications for the benefit of the humanity, science has also led to harmful effects on the natural environment through science-based technology, industry, and everyday consumption. Georg Henrik von Wright summarized this growing concern in The Preface to his collection *Humanismi elämänasenteena* (Humanism as an Attitude towards Life, 1981) with the following words:

> The more I have thought about it, the more I have led to question not only the impact of the scientific-technological revolution on human happiness and welfare but also on the form of life and society based on science and technology.

Sustainable Development

With the humankind's drastic influence on the globe, we live in a new epoch of "anthropocene". It is evident that for solving the "wicked" global problems caused by human activities we need research and education, backed up by international co-operation and agreements. The central keyword in this area is "sustainable development", introduced to the public in *Our Common Future* (1987), the Report of the World Commission on Environment and Development.[2] In spite of the signs of a world-wide ecological crisis (pollution of air and water, wastes and toxic chemicals, soil eroding, deserts expanding, the ozone shield diminishing, population exploding, gap between poor and rich widening), the Commission led by Norway's former prime minister Gro Harlem Brundtland was optimistic that international co-operation with joint efforts can secure the

[2] This term was first used in the *World Conservation Strategy* (1980) of the International Union for the Conservation of Nature and Natural Resources (ICUN). The publication of the book series *State of the World* was started in 1985.

continuing development of the humanity. The more pessimistic critics of the Commission, on the other hand, have argued that the program of sustainable development has not yet sufficiently analysed the economic and political causes of the ecological crisis, and assumes the Western paradigm of progress and linear growth.[3] In spite of the good intention that the community of nations "is, and must be, able to secure the continuing development of mankind", the report did not make sufficiently explicit how its recommendations are based upon factual and value premises. The overall attitudes about this issue resemble the late 19th century reactions to the disadvantages of capitalist economy:[4] (i) the "right wing" industries and governments irresponsibly without bad conscience continue to harm natural and social environment within the present way of life, (ii) the "moderates" are convinced that economic growth can continue when nations reach reasonable agreements – as highlighted in the Kyoto agreement on carbon dioxide wastes (1997) and the World Summits in Rio de Janeiro (1992) and Johannesburg (2002), (iii) the "radical left" regards sustainable development as insufficient and requires a new world order with "degrowth" measures.[5]

Sustainable development was defined by Brundtland's Commission as a dynamic process which "meets the needs of the present without compromising the ability of future generations to meet their own needs". A virtue of this definition is its concern for the future: we are not allowed to be selfish and short-sighted and look only after our own momentary profits, but our children and their descendants should also have sufficient conditions of living. A weak point is the reliance on the concept of "need", which is left without a proper analysis. In the kernel of the political debates about sustainable development, one can find the tension between the necessities of life (energy, nutrition, housing) and the new "needs" and practices of consumption created by rapid technological development.

---

[3] See Cooper and Palmer (1992).

[4] Both capitalists and socialists embraced naïve technological optimism by failing to realize the potential destruction that uncontrolled industry may cause to nature.

[5] Among philosophers, the third position has been advocated by Thomas Wallgren.

The report also overlooks the new developments in environmental ethics.

Brundtland's Commission was mainly concerned with ecological threats, but it recognized that human and social behaviour in both rich and poor countries have dramatic effects on our environment. The Johannesburg Summit identified three pillars of sustainable development: economic development, social progress, and environmental protection. Today UNESCO advocates a very broad definition where sustainability covers, among other things, the management and protection of natural resources, climate change, global warming, biodiversity, health, rural development and urban planning, poverty reduction, corporate responsibility, cultural diversity, education for all, free flow of information, human rights, gender equality, and peace. If one is worried that such an extensive list of virtually all important human goals loses the focus on environmental issues, the holistic approach to sustainability has the merit of recognizing the interrelations between ecology, economy, culture, and society.

Towards Rational Environmental Policies

How can scientific research and education best promote sustainable development? It is clear that all areas of natural and social sciences - especially when they work together in multidisciplinary projects - can give their contribution to this goal. Knowledge about the state and behavior of natural and social systems gives us evidence-based information about the prospects of rational environmental and social planning. It is also important to know what probable consequences our alternative actions will have. A systematic methodology for outlining such different scenarios has been developed in futures studies (cf. Chapter 12 above). But empirical and theoretical expert knowledge alone is not a sufficient basis of environmental and social planning and decision-making; we also need a clear value-based vision of desirable futures. Such value questions cannot be reduced to the empirical study of human needs, since they always include a personal commitment to what we regard as desirable or valuable. But values are not arbitrary subjective choices, either, since questions about values and norms can be rationally debated in philosophical axiology.

Environmental ethics is a branch of philosophy which considers questions about the intrinsic and instrumental value of nature. More generally, our ethical views concern the conditions of good human life and just society.

In sum, without factual knowledge provided by science environmental engagement may remain fanciful wishful thinking, and without enlightened ethics environmental studies have the danger of sliding to the admiration of technocratic expertise. Thus, a model for the interplay between research-based facts and human values in the domain of sustainability is needed:

(1) environmental research + environmental ethics = rational environmental policy.

This "equation" can be compared to the program of evidence-based medicine, which combines the value of health with the rational means of maintaining and improving this aim. It thus provides a link between belief rationality, instrumental rationality, and value rationality (cf. Chapter 8 above). *Evidence-based environmental policies* can be formulated by using G. H. von Wright's (1963) notion of *technical norm*, which expresses a relation between ends and means:

(2) If you want A, and you believe that you are in situation B, then you ought to do X

(cf. Chapter 11 above on design science).[6] The norm (2) is true, if X is a necessary condition for achieving the end A in situation B, i.e. X is a (necessary) cause of A in situation B.[7] Therefore, the technical norm (2) can be justified by giving scientific evidence for the causal claim "X is a cause of A given B".

For environmental research, this schema defines three important tasks: investigation of the state B, study of the causal connection X & B → A (i.e. X causes A in situation B), and

---

[6] My first formulation of the model of rational environmental policies was given in 1992 in the 3rd cooperation Conference of Circumpolar Universities in University of Lapland, Rovaniemi (see Niiniluoto, 1993, 1994). It should be applicable equally well to arctic tundra, Brazilian rain forests, urban California, the arid zone of Africa. or the oceans.

[7] Generalizations of (2) allow that X is a sufficient or probabilistic cause of A given B.

identification of possible side effects of X. For environmental ethics, this schema gives the task of analyzing the end A, but also the moral evaluation of the means X is relevant: if the necessary conditions of your end A are morally unacceptable, you should revise your valuation of A. The technical norm (2), as a *conditional recommendation of rational action*, yields a practical inference from the premises that A is accepted as a desirable aim and B is believed to be the prevailing situation to the "technical ought" X – and such rational commitments may be valid for individual, collective, and corporate agents.[8]

In environmental research, the situation B is often described by a value-laden term like "polluted", and the aim A is to get rid of this undesirable state. For an overall description of B, the Index of Sustainable Economic Welfare (ISEW) combines the amount and distribution of consumption with environmental factors like exploitation of natural resources, reduction of cultivated land, pollution of air and water, and warming of climate. Is this compatible with the demand that science should be objective or value-free? A positive answer can be sought by making a difference between general or "thin" value concepts (like good and right) and "thick" concepts (like polluted and cruel). The latter supervene on descriptive value-free terms: for example, to say that a lake is "polluted" expresses a fact (i.e. it is contaminated by poisonous chemical substances or micro-organisms, and their amount exceeds a preassigned limit or threshold). In a community which accepts this operational definition, the term "polluted" has an objective meaning, which combines a factual statement and negative valuation of this fact. Value-laden terms are also used in *environmental law*, which contains precautionary principles for the prevention of the "pollution" and "degradation" of natural environments and liabilities for parties who have caused such

[8] The formulation of the technical norm (2) presupposes a division of relevant factors into those which are held constant (B) and those which we can and may manipulate (X). For example, the environmental programs of the European Union have proposed drastic changes in industry, energy, agriculture, transportation, tourism, and taxation of carbon dioxide, but accepts free trade as the supreme principle.

harms or "contaminations" for air, soil, ground water, and ecosystems (like lakes and forests).[9]

The schema (2) can be illustrated by an example of lake management. Enäjärvi lake in Uusimaa, near Helsinki, is a famous for its bad health in the summer: eutrophication with algal blooms and even fish kills.[10] Research on water management tries to establish technical norms which express means to prevent the pollution by toxic algae. The state (B) of Enäjärvi has a history from the irresponsible dumping of waste waters from nearby fields and villages. The lake bottom is covered by a phosphorous sediment, whose fluffing by fish (especially roach and bream) boosts eutrophication. The aim A is generally accepted: improve the condition of the lake so that its fauna is healthier, smell is decent, and recreational activities (like swimming and fishing) are again possible. Specific research-based methods X for reaching this goal include elimination of wastes, oxidation, and mass removal of unwanted coarse fish by nets.

Sustainability Science

Following the model (2), let us review the main tasks of promoting sustainability by environmental research and education.

First, the methods of science provide tools for studying and describing the current *state of nature* and its development. Systematic observations and measurements in research stations help to follow e.g. the quality of water, flora and fauna in lakes and oceans. Atmospheric observations may concern the temperature and quality of air, the amounts of ozone and carbon dioxide, or the density of pollution and toxic substances. Time series of such observations may tell alarming news about changes in nature.

Secondly, environmental research may focus on *lawlike patterns of change* in natural phenomena, e.g. the interaction of

[9] Since 2010 citizen's movements have campaigned with the UN that "ecocide" (i.e. serious destruction of large areas of natural environment as a consequence of human actions) should be recognized as a crime by international law.

[10] My family has spent summer holidays near this lake since the 1950s.

atmosphere, land, and seas.[11] With knowledge about such laws of nature, it is possible to build theoretical, often mathematical models for explaining and predicting the temporal development of natural systems. The basic science in this area is ecology, which investigates the relationships between populations of organisms and their natural environment.[12]

Thirdly, a crucial challenge of multidisciplinary environmental research is to bring the *human agent* into the systems in consideration. What harmful changes in nature are due to human actions and interventions? What kinds of technologies might be helpful in the protection of nature? What cultural habits, social practices, and economic arrangements best support the goal of sustainable development? Here natural sciences, technological research, biology, agricultural and forest sciences, medicine, urban studies, cultural studies, economics and other social sciences have to work together. In this way, the subject matter of research is broadened to the interaction of man and nature, especially the rational utilization of natural resources and its cultural, social, and economic aspects and conditions. Recently the name *sustainability science* has been used for this multidisciplinary research program.

Fourthly, as far the concept of sustainable development contains other (*economic and social*) pillars besides the environmental perspective, medicine, the humanities, education, psychology, law, and social sciences are directly relevant to their study. Systematic inquiry can follow the current state of society and seek knowledge about its temporal changes. Sustainable development of a society requires that its economic basis is secure, but in the global world this goal is conditioned by international co-operation and competition. In spite of its importance, economic success has to be balanced with considerations of human welfare and social justice. On the other hand, economic stability in the long run depends on the ways in which natural resources are used and cultivated. Thus, we see again that research on social sustainability remains one-

---

[11] Markku Kulmala from the University of Helsinki, the world's most cited scientist in geosciences, has investigated the interaction between air (esp. aerosols) and forests.

[12] Ilkka Hanski (1999) has studied the development of metapopulations of butterflies.

sided or incomplete if it does not bring to its focus the interactions between nature, economy, culture, and society.

In education for sustainable development, multidisciplinary environmental research and ethics are important subject matters. Introduction to cognitive and moral attitudes about natural and social environment are needed in schools - and already at home and Kindergarten. International organizations, national governments, business firms, the media, and NGOs have all their roles in our common political and moral task of sustainability. Universities and institutions of higher education have particularly important tasks. As we have seen, all scientific disciplines have a potential for increasing our understanding of the conditions of sustainability in the wide sense. In a more specific sense, UNESCO has created programs, networks, and chairs devoted to problems of sustainable development. Together with the International Science Council (ISC), it supports the research network Future Earth. The Baltic Sea Project is a fine example of regional co-operation in research and teaching. The University of Helsinki has established - besides its traditional faculties of bioscience, agriculture and forestry, and social sciences - a research network of Environmental Research (HERC), a chair in environmental policy, and a teaching program on environmental problems for students from all faculties. In 2018 HERC was integrated to the new Helsinki Institute of Sustainability Science (Helsus).[13]

Environmental Ethics

The traditional domain of ethics consists of relations between human beings or persons: Aristotle's ethics is based on human virtues, Immanuel Kant's deontological ethics appeals to the categorical imperative, "treat humanity never merely as a means to an end", and John Stuart Mill's utilitarianism recommends actions which bring about the maximal amount of human happiness. Environmental ethics extends this domain to include animals, living beings, and nature as a whole. We shall see that such an extension can be defended in some respects, but its extreme forms are implausible.

---

[13] A key figure in this development was rector Jari Niemelä.

A positive concern for nature presupposes that the natural environment has some *value*, which can serve as a basis of human ethical attitudes, commitments, and responsibilities for nature. Ethical concern (combined with environmental research) can then manifest itself in practical action in *environmental politics*, especially in the conservation and protection of nature. Environmental ethics in its classical period from the 1970s to the 1990s - the journal *Environmental Ethics* started to appear in 1979 - was primarily occupied with the axiological question about the value of nature.

Clifford Hooker (1992) classifies rival ethical positions by the notion of *responsibility*, as exemplified by the schema:

(3) C takes responsibility for Y to D for reasons Z,

where substitutions to C, Y, D, and Z give answers to the questions: who, for what, to whom, and why.

One model takes man to be a *despot*, who rules nature with narrow and arbitrary self-interests without any responsibility. This seems to be White's (1967) interpretation of the human task given in the *Genesis*: "multiply and subdue the earth and have dominion over it". But Francis Bacon understood better that "to be ruled, nature must be obeyed".

The *prudential user*, as advocated by John Passmore (1980), sees nature as a means of livelihood (hunting, fishing, mining, agriculture, forestry, industry), so that it is rational to protect nature for human economic purposes. Thus, the humans are responsible for themselves to themselves on non-ethical prudential interests. Passmore also mentions another form of utilitarianism which emphasizes our ethical responsibilities to other people, especially future generations who should have the possibility of cultivating nature for their own purposes. Brundtland's commission formulates this view by appealing to our responsibility for the needs of future generations.

Nature can be seen as a source of *aesthetic* and *emotional* experiences, such as beauty, health, and happiness. This is the basis of environmental aesthetics.[14]

The *humanist* approaches have concentrated on those features which distinguish humans from other animals, so that ultimately we are responsible for nature on the ground of

[14] See Sepänmaa (1986), Berleant (1992). Cf. Chapter 3 above.

human interests. But care of nature can be justified along these lines. A typical example in Finland was the Christian humanist and historian Zachris Topelius who founded in 1870s the first societies for the protection of small birds: care of animals taught children to love their fellow humans.

According to *pastoralism,* man is God's gardener or steward, whose prerogative is to cultivate and protect nature. This "stewardship ideology" is usually based on the religious doctrine that nature and its creatures are of divine origin, so that the humans are collectively responsible for nature to God. Many Eastern and Western religions wish to promote, in this sense, the harmony of human beings with the world. A variant of this view sees the world as a living organism, like Lovelock's Gaia or a feminine "Mother Nature".[15] But pastoralism can be formulated without religious or metaphysical assumptions: according to Robin Attfield (1991), we are responsible to all moral agents for the protection of the inherent value of people and other creatures.

Despots, prudentialists, aestheticians, and humanists all favor *anthropocentric* approaches which relate the value of nature to human interests. The official definition of sustainable development is likewise anthropocentric in the sense that it is concerned with the living conditions of the human species on earth. Thus, its value orientation is the instrumental or prudential utilization of nature for human purposes. This is one way of justifying environmental protection as a form of man's responsibility for nature. During the last decades, ecological movements and the new philosophy of environment have promoted *ecocentric* views which defend the intrinsic value of life and nature.

One of these biocentric approaches is the *animal rights* movement, started by Peter Singer (1975), which treats the individual members of other species as equally valuable as human beings. This view is usually based on the assumption that animals are sentient beings with the ability to feel pleasure and pain. But as Tom Regan (1984) points out, this does not mean that animals are "moral agents", who as self-conscious persons

---

[15] See Merchant (1983), who complains that modern science with its materialism brought about the "death of nature". Ecofeminism draws a parallel between nature and "wild" women (Plumwood, 1993).

can make moral judgments and choices; rather animals are "moral patients", whose rights are in fact human obligations to treat them well without causing unnecessary pain or harm.[16] Another important view was formulated by G. H. von Wright (1963, 50): even though he was most interested in "the good of man" (i.e. welfare and happiness) (cf. Chapter 19 above), he pointed out that the phrase "the good of X" can be non-metaphorically applied to all living beings, i.e. "who can meaningfully be said to be well or ill, to thrive, to flourish, be happy or miserable".

Another biocentric approach attributes intrinsic value to the *species* of plants and animals instead of their individual members. This is expressed by the concept of *biodiversity*. Declining biodiversity due to human activities (agriculture, industry, consumption, climate warming) is an alarming aspect of the ecological crisis: the UN nature panel IPBES warns that of the eight million species of plants and animals about one million are endangered within the next decades.

More radical *naturocentric* approaches attribute intrinsic value to the natural world as a whole: the "deep ecology" of Holmes Rolston III (1988) and the "ecosophy" of Arne Naess (1989). As a variant of deep ecology, Aldo Leopold's "land ethics" assign intrinsic value to untouched natural landscapes.

Instrumental and Intrinsic Values

Debates in environmental ethics are often couched in terms of the distinction between instrumental and intrinsic values. Practically speaking all agree that nature has instrumental

---

[16] Argentine granted unprecedented legal rights to a captive orangutan in 2015, but it is debatable what this really means in practice. As a natural extension of the concept of a human person, *legal personhood* is granted to companies, corporations, and societies, which have autonomous decision-making bodies. Such legal persons are moral agents with accountability and liabilities e.g. in economic and environmental matters. On the other hand, Saudi Arabia's decision to grant citizenship to Sophia the Robot in 2017 is questionable, since it is based on the speculative view that human-like robots really have intelligence, emotions, and consciousness. In the EU Civil Law Rules on Robotics in 2017 it was suggested that robots could have the special status of "electronic persons", but the implications of this new category are still unclear.

value for the humanity, and Passmore (1978) argues that rational environmental policies can be based on such human-centered values. The Brundtland report, which emphasizes sustainable development serving human needs, did not discuss the intrinsic value of nature, which has been a major philosophical theme in environmental ethics.

It is important to see that instrumental and intrinsic values do not exclude each other: truth is a basic value in science and beauty in the arts, but both of them can be useful e.g. economically. Thus, in a human value system, natural objects may have intrinsic value in themselves but also be instrumentally useful for other values. Further, it is significant that the intrinsic-instrumental divide is independent of the philosophical question whether values are human-made or in some objective sense pre-existing in nature. The latter view has been defended by many authors. Holmes Rolston III (1988) argues that the intrinsic value of environment should be "naturogenic" rather than "anthropogenic", which presupposes a strong form of value objectivism. Tom Regan (1988) defends the "inherent" value of animals which are "subjects-of-a-life", and for Paul Taylor (1986) respect for nature is based on the good of a living organism which allows its flourishing. Arne Naess' (1989) holistic ecosophy accepts the egalitarian intrinsic value for all species.

In my view, all values are human constructions in the Popperian world 3, so that there are no values generated by nature itself. Animals lacking conceptual thinking and intentionality do not satisfy the conditions for value creation - preferences and goal-directedness are not enough. Animals may have needs, emotions, wants, and interests, whose satisfaction has survival value for them, but to elevate them to some sort of "values" is an instance of what G. E. Moore identified as the naturalistic fallacy in evolutionary ethics. "Z is good for A" is not the same as "Z is a value for A" (e.g. human blood is good for mosquitos, mosquitos are good for birds …). Animals are moral patients rather than moral subjects, and "animal rights" are "immunities" in Hohfeld's sense: our duties to treat animals without unnecessary suffering. A forest has no "right" in itself to exist and to be protected, but we may have a moral or legal responsibility to respect a conserved area of land. The 2008 Constitution of Ecuador states that nature has "the right

to integral respect for its existence", but this means that all person and communities can demand in courts that Ecuadorian authorities enforce the restoration and regeneration of nature.[17]

Human-made intrinsic values need not be human-centered (anthropocentric), but may be biocentric or naturocentric as well. There would be no difficulty to extend Kant and Mill to all animals: do not treat animals merely as means to an end, and approve acts which increase the happiness among all animals. In my view, we have good reasons to defend the intrinsic the value of biodiversity and the health of ecosystems. But Richard Routley's "last person argument" does not prove the existence of objective values,[18] as e.g. Attfield (1991, 155) thinks, since even the last human person may give value to a future world without human beings: a world with peacock butterflies is better than without this species.

Intrinsic values are prima facie morally relevant as the aims of our actions and responsibilities, but the means for achieving them may be morally questionable. Intrinsic values in a value system may be hierarchically ordered: a farmer may value the diversity of animals and plants, but not so much as to restrict his interest for exploiting his land for the income of his family. Even if all animals may have some intrinsic value, still we prefer human beings to mosquitos (*pace* Naess). Similarly, when conflicts may rise with the intrinsic value of nature and its instrumental value (e.g. building a road through a forest with an endangered species like the flying squirrel), legal and political solutions may needed.

On the other hand, instrumental values are usually not morally binding. If Y is good for A, then a human agent B need not

---

[17] New Zealand granted in 2014 the Whanganui river the status of a legal person (see O'Donnell & Talbot-Jones, 2018; cf. footnote 14). In my view, this is a category mistake in ontology, based on the religion of the indigenous Maori people, which regards rivers as living beings. In the same way, Mississippi has been poetically described as "Old Man River", but environmental law should not be based on religion or poetry. There are other ways of securing the political and legal rights of the indigenous people to their land.

[18] Routley (later Sylvan) presents this argument against "human chauvinism", but leaves it open whether values exist objectively without human acts of valuation (see Routley and Routley, 1982).

have a moral duty to realize Y, as O'Neill (1992) points out against Taylor (1986): for example, human blood is good for mosquitos, but I have the right to kill a biting mosquito.[19] But if B is an altruist who takes the well-being of A to be her intrinsic value, then Y is an instrumental value for B. Examples include the care of family pets and domestic animals and the protection of wild endangered animals.

Policy Conflicts

It often happens that different parties disagree about their preferred environmental policies. The schema (2) shows that this conflict may come from different factual beliefs about the situation B or about the regularity X & B → A. Such disagreement can eventually be resolved by scientific research. Sometimes the different groups have failed to see the unintended side-effects of the act X (e.g. DDT was effective as an agricultural toxic against insects, but it killed birds as well; fertilizers help fields to grow, but pollute waters). Even these issues can be solved by more careful testing of the consequences of alternative policies.[20]

Some policy conflicts do not concern factual questions, but involve different valuations A. Further, there may be differences in the question whether the situation B will be changed or to some extent modified (e.g. the acceptance or reforms of the market economy). Such value conflicts can easily arise between different interest groups - e.g. land owners and active citizens in the protection of healthy natural environments (forest, urban areas, seas). The demand of sustainability might turn out to be inconsistent with the preservation of old cultural traditions or the way of life of some ethnic group or profession. In democratic societies, different moral opinions are accepted, and such controversies are reconciled by legislation and by the political system (even the parliament).

---

[19] Conversely, if a seagull catches a fish, I have no obligation to rush to protect the pray. Kyle Johansen (2020) defends the stronger thesis that humans have ethical duties to reduce the suffering of wild animals even by gene technology, but I am afraid that this kind of interference in the struggle of survival between wild animals leads to a reductio ad absurdum.

[20] Medicines and vaccinations are not allowed for public use without careful testing.

Sometimes selfish interests - following Roger Scruton's (2011) "oikophilia" or "friends of home" - will lead to rational behavior: awareness of environmental problems arise when the water of *my lake* or the air of *my city* is polluted. But, as Liisa Uusitalo (1989) has convincingly argued, environmental problems typically arise from the *free rider* problem, which is an instance of the prisoner's dilemma in game theory. A free rider calculates selfishly that he need not participate in a joint project of saving a lake, and when everyone follows this reasoning then the worst thing happens. This conflict between individual and collective rationality is a powerful argument against Adam Smith's classical thesis about the benefits of egoistic strategies by "the invisible hand" (cf. Chapter 20 above).

Besides value conflicts in environmental policy, there is the problem that good intentions do not always lead to concrete actions. Most people today would *say* that they are in favor of the protection of nature, but are they willing to *do* something in their own lives for this purpose? Therefore, we need systematic measures and directives in order to promote behavior which is at the same time individually satisfying and altruistic or collectively rational. This is especially relevant in the hard case of climate change.

Climate Change as a Collective Harm

The traditional discussion about the ecological crisis concerned the pollution of air and seas and the exploitation of non-renewable resources. These issues are still serious: WWF calculates that we use every year natural resources the amount of 1,7 planets. But a new theme about the "greenhouse effect" started in the 1980s, and the UN warned about "dangerous climate change" in 1992. The matter was debated in the summits of Rio in 1992, Kyoto in 2005, and Copenhagen in 2009; the Paris Agreement in 2016 defined the aim of keeping the global temperature rise well below 2 degrees Celsius above pre-industrial levels (now about 1.1 degrees). The seriousness of climate change was finally realized when the Intergovernmental Panel on Climate Change (IPCC) collected from hundreds of research articles convincing scientific evidence that the dominant cause of global warming is the human use of greenhouse gas concentrations (CO_2 emissions) as well as black coal in the

Arctic region. Climate models predict catastrophic global ecological, social, economic, and political consequences: with accidental local gains (vineyards in Northern Europe?), the main losses and damages include extreme weather, fires, hurricanes, storms, droughts, desertification, melting glaciers, rise of sea level, economic losses, human suffering (lack of water and food, diseases, climate refugees), animal suffering, losses in biodiversity, and extinction of species. Climate warming is now recognized as the most urgent threat to sustainable development. Current plans and efforts are not sufficient for mitigation, but e.g. Finland has accepted the ambitious plan of being carbon neutral in 2035. The European Union's Green Deal aims to make Europe climate neutral by 2050, and this "green transition" by means of sustainable industry and transport without fossil energy is supported by a massive recovery and resilience plan.

Philosophers of science have been active in the assessment of the evidence for the role of human activities and the reliability of climate models and their predictions.[21] But around 2010 *climate ethics* has been established as a specialty within practical philosophy. It has a similar role as environmental ethics in the model (2): it is important to combine up-to-date scientific research with ethical goals, since environmental politics without (natural and social) science is blind to facts and without ethics blind to values. The main philosophical issue in climate ethics has been the question who is responsible for climate harms (individuals or collectives?) and what principles of justice can govern the transition to the new carbon neutral age.

Global warming is a collectively produced harm – caused by emissions produced by agriculture, industry, services, and consumption. It is a moral harm, as it threatens human rights to life, health, and subsistence. It fosters injustice, as its effects harm especially the poor. As an unintentional long-term side-effect of numerous individual actions, it is difficult to conceptualize it in traditional terms of moral agency. The tendency to evade moral responsibility – it's not my fault! - is explained by psychological mechanisms of moral disengagement.[22]

---

[21] See e.g. Winsberg (2018).
[22] Peeters et al. (2016).

However, as Shue (2014) argues, by appealing to measures of carbon footprint, one should distinguish subsistence emissions (necessary for survival) and luxury (dispensable) emissions. Peeters et al. (2016) suggests that emitters can be assigned remedial responsibility for climate change on the basis of their luxury emissions.

Christopher Kutz (2000) characterizes the shared responsibility of individuals for a collective harm with the term *complicity*. In her dissertation, Säde Hormio (2017) elaborates this conception on three levels: direct risky impact of individuals qua individuals, shared responsibility as members of collective agents (citizens, employees), and shared responsibility as constituents of unorganized collectives (consumers, polluters). It is known that e.g. in Finland about two third of the emissions result from households, which gives a basis for a promising opportunity: a sufficiently great number of citizens and consumers in different countries can together make a difference on global warming by changing their consumerist form of life.

The responsibilities of global warming cannot be reduced to individuals only, but they are shared also by collective agents: corporations operate with demands of social responsibility but also see new business opportunities in sustainable technologies (e.g. electric cars), cities and municipalities develop their environmental strategies (e.g. energy solutions without carbon and coal), and states influence climate change by legislation, taxes, directives, incentives, and investments in R&D. We and our children are all potential victims of global warming, and we need to reconsider our own values and behavior and to elect wise political leaders and decision makers to join us in promoting sustainable development.

Note. This chapter uses material from my articles "Science and Sustainability", in Taina Kaivola & Liisa Rohweder (eds.), *Towards Sustainable Development in Higher Education – Reflections*, Helsinki: Ministry of Education, Finland, 2007, 38-41, and "Luonnon arvo ja ihmisen vastuu", in Arto Haapala & Markku Oksanen (eds.), *Arvot ja luonnon arvot*, Helsinki: Gaudeamus, 2000, 55-67. I have also used my lecture "Reflections on Environmental Ethics", in the Entrètiens of Institut International de Philosophie (IIP) in Beijing, August 15, 2018 (in connection with the XXIV World Congress of Philosophy).

Literature

Attfield, Robin (1991). *The Ethics of Environmental Concern*. Athens: University of Georgia Press.

Berleant, Arnold (1992). *The Aesthetics of Environment*. Philadelphia: Temple University Press.

Cooper, D. E. & Palmer, J. A. (eds.) (1992). *The Environment in Question: Ethics and Global Issues*. London: Routledge.

Hanski, Ilkka (1999). *Metapopulation Ecology*. Oxford: Oxford University Press.

Hooker, Clifford (1992). "Responsibility, Ethics, and Nature". In Cooper & Palmer (1992), 147-164.

Hormio, Säde (2017). *Marginal Participation, Complicity, and Agnotology: What Climate Change can Teach us about Individual and Collective Responsibility*. Helsinki: University of Helsinki.

Johannsen, Kyle (2020). *Wild Animal Ethics: The Moral and Political Problem of Wild Animal Suffering*. New York: Routledge.

Kutz, Christopher (2000). *Complicity: Ethics and Law for a Collective Age*. Cambridge: Cambridge University Press.

Passmore, John (1980). *Man's Responsibility for Nature*. Duckworth: London.

Merchant, Carolyn (1983). *The Death of Nature: Women, Ecology and the Scientific Revolution*. San Francisco: Harper and Row.

Naess, Arne (1989). *Ecology, Community and Lifestyle: Outline of an Ecosophy*. Cambridge: Cambridge University Press.

Niiniluoto, Ilkka (1993). "Kestävä kehitys, tiede ja ympäristöetiikka". In Helena Kekkonen (ed.), *Vastuulliseen teknologiaan*. Helsinki: Tekniikka elämää palvelemaan.

Niiniluoto, Ilkka (1994). "Nature, Man, and Technology - Remarks on Sustainable Development". In Lassi Heininen (ed.), *The Changing Circumpolar North: Opportunities for Academic Development*. Rovaniemi: Arctic Centre Publications, 73-87.

O'Donnell, E. J. & Talbot-Jones, J. (2018). "Creating Legal Rights for Rivers", *Ecology and Society* 23 (1).

O'Neill, John (1992). "The Varieties of Instrumental Value", *The Monist* 75, 119-137.

Peeters, Wouter, De Smets, Andreis, Diependaele, Lisa & Sterckx, Sigrid (2015). *Climate Change and Individual Responsibility: Agency, Moral Disagreement and the Motivation Gap*. Houndmills: Palgrave Macmillan.

Plumwood, Val (1993). *Feminism and the Mastery of Nature*. London: Routledge.

Regan, Tom (1984). *The Case of Animal Rights*. London: Routledge.
Rolston, Holmes III (1988). *Environmental Ethics*. Philadelphia: Temple University Press.
Routley, Richard & Routley, Val (1982). "Human Chauvinism and Environmental Ethics". In Don Monnison, Michael McRobbie & Richard Routley (eds.), *Environmental Philosophy*. Canberra: Australian National University.
Schwarzenbach, Alexis (2002). *Saving the World's Wildlife: WWF – The First 50 Years*. London: Profile Books.
Scruton, Roger (2011). *Green Philosophy: How to Think Seriously about the Planet*. London: Atlantic Books.
Sepänmaa, Yrjö (1986). *The Beauty of Environment: A General Model of Environmental Aesthetics*. Helsinki: Annales Academiae Scientiarum Fennica.
Shue, Henry (2014). *Climate Justice: Vulnerability and Protection*. Oxford: Oxford University Press.
Singer, Peter (1975). *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York: Avon.
Taylor, Paul W. (1986). *Respect for Nature: A Theory in Environmental Ethics*. Princeton: Princeton University Press.
Uusitalo, Liisa (1989). "Economic Man or Social Man – Exploring Free Riding in the Production of Collective Goods". In K. Grunert & F. Ölander (eds.), *Understanding Economic Behavior*. Dordrecht: Kluwer, 267-283.
White, Lynn Jr. (1967). "The Historical Roots of Our Ecological Crisis", *Science* 155, 1203-1207.
Winsberg, Michael (2018). *Philosophy and Climate Science*. Cambridge: Cambridge University Press.
World Commission on Environment and Development (1987). *Our Common Future*. Oxford: Oxford University Press.
von Wright, G. H. (1981). *Humanismi elämänasenteena*. Helsinki: Otava.

Chapter 18: Improving the Quality of Life

Introduction: Three Questions

The World Conference on Science for the Twenty-First Century, organized by UNESCO and ICSU in Budapest in the summer of 1999, adopted a *Declaration on Science and the Use of Scientific Knowledge* and a related *Science Agenda - Framework for Action*. According to these documents, "scientific knowledge has led to remarkable innovations that have been of great benefit to humankind", but scientific advances have also led to "environmental degradation and technological disasters, and contributed to social imbalance or exclusion". As the benefits of science are "unevenly distributed", it is time to agree on a "new commitment": "the sciences should be at the service of humanity as a whole, and should contribute to providing everyone with a deeper understanding of nature and society, a better quality of life and sustainable and healthy environment for present and future generations".

These statements are important, since they clearly recognize that scientific research has been a source of both benefits and harms. It is also significant that here the goals of science are not reduced merely to the increase of economic wealth and prosperity, as in many national and international programs of science and technology policy, but rather refer to more general values like the quality of life and sustainable development (cf. Chapter 17 above). Further, the benefits of science are expected to serve the whole of humanity, including future generations. These observations and requirements define at least three different kinds of problems.

First, *what is meant by the quality of life?* This issue has been discussed in disciplines like social psychology, sociology, economics, development studies, and futures studies, but ultimately it is a question of ethics and political philosophy.

Secondly, *in which ways can science influence our life and thereby shape the future?* This question is sometimes directly formulated as the problem of choosing the specific target areas of research that are most likely to contribute to the improvement of the quality of human life. But to understand this issue, we should start by considering the mechanisms by which science can have an impact on our conditions of living.

Thirdly, *what kind of organization of the research system is most helpful in enhancing a better quality of life?* This is a fundamental problem of science and technology policy: what is the best way of setting up the interchange between scientific inquiry and human well-being?

In the following, I try to say something about all of these three questions. I start with a survey of the approaches to defining the quality of life. Then I outline ten ways in which science can shape the future (where "science" is understood in the broad sense as including natural science, medicine, social science, and the humanities). Finally, I proceed to ask what kind of "social contract" between science and society would be most appropriate in the present world. Even though these questions are relevant to the whole system of research, we shall see that they define urgent challenges for academic research and education within the universities.

How to Define the Quality of Life?

The simplest way of defining the quality of life is to treat this issue as a matter of purely *subjective preferences*: a person's well-being means that his or her wants or wishes are satisfied. A person is unhappy, if his or her wants fail to be satisfied. This individualist and relativist view has been inherited from classical utilitarianism to hedonistic conceptions of consumer society: individual citizens seek happiness and pleasure by buying commodities in the free market. Liberalist utilitarians also thought that the maximal sum of happiness is guaranteed, if all citizens are free to compete for their own interests without constraints from others.

However, the liberalist John Stuart Mill already noted in his *On Liberty* in 1859, however, that all pleasures need not count as equal: it is better to be Socrates dissatisfied than a fool or a

pig satisfied.[1] Others have argued that some individual wishes (e.g. the desire to smoke, the urge to dominate others) may be harmful to the individual or detrimental to the relevant environment, so that the satisfaction of such a desire need not be ethically right at all. Some wishes may be based upon mistaken assumptions and beliefs. Further, individual preferences may be easily influenced by new technological innovations and manipulated by advertising (cf. Chapter 20 below).

Objective accounts of welfare replace subjective preferences by some criterion which is assumed to have a basis in the human nature or in the conditions of human life. Aristotle defined man as the rational animal whose potential of reason and *eudaimonia* (happiness) can be actualized by education. Thus, in the Aristotelian view, revived in an eloquent way by Martha Nussbaum (1986), good human living is constituted by wisdom in knowledge-seeking and by rational choices guided by characteristically moderate human *virtues*.

Naturalist accounts of well-being refer to an assumed system of human *needs*. It is easy to agree that there are some "basic needs" that are necessary for life in general (e.g. sexuality, food, drink, dwelling), and an inadequate supply of them leads to bad or unsatisfactory human life. But lists of human needs, proposed by philosophers and psychologists, usually add to such "material" or "animal needs" a variety of other desirable conditions like security, learning, friendship, and self-respect.

Among the sociologists, Erik Allardt (1993) has proposed an account of three dimensions of welfare. First, *having* includes indicators like economic resources, housing conditions, employment, working conditions, health, and education, but he also suggests that conditions concerning the state of the physical and biological environment should be added to this list. Secondly, *loving* refers to the relations between people - such as family, friendship, and associations. (In the Scandinavian countries, the factor of companionship has a zero correlation with the material level of living.) Thirdly, *being* stands for identity and self-governance - e.g. opportunities to enjoy nature, participate in political decisions, and maintain personal growth. Allardt's three values having – loving –being are

[1] See Mill (1956).

clearly related to the Nordic model of the welfare state (cf. Chapter 19 below).

Economic approaches to welfare have usually been interested in the question of how the citizens possess material *wealth*: the quality of life amounts to being well off. Traditional theories of *justice* are concerned with the distribution of such wealth. While the egalitarians demand an equal distribution, John Rawls (1971) adopts a Difference Principle stating that – given a primary condition of freedom - an unequal distribution of "primary goods" is justified if and only if this is for the benefit of those who are in the worst position in society.

Amartya Sen has criticized Rawls and other welfare theorists for their focusing on the distribution of resources, goods, and commodities. He argues that social equality and the quality of life should be defined in terms of the *capacity* or *capability* that people have in using their resources to satisfy their needs.[2] For example, a physically handicapped person may need more resources as a compensation to achieve the same level of functioning as other people. More generally, we should distinguish between the objective opportunities that the citizens have for their actions and the actual free choices that they make in their lives. This emphasis on what people can really do on the basis of their resources comes close to the idea of "positive freedom" (in one of the meanings of this term).[3]

The capability account of welfare comes also close to some proposed definitions of *health*. Some philosophers have proposed relational accounts where the health of a person is not a state of her physical condition but a relation between the goals and means that she has.[4] But, as aspiration levels have a large variance (during a single person's life span or between different persons), this may lead to unnatural or absurd consequences: a person with expensive tastes is never healthy, and a completely disabled or frustrated person without any goals in life at all is healthy. Therefore, as suggested by Lennart

---

[2] See Nussbaum & Sen (1993).

[3] See Berlin (1969) on positive (free for) and negative (free from) freedom.

[4] The negative definition of health as lack of sickness was replaced by the World Health Organization (WHO) by the positive definition as "the complete state of physical, mental, and social well-being", but it is doubtful that any real person can satisfy such an ideal condition. For a recent program in the "science of well-being", see Alexandrova (2017).

Nordenfelt (1987), a better formulation of the relational account states that the health of a person consists in her capacity to reach those goals that are necessary to her adequate functioning or to her long-term happiness.

These attempts further illustrate the interplay between subjective and objective factors in analyzing what G. H. von Wright (1963) calls "the good of man".[5] Some subjectivist approaches would allow a state where most individuals feel happy but the total condition of the society is a disaster.[6] Some objectivist approaches would allow cases where the society as a whole enjoys a high level of welfare but most people feel miserable.[7] Therefore, the quality of life should in some way combine subjective elements (such as personal feelings of satisfaction) with objective measurable conditions (such as the level of resources and capabilities).

Subjective well-being may be caused by objective social conditions, since life satisfaction depends on the realization of things that one finds valuable. Such conditions may momentarily arise from simple little things in life[8], but more generally they may include personal health and social relations, social trust and security, quality of work, freedom of choice, and democratic political participation. Wilkinson and Pickett (2009) argue that more equal societies almost always do better than less equal ones. It is no wonder that *World Happiness Report* has ranked Finland and other Nordic welfare states as the happiest countries in the world.

Economists, who define happiness as life satisfaction, have also contributed in this field by asking whether money makes us happy.[9] What is known as Easterlin's paradox, the Western countries have become step by step richer in the last decades, but their happiness has not grown in tune. The conclusion is

---

[5] Cf. the interplay of subjective and objective dimensions in Reid's notion of beauty (Chapter 4 above).

[6] Aldous Huxley's novel *Brave New World* (1932) describes a dystopian society where the citizens are kept happy with a soma vapor.

[7] Soviet Union invested in technological progress in a system of state socialism, but the lack of individual freedom made many of its citizens unhappy.

[8] "Happiness is a warm puppy", told Charles M. Schultz's Charlie Brown in 1962.

[9] See Layard (2005) on the economics of happiness.

that, while sufficient wealth removes poverty and improves security, there is a saturation point where the marginal utility of increasing incomes disappears without any more leading to higher levels of life satisfaction. The caricature portrays the life of Scrooge McDuck, the world's richest duck, as miserably overwhelmed with worries.

Money is an instrumental value, not an intrinsic value in itself. In my view, the same holds of happiness as a hedonistic conception. The ultimate aim of life is not to maximize momentary feelings of pleasure and happiness, but to develop one's abilities and capacities and thereby to contribute to a common good. Here deontological ethics surpasses eudaimonistic ethics: "real" happiness as life satisfaction is not an aim but a consequence of the awareness that one has done the best to realize the existing potentials.

Attempts to measure the factors mentioned in the proposed definitions of the quality of life lead to a great variety of *social indicators*. Apart from the problems of operationalization, the obvious difficulty for purely objective accounts of welfare is that it seems very difficult to agree on some list of such indicators and their weights of importance. The choice of such a list already involves an ethical or political standpoint. In this sense, I think we have to accept *moderate relativism* concerning the ultimate aims of human life: different individuals and different cultures may set up these aims in alternative ways, and no dictator (a philosopher, an economist, or a political leader) can fix them for the others.[10]

However, this moderate relativism does not preclude the possibility of agreements on the basic *rights and responsibilities* of all human beings, where the agreement is binding for all those who voluntarily share its commitments. One example of such agreements is given by the ethical codes accepted by many professions. Another is the extensive list of political, social, and cultural rights granted for all Finnish citizens in the constitution of the republic of Finland. As a result of democratic decision-making in the parliament, it is a reflection of the self-understanding of the Finns about the most important ingredients of the quality of human life. Another example is the

[10] Edward Westermarck (1932) argued that moderate ethical relativism sustains tolerance in society.

Universal Declaration of Human Rights adopted in 1948 by the United Nations.

Declarations of human rights are important as expressions of our best wishes, but they also provide a basis for more operational definitions that allow us to observe the situation of different nations with respect to their recommendations. In this sense, highly significant work is done by the United Nations Development Program. Its *Human Development Report* considers the state of almost two hundred countries with respect to the human development index (HDI), defined as the average of three factors (taking numerical values between 0 and 1): health (measured by life expectancy at birth), education (measured by adult literacy rate and the enrolment ratio in first-, second- and third-level education), and wealth (measured by GNP per capita). The highest HDI-index has been obtained by Australia (in 1990-92 and 1995-96), USA (in 1993), Norway (in 1994, 1997-2009, 2011, 2014-19), and Switzerland (2010, 2012-13). Even though some historical progress can be observed by the series of the HDI-index, the marginalization of poor countries can be seen clearly. Later the index was modified to take into account the intra-national diversities in the quality of human lives.

The work on HDI was continued in 2000 with the UN agreement on eight *Millennium goals*:

- eradicate extreme poverty and hunger
- achieve universal primary education
- promote gender equality and empower women
- reduce child mortality
- improve maternal health
- combat HIV/AIDS, malaria, and other diseases
- ensure environmental sustainability
- develop a global partnership for development.

Until 2015, progress in realizing these goals was slow – with the positive exception in the reduction of poverty and child mortality. One of the reasons was the lack of a transparent systems of performance assessment. In 2015 a new system of 17

sustainable development goals with 169 specific targets was adopted by 193 member countries of the UN. This *Agenda 2030* includes as its goals no poverty, zero hunger, good health and wellbeing, quality education, gender equality, clean water and sanitation, affordable and clean energy, decent work and economic growth, industry and infrastructure, reduce inequality, sustainable cities and communities, responsible consumption and production, climate action, life below water, life on land, peace and justice, and partnership for goals. These lists, with their specifications, are important guidelines to nations who wish to promote quality of life for their citizens.

How can Science Influence Human Life?

Science can influence human life and shape our future in several direct and indirect ways. I distinguish here ten of such patterns.

First, ever since the birth of philosophy and science in the ancient Greece, as separated from myths and religions, science has provided *a body of knowledge*. According to Plato's classical definition, knowledge is constituted by justified true beliefs. For Aristotle, rationality - our ability to use reason in theoretical and practical matters - is a defining characteristic of human beings which distinguishes them from other animals. Science includes not only "know-that", i.e. well-established information about the facts and regularities of the world, but also "know-why", i.e. explanation and understanding of the permanent and changing features of the reality.

Secondly, while the medieval tradition usually conceived *scientia* as the possession of well-established knowledge, and the method of science as a tool in organizing and teaching already known truths, modern philosophers from Francis Bacon and René Descartes emphasized the dynamic nature of science as the pursuit of *new truths*. Fallibilism, as characterized by Charles Peirce and Karl Popper, admits that all of our efforts in truth-seeking are liable to error, so that even our most convincing results in science are always in principle open to criticism, correction, revision, or rejection (cf. Chapter 7 above). Hence, the method of science should be seen as *a tool of systematic and critical thinking*. It is this aspect of science which can be

viewed as its most enlightening ingredient: not the possession of truths, but the diligent search for truth.

Thirdly, even though basic value commitments are ultimately personal choices that cannot be proved as valid by empirical science, philosophical and critical thinking helps us in many ways to assess and *reconsider our system of values* (cf. the discussion of the quality of life in the preceding section). Social sciences can give us information about the diversity of ethical principles in various cultures. Scientific inquiry can change our instrumental values by establishing or disproving some putative connection between means and ends. Science can also show that some goals are not worth pursuing, as they are inaccessible by the available means. Further, knowledge about the consequences of adopted values within the history - e.g. awareness of the ecological crisis due to our wish to dominate and exploit nature (cf. Chapter 17 above) - gives us motives for changing even our intrinsic values.

Fourthly, through scientific education academic masters and doctors acquire *skills* needed in the professional life. These professional skills may be related to specific occupations (including academic teaching and research). But the ability of critical thinking should also provide a general aptitude for learning. The unity of inquiry and learning was expressed already in Humboldt's slogan "Bildung durch Wissenschaft", and it is a central idea of many contemporary reforms of university education (cf. Chapter 13 above). This can be seen in the similarities between effective learning strategies and the methods of scientific research.

Fifthly, science may allow us to make *predictions* and *forecasts* about the future behavior of natural, human, and social systems. Such a predictive power requires knowledge about the present state of the system and about the laws governing it. Forecasts may also be based upon trend exploration by using advanced statistical methods. They help us to anticipate future courses of events and to prepare us for their possible benefits or disasters.

Sixthly, for chaotic natural systems, whose behavior cannot be predicted from any finite knowledge about their present state, and for complex social systems, whose behavior depends on the choices that human beings will do in the future, the goal of prediction can be replaced by a more general approach: *the*

alternative futures are represented as a branching tree, and the probability and the desirability of these branches is systematically investigated. If we know, where we want to be in the future, the tree helps us to formulate a strategy which delineates a path from our present situation to the desired goal. This is the basic idea of the scenario method employed in futures studies (cf. Chapter 12 above).

Seventhly, scientific knowledge about natural and social regularities enables us to formulate conditional recommendations of action of the form "If you wish A, and you believe you are in situation B, you should do X". These rules, called *technical norms* by G. H. von Wright, are typical results of applied research (cf. Chapter 11 above). In fields of applied or mission-oriented research, some goal is assumed or taken for granted (e.g. environmental protection, agricultural production, occupation health, social welfare, gender equality, peace), and a targeted research program or a research institute is established to work towards the achievement of this goal. But, as the history of science shows, practical progress in finding technical norms often comes in an unexpected and unintended way from the results of basic research.

Eightly, scientific knowledge can be used as a resource in *rational decision-making and planning*. Formal models of decision-making include factual beliefs (probabilities) and valuations (utilities), and the task of science is to provide research-based facts for decisions. Such decisions may concern our vitally important choices of individual action, but they may also involve significant collective decisions about the improvement of our natural and social environment. Thus, the relevant results of science should be available, by science advice, to the most central decision-makers in the society.

Ninthly, scientific knowledge may be applied to develop new *technological tools*, procedures, and products. According to the traditional "science-push model", an "innovation chain" starts from the discoveries in basic research, proceeds via applied research and industrial "development" to new products. The final stage of "innovation" transforms such technological possibilities into commercially profitable commodities in the market. After the success of Los Alamos, research and development (R&D) are today often performed as united team works within the same institutions, so that the innovation

chain is as it were compressed into the activities of a single research group, combining activities in producing new knowledge and finding out the potentials of its commercial utilization.

Tenthly, through its connection to rational decision-making, administration, technological change, and commercial innovation processes, science serves to increase the *income* of nations and corporations and to strengthen the competitiveness of national economies.

To summarize, science (1) changes our world view by the production of new knowledge, (2) teaches us new methods of critical thinking, (3) allows us to rationally revise our values, (4) promotes learning strategies and professional skills, (5) helps to predict future events, (6) assists us in the systematic investigation of alternative possible futures, (7) justifies conditional recommendations of action, (8) provides a rational factual basis for decision-making and social planning, (9) gives rise to new technological innovations, (10) activates economic competition and increases the wealth of nations.

Science as an Intrinsic Value

Given the multiple ways in which science shapes the future, it is important to ask whether we can in some way and to some extent guarantee that this influence also advances human life.

The traditional optimistic view about science holds that the progress of scientific knowledge, method, and thinking (cf. the first four items in the list of the preceding section) is intrinsically a good thing - and thus has an immediate influence on the quality of our life. Against the radical sceptics, who sought peaceful tranquility in the denial of human knowledge and recommended the suspension of judgement in order to avoid dogmatism, and against religious thinkers, who urged that the scientific study of nature by men is dangerous or illegitimate, most ancient philosophers argued that knowledge of important truths is a precondition of individual human happiness. Also modern *Enlightenment* has been based upon the conviction that it is better to know than to be ignorant, and therefore it is our natural prerogative to pursue new truths (cf. the Chapter 15 above). All human beings have not only the natural desire to know, but in democratic societies they also have

the basic right to science-based education which replaces their prejudices and other false beliefs by the scientific world view.

If knowledge and its pursuit are *intrinsic values*, then universities as institutions of academic research and education face three important challenges. One concerns *the quality of research work*: in order to reach well-justified and truthlike results, academic research should keep up high professional standards of inquiry in all circumstances. This has been a central function of the universities since the unity of research and education was adopted in the Humboldtian university in the early nineteenth century. The Mertonian virtues of originality, universalism, disinterestedness, honesty, fairness, and openness to public criticism have served to strengthen then efficacy of the scientific community as an autonomous social institution.[11] The employment of peer review methods for evaluating publications has more recently been complemented by similar assessments of research groups and institutions. For example, the University of Helsinki organized in 1999 its first research assessment exercise where international panels evaluated the scientific quality of all of its disciplines.

The second challenge is to keep us high *standards of academic education*, both in its up-to-date research-based contents and in the use of most advanced learning strategies. Today this activity is also extended outside the circle of own students via open university, further education, science centers, science fairs, and popular scientific journals. Even though the so-called Information Society, both in Europe and North America, seems to be leaning towards entertainment and rather trivial surface forms of knowledge, the applications of information and communication technologies create many-faceted opportunities to use the new media for education and "edutainment".

Moreover, international co-operation tries to balance the world-wide inequalities in the access to education. Various joint efforts have been in operation with the developing countries, such as North-South co-operation in higher education and education export from Europe to the Middle East and Asia, as well as the program of improving literacy among girls.

---

[11] The sociologist of science Robert Merton formulated "the ethos of science" in 1942. See Merton (1973).

The third challenge concern the *ethics of science*. Here the four basic principles of research integrity and good scientific conduct are *honesty* (against fabrication of data and misrepresentation of results), *fairness* (against plagiarism and misappropriation, i.e. stealing the merit of new discoveries), *protection* (against harms to objects of research like nature, animals, human beings, and social groups), and *responsibility* (against harmful applications of science to the environment, culture, and society at large).[12] These principles are vital to the success of science to provide true or truthlike results, so that science serves both belief rationality and instrumental rationality (cf. Chapter 8 above). But they are also significant for the public trust of science in society. In Finland, all universities and research organizations have committed themselves to follow the guide *Good Scientific Practice and Procedures for Handling Alleged Misconduct and Fraud in Science*, issued by the National Advisory Board on Research Ethics.[13] In the European Union, a corresponding document of the self-regulation of the scientific community is *The European Code of Conduct for Research Integrity*.

The view of knowledge and education as intrinsic values is based upon the philosophical thesis that science is in fact a progressive enterprise in its attempt to approach the truth. I have defended this thesis in my book *Critical Scientific Realism* (1999). But it has also been under a heavy attack from the "postmodern" sceptics, relativists, and anti-realists. Somewhat surprisingly, in the STS studies of science, technology, and society it has been popular to claim that scientists are "lobbyists for their own" who are not able, nor willing, to search for objective truth.[14]

I think that scientific thinking should tolerate and even encourage a self-critical reflective attitude towards science itself. But open hostility to science, even by professional studies of

---

[12] See Niiniluoto (2007).

[13] As the Chancellor of the University of Helsinki in 2008-13, my duties included the handling of the violations of research ethics. Cases of fraud were extremely rare, but typical complaints concerned disputes about the rights to scholarly authorship.

[14] See Jasanoff et al. (1995).

science, is not fair – and the debate known in the 1990s as *Science Wars* led to unfortunate dogmatic disputes.[15]

The Social Contract between Science and Society

Ever since Francis Bacon, scientific inquiry has been compared to the explorers and conquerors of new territories (cf. Chapter 15 above). This simile was made popular by Dr. Vannevar Bush in his memorandum *Science: The Endless Frontier* to President Roosevelt in 1945.[16] Bush argued that the Federal Government should promote curiosity-based research by free intellectuals in universities. This helps to open "new frontiers of science" and leads to success in industrial research as well as promotes health, wellbeing, and security. On Bush's advice, the National Science Foundation (NSF) was established as a research-funding organization in 1950. The same model – with an attempted balance between fundamental and applied research and technological development (cf. Chapter 11 above) - was followed since the 1960s by many countries in their systems of national science policy.

The European Union has organized its research and technological development since 1984 in the form of Framework Programs (FP), whose expressed aim is to improve the economic competitiveness of the community. As exceptions to this aim, since 2007, are the grants for "frontier research" awarded by the European Research Council (ERC) on the basis of the excellence of individual applications independently of their practical utility.

An important change of rhetoric was proposed by the first report of the new European Research Area Board (ERAB) in 2009. In its view of "a new renaissance", ERAB suggested that the European Research Area should be driven by societal needs, but instead of economic competition it drew attention to "wicked problems" and "Grand Challenges": climate change, energy supply, water resources, ageing, healthcare, sustainable prosperity to all.[17] In the Horizon 2020 program this list of societal challenges is elaborated in seven points:

---

[15] For a summary with references, see Niiniluoto (1999), Ch. 10.

[16] Vannevar Bush was an electric engineer, not related to the later presidential family with the same name.

[17] See ERAB (2009). The Finnish member of this Board was Marja Makarow.

- health, demographic change and wellbeing
- food security, sustainable agriculture and forestry, marine and maritime and inland water research and the bioeconomy
- secure, clean and efficient energy
- smart, green and integrated transport
- climate action, environment, resource efficiency and raw materials
- Europe in the changing world – inclusive, innovative and reflective societies
- secure societies – protecting freedom and security of Europe and its citizens.

These headlines are not yet research problems, but they can be approached by developing multidisciplinary research programs.

Michael Gibbons, a former director of the Science Policy Research Unit at the University of Sussex and the secretary-general of the Association of Commonwealth Universities, published in 1999 in *Nature* an article on "science's new social contract with society". Gibbons argued that "the old image of science working autonomously will no longer suffice": the clear demarcation lines between university science and industrial science are disappearing. According to the new "social contract", not only can science speak to society as a contributor of "reliable knowledge", but "the society can now 'speak back' to science" by demanding "various innovations" and "socially robust knowledge". Such knowledge should be "transparent and productive" in the sense that its acceptability is tested "not only against nature, but against (and hopefully also with) other people".

In my view, Gibbons refers to important developments within the relations between science and society, but his main conclusion is somewhat misleading. It is a fact that university science is today only a small fraction of all research activities. For example, in Finland, where the total R&D expenditure is about 3% of GNP, almost 70% of this R&D is done in private

corporations.[18] Of the publicly funded 30% of R&D, about one half consists of support for technological development in industrial firms and sectorial research carried out in governmental institutions under the auspices of different ministries. A new instrument of "strategic research" was established in 2013 for the study of societally important problems. About 40% of the public resources go to the universities - either directly or via the Academy of Finland. Today academic research work in the universities increasingly relies on external funding - its sources include contract research and the research programs of the Academy of Finland.

These figures already suggest that today science is primarily considered as *an instrumental value*. But even though the residue of "free" academic basic research, founded upon the duties of university professors and research groups financed on a competitive basis, may seem to be small in comparison with other areas of R&D, it still has (as noted above) very important tasks in promoting science-based world views and education. In spite of the emphasis on research with useful applications, Finland has maintained an important role for basic research – and (*pace* Gibbons) decisions about it should be based on internal epistemic criteria of excellence. This is in agreement with Vannevar Bush's policy recommendations. Academic research should still serve as a source of new theories and innovative methods. The applications of the results of basic research, as well as research in the service of the education of professional skills, have a great instrumental value. The power of fundamental research was seen in the ability of the medical community to develop in a short time vaccination for the covid-19 pandemic in 2020. The autonomy and integrity of academic research are important for the universities also in order to maintain their critical potential towards dominant cultural and social trends. For example, the universities should be able to study the mechanisms that sustain the present inequalities among the nations or among their citizens, and have the liberty to raise their voices against the global threats in the present world. In its *Strategic Plan 2021-2030*, the University of Helsinki combines excellence in fundamental research with ambitious

[18] In the heydays of Nokia's success, its investment in R&D was larger than the total budget of the Finnish universities.

statements about its role in promoting sustainable development and global responsibility "with the power of knowledge - for the world".

Sectorial and mission-oriented research can be directed towards goals that are useful in planning the future, improving administration, protecting environment, and making rational decisions - and thereby they are instrumentally and socially relevant for the satisfaction of human needs (cf. the items five to eight in my list). But, in order to avoid socially important but low-quality research, here it is again important to emphasize the necessity to secure the scientific standards of such work.

Many national programs of science and technology policy look at the benefits of research only in terms of the ninth and tenth items in my list, viz. technological progress and economic wealth. Technology is seen as a way producing and distributing commodities in the free market. Firms, based on private ownership, produce goods with the aim of maximizing profit, and the consumers have complete freedom in making their own choices according to their own needs and preferences. Science, both basic research and applied research, serves as a basis of technological development. This motivates the treatment of scientific research and higher education as "investments" which should yield economic profit in the short or at least not-too-long run. This means that both science policy and technology policy are understood to be parts of the "national innovation system" which ultimately aims to promote commerce and industry.

I agree that the universities, too, have a role in the creation of technological innovations - and this is one way in which they can be socially relevant for human life.[19] But, in my view, the rationality of scientific inquiry should not be reduced to the commercial principles of technology policy (cf. Chapter 20 below). It is still significant to make a distinction between technology and science: technology does not produce knowledge by inquiry like science, but rather designs new artefacts, tools, and machines. Artefacts are not constrained by *truth* in the same way as knowledge claims, but by what is physically and

[19] Societal interaction, including participation in innovation activities, was added to the University Act of Finland in 2005.

economically *possible*. We do not decide what is real in nature, but we can choose what artefacts we wish to produce.

For these reasons, I think it is misleading that Gibbons does not distinguish between the acceptability of a research project (e.g. the Superconducting Super Collider), a knowledge claim (e.g. whether genetically manipulated organisms or GMOs affect our health), and a technological artifact (e.g. whether GMOs should be allowed to be sold in the market).

Technology is "value-laden" in a way which differs from science. Commercial value in the market is one of such values, in our Western society even the primary criterion for the innovation and diffusion stages in the life-cycle of artefacts. But it need not be the only relevant factor in technology assessment (see Chapter 20 below). Technology assessment should be organized in co-operation with laymen who are potential consumers of technological tools and methods.

In a democratic state, society indeed "speaks to science" in the sense that the citizens should have the right to participate in decision about how large percent of the GNP is spent on research and education, and whether public funds are allocated to military research. More generally, rational science and technology policy in a democratic state should be based upon co-operation between the public, private, and "third" sectors. Not only the government and corporations are included among the solvent clients of science. There should also be channels by means of which interest groups and non-governmental non-profit organizations can gain advance from the results of science. One interesting attempt in this direction is given by Science Shops in the Netherlands.

Regularly conducted "science barometers" in Finland have consistently shown that the social contract between science and society is functioning well.[20] Most Finns are interested in science. Concerning the very or fairly high trust in social institutions, universities are in the third place (77%) after police (83%) and defense forces (79%), followed by the judicial system (69%) and the scientific community (69%). The Finns expect science to be successful in problem-solving in treating diseases (85%), energy production (67%), prolonging the human life span /65%), and improving quality of life (60%).

[20] For the results in 2019, see Varpula (2019).

Conclusion

I conclude that, just as the Budapest declaration on science demanded, the sciences can be at the service of humanity as a whole. But we should also be aware that science and the related academic institutions alone cannot do very much: a global political will is needed as well.

In its search for "globalization with a human face", the UN Development Report builds its faith in the democratic extension of global governance. The Report estimates that extreme poverty in the world could be eliminated by a sum which is less than the expenses per year spent on cosmetics in the United States or on ice cream in Europe. The costs of basic education for all people in the world and basic health care and food for everybody correspond to the international trade with weapons, and they could easily be covered by a small tax on speculative currency exchanges, as suggested by James Tobin.

If solutions to global political problems could be found, institutional networks of higher education and research centers could then be employed to secure permanent capacities of nations to sustain and improve the quality of life of their citizens. It is our common responsibility to join efforts in working for this goal.

Note. This updated chapter is based a lecture in the seminar "Science Shaping the Future" at the University of Helsinki, March 28, 2000.

Literature

Alexandrova, Anna (2017). *A Philosophy for the Science of Well-Being*. Oxford: Oxford University Press.

Allardt, Erik (1993). "Having, Loving, Being: An Alternative to the Swedish Model of Welfare Research". In Nussbaum & Sen (1993).

Berlin, Isaiah (1969). *Four Essays on Liberty*. Oxford: Oxford University Press.

Bush, Vannevar (1945). *Science: The Endless Frontier*. Washington: United States Government Printing Office.

ERAB (2009). *Preparing Europe for a New Renaissance: A Strategic View of the European Research Area*. European Commission: Brussels.

Gibbons, Michael (1999). "Science's New Social Contract with Society", *Nature* 402 (2 December 1999).

Jasanoff, Sheila, Markle, Gerald, Peterson, James & Pinch, Trevor (1995). *Handbook of Science and Technology Studies*. London: SAGE Publications.

Layard, Richard (2005). *Happiness: Lessons from a New Science*. New York: The Penguin Press.

Merton, Robert (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: The University of Chicago Press.

Mill, John Stuart (1956). *On Liberty*, ed. by C. V. Shield. Indianapolis: Bobbs-Merrill.

Niiniluoto, Ilkka (1999). *Critical Scientific Realism*. Oxford: Oxford University Press.

Niiniluoto, Ilkka (2007). "Ethical Issues in Research and Education". In Kirsi Tirri (ed.), *Values and Foundations in Gifted Education*. Bern: Peter Lang, 9-14.

Nordenfelt, Lennart (1987). *On the Nature of Health: An Action-Theoretic Approach to the Philosophy of Medicine*. Dordrecht: D. Reidel.

Nussbaum, Martha (1986)). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.

Nussbaum, Martha & Sen, Amartya (eds.) (1993). *The Quality of Life*. Oxford: Clarendon.

Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Varpula, Vesa (ed.) (2019). *Summary of the Finnish Science Barometer 2019*. Lahti: Tieteen tiedotus.

Westermarck, Edward (1932). *Ethical Relativity*. London: Kegan Paul.

Wilkinson, Richard & Pickett, Kate (2009). *The Spirit Level: Why More Equal Societies Almost Always Do Better*. London: Allen Lane.

von Wright, G. H. (1963). *The Varieties of Goodness*. London: Routledge.

Chapter 19: Justice, Law, and Welfare Society

The first section of this chapter documents what I wrote in October 1994, when Finland was voting for the membership in the European Union. The next section summarizes what happened thereafter. The final sections return to more principal questions about justice, law, and welfare society.

Vision in 1994: Europe at a Cross Road

A story, well-known and popular among the modest and self-ironic Finns, tells that about two thousand years ago the tribe of our ancestors marched from the east, reached the Baltic Sea, and found a crossroad with a sign pointing to the south: "To fertile lands". Those who could not read turned to the north and came to Finland.

After a millennium, Finland became the north-east corner of the Western Europe, a border area between the Roman and Byzantine churches. As part of the Swedish kingdom, Finland became a member of the culturally unified Europe through its government, legal system, and university education. As an autonomous Grand Duchy of the Russian empire in the 19th century, Finland reached its self-consciousness or identity as a nation with its own language, folklore, literature, and art. The University of Helsinki played a central role in this process. As an independent nation, Finland was built to be a democratic state with a rule of law; after the Second World War, Finland skillfully operated with a principle of active neutrality in foreign affairs, and successfully developed an economically prosperous welfare state which maintained a balance with agricultural and forestry economy, industrialization, and urbanization.

Our success story continued to the year 1990. With overwhelming enthusiasm and optimism, Presidents Bush and Gorbachev met in Helsinki to declare the opening of a new era without cold war. On the same September day, I gave the main address at the celebration of the 350th anniversary of the University of Helsinki, demanding the continuation of the modern cultural project of Enlightenment, but urging that a warning signal for the "consumption feast" has been given. This was just the time when the overheated Finnish economy - strained also by the losses in the trade with Soviet Union, economic fluctuations in Europe, and eventually failing attempts to maintain a strong currency - started to fall. Finland entered a deep recession, a period of economic, social, and political crisis. In spite of the recent recovery of export industries, with the help of the government, the problems have not been solved: the rate of unemployment is still almost 20%.

Finland is by no means the only locus of crisis in Europe today: unemployment in Spain has reached 25%, neofascists join the government in Italy, racism and violence increase in Germany, the former socialist countries have difficulties in adapting to the market economy, economic insecurity fosters militant nationalism and organized crime in Russia.

The development now displays many opposite trends. A great eastern imperium has been dissolved, but a large and powerful western union is in the making. Finland is among the countries seeking membership in the European Union, but - here as elsewhere - the issue has divided opinions among the citizens. Just yesterday the Finns have made a historical choice by voting yes for the EU with a majority of 56.9 per cent.

The economic and social turmoil has intensified debates about the fate of the welfare society, especially the state-centered model developed in the Nordic countries. The economic integration, characterized by internationally moving capital without a homeland, seems to make nation-states obsolete. Those who used to be called political "conservatives" are now demanding thoroughgoing revolutions in the social structures, while the role of the "radical" left-wing parties is to restrain change in order to protect the health of the social system (in the same way as the green movement is opposing uncontrolled technological change in order to protect nature).

Still some years ago it was a commonplace to believe in a law of progress. The Western societies were supposed to have a destined direction towards ever increasing wealth and happiness. The socialist countries were officially on a road to the paradise of communism. Finland took Sweden to be a model to be followed with a lag of a few years. In future studies, conceptions of postindustrial society (Bell), information society (Masuda), third wave (Toffler), and megatrends (Naisbitt) formulated optimistic visions of a qualitatively better society of a new kind. Now the dramatic historical events have cooled down or disrupted these high expectations. For example, the economic recession led soon to governmental programs which favor "neo-industrialization" and cut down the public sector (i.e. precisely those "information occupations" that were supposed to constitute the major class of the postindustrial society).

It seems that the whole Europe has come to a crossroad. We have been at loss, without knowing who we are, where we are, where we want to go, and how to go where we wish. There are no signs which tell us where the fertile lands are. We have to find our own way to the future.

The building of the European Union can be seen as an expression of the optimistic program of economic progress. Its critics argue that the EU-enthusiasts have failed to understand the great changes in the world history: the Modern Age has come to its end. But even many of those sceptics, who think that we are now facing what Spengler called the decline of the Western culture, accept as a fact that our best strategy is to be active in the planning and building of a new Europe - rather than trying to separate us from the flowing trends of history. So far, at least in Finland, the discussion about the EU has concentrated on the payments and on the farm subsidies that we might get as a member state. Now it is time to start to find some positive content, some positive goals, that we wish to achieve together in the future.

We have to recognize that the European people and societies differ largely from each other. The features that might unite us should be sought from our common traditions - especially the ideal of education or *Bildung*. From this perspective, the Western or European ideal of a human being gives

emphasis to such characters as rationality, freedom, activity, optimism, learning, individualism, and social organization.

Hence, our immediate task is to analyze the challenge, opportunities, and threats of integration on various levels - not only economic, but also political, legal, social, cultural, and linguistic. The parts of Europe have developed from different backgrounds and circumstances to a variety of cultural and social systems. Besides variations in the conditions of local environment, these systems are based upon characteristic assumptions and valuations. Any attempt at integration must pay attention to such differences in beliefs and values. For example, it is crucially important to recognize that differences in social security policies may reflect conflicting views about the present trends and future development of society. Similarly, differences in government and legislation may be based upon opposing conceptions of morality and justice.

And What Happened Next …

Heated debates on the situation in Finland were conducted in a working group that I chaired in the autumn of 1993 by the appointment of the Prime Minister Esko Aho. Our group never reached a consensus about the present state of the Finnish mentality and its future prospects.[1] The way out of the crisis with great expectations was opened, when Finland joined the European Union in 1995, and the Nokia company had spectacular success as the market leader with mobile phones. The EU turned out to be a useful framework for supra-national activities in security, environment, trade, student mobility, and refugees, but hopes for good relations with Russia were gradually frustrated. China raised to a new status as a super-power. After a bright Millennium period with the program of globalization, the finance crisis emerging from the USA hit the world in 2008, Nokia collapsed as a producer of mobile phones, and in 2020 the corona pandemic upset the life of most nations. And just when the covid-19 virus started to weaken its threat, Russia shook the world by attacking Ukraine on February 24, 2022. Within less than three months, Finland and Sweden submitted membership applications to NATO on May 18, 2022.

---

[1] See Niiniluoto and Löppönen (1994). Our next books discussed Europe and the global world.

Morality, Justice, and Law

To begin our philosophical discussion, it is in order to make some comments on the relations between morality, justice, and law.

Social groups, from small communities to firms, professions, societies, states, and cultures, have common customs or habits (*mores* in Latin) which regulate the behavior of their members. Desirable and undesirable patterns of behavior are associated with positive and negative emotions and attitudes; thereby the cohesion and strength of the group in enhanced. The *morality* prevailing in the group is constituted by such influential but mostly unwritten principles. They are codified or made explicit in systems of values and norms formulated by religious or philosophical doctrines; *ethics* is the philosophical theory that tries to systematically analyze and justify such moral principles. An important part of ethics, emphasized by Aristotle, is provided by the virtues that a good person, or a person having a good life, should have. In particular, a sense of justice can be understood as a fundamental mental disposition of a virtuous person.

Justice is also an attribute that can be attached to social situations, structures, and institutions. Based upon implicit or explicit moral principles, the members of a social group have an ability to judge whether a situation is just or not. This ability can be called a sense of justice. This *Rechtsgefühl* is no historical constant: it changes over time, when moral values develop in new circumstances. One motive for change may result from the application of ethical theories to questions concerning justice. In practice, philosophical theories of ethics and justice are in a dynamic mutual interaction with the prevailing morality and sense of justice.

Immoral or unjust action causes moral disapproval in the community, but this is in many cases only a weak threat or sanction. Legislation can be seen as a more formal tool of defining the rights and duties of citizens and social institutions, whereby their behavior is regulated by judicial sanctions decided by the authority of courts.

In some discursive theories of law (inspired by Perelman and Habermas), the validity of a legal norm in a community is defined as the acceptability in principle of that norm (or its

interpretation) by communication and argumentation in ideal conditions. For example, Aulis Aarnio (1987) recommends that "legal dogmatics ought to attempt to reach such legal interpretations that could secure the support of the majority in a rationally reasoning legal community". I think in such theories "validity" comes to mean something like "being just according to the best theory of justice".[2] This view is radically opposed to the doctrines of legal realism and positivism, which define the valid legal order as the content of those enacted statutes that are actually followed or respected in the community.

The Nordic tradition has been strongly influenced by legal realism. As law-abiding Protestants, we take legal statutes seriously. Therefore, the Finns tend literally and stubbornly follow all the EU directives, while the more flexible attitude in the southern Europe is more willing to ignore some of them.

It is natural that the legal order should reflect the sense of justice in the society. But, in my view, realism is the right view in the sense that the actually valid content of the legal order (e.g. in Finland) is always to some extent different from the ideal system of justice. Therefore, existing laws may always be evaluated and criticized by the moral sense of justice of the citizens and by the theories of justice of philosophers. This is an important issue, when the laws and directives of the European Union are discussed.

Welfare Society: The Nordic Model

The building of the Nordic model of *welfare society* was started in the 1960s. In Finland, it was mainly a joint project of the Social Democrats and the (formerly agrarian) Center party. It has involved a conception of a strong state and large public sector which takes care of security networks, allowances, and many other welfare services to all citizens. In contrast, the Catholic model of welfare has given more emphasis to the family and the church, while the American model relies largely on private insurance policies and the market.

In my personal view, the Nordic model is the best social organization found so far (at least for economically well-developed countries). It has avoided the depressing lack of freedom

[2] Cf. the debate in Niiniluoto (1985) and Aarnio (2011).

in the earlier Eastern socialism and the immense inequalities of wealth in the American capitalism. The Finns, in spite of fluctuating economic recessions, still value highly the benefits of this model. In a time, when the feeling of insecurity and loss of control of one's own life are leading to frustration and alienation, i.e. to a kind of mental recession, the desire for solidarity and responsibility is again increasing.

This is not to say that the Nordic model has ever been perfect: some of its problems were discussed by the social scientists already in the 1970s. Moreover, it has to be transformed to account for the recently emerged problems. If we cannot afford to have all the good things, we urgently need value discussion to put our goals in a preference ranking. This has proved to be extremely difficult, since every interest group in the society wants to preserve its already achieved status.

A social organization may be evaluated on two grounds. First, is it *effective* in producing the desired results (e.g. public services), and does it perform this task *efficiently* in the cost-benefit sense? Secondly, is it *just* or at least more just than its alternatives? I think the Nordic model scores on both of these points - and the recent criticism against it is in many ways unfair.

From Ronald Reagan's USA and Margaret Thatcher's England, the neoliberal New Right movement reached Scandinavia in the 1990s. It gained impetus from the collapse of the socialist system and the difficulties created by the economic recession. It started an ideological campaign against the welfare state: the public sector is too large and ineffective, the taxes are too high, the role of the state and the local government is bureaucratic and paternalistic, and prevents the active operation of the free market and enterprising citizens. These claims are today repeated, in radical and moderate forms, by some thinkers and new political movements in Finland.

I think it is fair to say that in some areas our system has been too centralized and too much governed by restrictive directives and regulations. There was a time in the 1970s when it was thought that the state should take care of almost everything - rather than give opportunities to "undemocratic" private companies, societies, and individuals. For example, there was a state committee that wished (but fortunately failed) to abolish the activities of private non-profit foundations. The

reforms which increase the autonomy of universities in their financial decision making and allow the elementary and secondary schools to be more initiative in planning the teaching and the curriculum are in the right direction. Moreover, the principles of equality should not demand that all citizens are identical, or that no one is allowed to be better than others in any respect. It is appropriate to acknowledge the differences in the talents and skills of young people, and take this positively into account in the educational system.

However, I think it is a mistake to claim that the welfare system was the cause of the recession in the early 1990s. Finland was able to build its welfare society until the end of the 1980s without a large foreign loan. What went wrong in the latter half of the 1980s was the overheated economy, due to selfish and greedy values in the market opened for free competition. (The opposite view claims that the lack of competitive spirit drove Finland to recession.) When the time of paying the bills came, the illusions were broken, companies were bankrupt, unemployment exploded, and the banks entered a crisis, which the government tried to solve by huge subsidies and by increasing national debt. The growth of expenses in the social sector is an effect, not the cause, of these events.

The overall claim that publicly produced services are always ineffective and inefficient is also doubtful, if presented in a dogmatic manner. Similarly, the claim that they are always effective is likewise dogmatic. It is an empirical question, to be studied separately for each case, which services can most profitably be arranged by the state or local community sector and which should be privatized. For instance, the US and British health care systems suggest that the costs of private social welfare programs may be higher than public ones. The mechanical policy of subjecting the whole public sector to accountability in terms of cost-benefit-efficiency has proved to be problematic in many cases (e.g. university, army, police, school): the desired "results" are difficult to define and measure; the real costs (including increased working load and bureaucracy) are hard to count. The transformation of public offices and institutions into business corporations (e.g. post, train) has resulted in diminishing available services and to lower quality of life. Further, it is not always the case that better results are obtained by pushing the decision-making power as close to the potential

customers as possible. For example, in Finland the state subsidies for municipalities and counties were formerly divided and "earmarked" in different areas; when the local councils were allowed to make free decisions on these funds, the programs for cultural activities were severely cut down. The model of changing the support of high-quality art (opera, orchestras, films, etc.) to vouchers distributed directly to the citizens would probably lead to disastrous results. (Just see the polls of the most popular TV-programs!)

There is thus much to be said in favor of the Nordic welfare model in terms of its efficiency. This becomes clear also if we note how it may be able to sustain the equality between the genders, rights of cultural minorities, social security and safety of the citizens, balance between social groups, lack of social diseases like poverty, violence, corruption, and criminality. This fact can be seen in the success of the Nordic countries in the measurements of happiness and quality of life (see Chapter 18 above).

The Nordic model of welfare society has given room for a large "third sector" - besides the state-centered public sector and the private business sector. When the formerly socialist countries of Eastern Europe got rid of the strict rule of the centralized state and one party, their urgent task was to develop legislation for social security and networks of the civil society, where the citizens engage themselves in voluntary, non-governmental, non-profit activities in order to pursue personally and socially important goals. Finnish expert group have worked in Estonia and Bulgaria to create systems and statutes of social welfare. Unfortunately, these programs have not been successful in Russia – and recently in Hungary and Poland – where the civil sector and political opposition have been repressed by the ruling party.

Finland - which has a long tradition of non-governmental organizations (NGOs), non-profit associations, societies, and foundations - can serve as a useful example also in the development of the third sector. The private non-profit foundations have effectively served, already for a half century, as an important grant-giving organization for funding scientific research, artistic creation, and other cultural activities. The legislation for foundations was accepted in 1931. The universities, scientific societies, and art associations have established

national and international network on the basis of personal contacts and agreements. The activities of citizens are also enhanced by systems which enable them to participate in different levels in national and local administration and decision making (e.g. referendums, political movements and parties, laymen members in courts, consumer boards, technology assessment panels, scientific and art councils, cultural boards, industrial democracy, self-management, participation in municipal planning).

The voluntary third sector, if cultivated with skill and imagination, gives opportunities for the creative activity of the citizens. In a future welfare society, this sector will play an increasingly important role. This can be seen already today, when the funds available in the state sector and individual consumers are decreasing. If the third sector wishes to remain as an autonomous organ for socially responsible goals of the citizens, it has to take care of its independence: it should not allow its functioning to be dominated by the national political interests or by the commercial aims of the business world.

But the third sector as such is in no contradiction with the other two social sectors, public and private, but can also work in co-operation with them. For example, the Finnish Cultural Foundation receives donations from individual citizens, and invests the funds mainly in the stock market, but decisions about grants are made by independent boards strictly on the basis of novelty and high quality of the applications. The grant programs of foundations complement the state systems of science and art councils.

Justice and Freedom

Besides efficiency, justice is a crucial standard of evaluating social systems and structures. Problems of retributive justice were the first to be treated in the laws of the early societies (e.g. the Code of the Babylonian king Hammurabi: "Eye for eye, tooth for tooth"). Justice in the political system was first systematically discussed in Plato's *Republic*. The dialogue is opened by the sophist Thrasymachus who asserts that "the just is nothing else than the advantage of the stronger" (338c). Socrates, in his turn, argues that a ruler should not "consider and

enjoin his own advantage but that of the one whom he rules and for whom he exercises his craft" (342e).

Plato's own solution to the organization of the state is based on a functional class division between workers, soldiers, and masters: "the having and doing of one's own and what belongs to oneself" is justice (434a). Plato - with Hegel and Marx - has been criticized by Karl Popper as a totalitarian and undemocratic "enemy of the open society" (cf. Chapter 21 below). In fairness, it should be added that Plato's master philosophers gain their position only through long education, which is open to men and women; they get acquainted with the idea of goodness and learn to regard "justice as the chief and the one indispensable thing" (540e). Pursuit of common or general good, instead of the selfish interests of the governors, is characteristic also to the political systems accepted in Aristotle's *Politics*: monarchy, aristocracy, and polity.

Modern political debates have centered on the relations between the state and its individual citizens. The Hegelian tradition, continued by Karl Marx, sees the state as a kind of super-individual which is more real and important than its individual members. The social and educational thinking in Finland was decisively influenced by our "national philosopher" Johan Vilhelm Snellman, who published his *Läran om staten* ("Theory of State") in 1842.[3] As a Hegelian, he made a distinction between the family, civil society, and the state. The civil society (*bürgerliche Gesellschaft*) is the domain of social life where citizens enter as individual bearers of rights; in pursuing their needs their success requires co-operation. Here Hegel and Snellman follow the Aristotelean view that man is not an atomistic individual, but a social being, a "political animal". Even though Hegel ultimately asserted that the potentials of a human personality - *Bildung*, self-consciousness and morality (*Sittlichkeit*), through education and knowledge - can be fully realized only in a state, and thereby differed from the liberal tradition of individualism, his conception of a civil society is an indispensable element of his social doctrines.

The modern individualist conception of human beings led first, in Machiavelli's *Il Principe* (1513), to the return of the doctrine of Thrasymachus: a ruthless prince gains power and

---

[3] For Snellman's account of the university, see Chapter 13 above.

prosperity for himself by exploiting his adversaries. Hobbes suggested, in his *Leviathan* (1651), that in the natural state "man is a wolf to another man", and therefore the task of the state is to protect citizens from other citizens. The tradition of liberal democracy, from Locke to Mill, led to ideas of special political rights (such as freedom of thinking, universal suffrage), expressed in declarations, laws, and constitutions of nation states.

The most radical form of individualism, in its fierce opposition to organized state, leads to anarchist doctrines. With injustice to Charles Darwin, the so-called "social Darwinist" views emphasize man's will to power: society is a jungle with a merciless struggle for existence and competition between all.

Today the welfare state is often blamed for excessive paternalism and domination which subjects its citizens to serfdom. This is the view of the libertarian (or neoliberalist) philosophy: it wants to restrict the tasks of the state merely to a "night guard"[4], which only guarantees the continuing opportunities of selfish competition in the free market. This emphasis on the value of liberty fails to distinguish between socialism and welfare society - when socialism as "the enemy" has disappeared from the scene, the new target is chosen to be the Nordic model (and similar principles of justice and equality in other societies).

This criticism confuses two notions of freedom: *negative freedom* ("freedom from") as the absence of constraints and prohibitions, and *positive freedom* ("freedom to") as the possibility and capability of doing something.[5] The basic rights of the citizens of liberal Western democracies include many negatively defined freedoms - such as the freedom of speech, thinking, and trade. The view of a libertarian is restricted to this domain of rights. On the other hand, welfare society is characterized by principles of positive freedom as well - such as the rights to education, work, childcare, aid to the elderly, and health care. Such social and educational rights were discussed in Finland in the 1990s by the commission of basic rights chaired by Pekka

---

[4] See Nozick (1974).

[5] Feinberg (1971) has argued that there is only one concept of freedom, since positive freedom can be defined as the absence of negative constraints. But I think it is artificial to describe, e.g., wealth as the lack of lack of money.

Hallberg. The existence of such rights gives opportunities and capabilities to the citizens and thereby improves their quality of life. It is of course an important issue how large domain of positive freedom a society can afford to guarantee, but it is clearly a propagandist mistake to urge that such freedoms or rights would be examples of domination and serfdom.

The Nordic model has thus justice on its side. This is an aspect which is often neglected in the debates, where the organization of society is seen merely as a kind of engineering problem.

These observations give us a perspective to evaluate the libertarian claims that the social security system and the large public sector of the welfare society make the citizens passive, institutionalized, and unmotivated. What is proposed as a cure is individual responsibility, initiative, and enterprising spirit. I think it is easy to agree on the desirability of the citizens' activity - this follows from the European ideal mentioned above. However, it should be clear that a simple recommendation or command "Why don't you try! Be active! Do something for yourself!" - directed at a person who may be uneducated or well-educated, unemployed or burdened with work, sick or healthy, poor or rich, homeless or wealthy, oppressed or governor - may be sometimes encouraging, but is often insulting and ineffective paternalism, if it is not associated with some positive element. The activity of citizens cannot be created from nothing without favorable conditions, positive rights and freedoms, that can be established through legislation and public and private funding systems.

Today the defenders of the welfare society are sometimes labeled as "value conservatives". Maybe this is appropriate. But I should like to add that the opposite view of libertarianism, often marketed as "progressive", seems to be a revival of old 19th century ideas of social Darwinism: society is seen as a jungle of egoistic beast playing a zero-sum game of the survival of the fittest. Against this emphasis on struggle and competition, the communitarian ethics seeks advantages in mutual co-operation and respect between citizens and nations.

Social Justice

The definition of social justice is of course a philosophical question, which requires an ethical position. In the theories of distributive justice, the problem is to find the most appropriate allocation of social goods and harms, benefits and burdens, to the citizens.

The basic formal ingredient of justice is the Rule of Equity: persons in the same or similar situations should be treated equally. As Feinberg puts it in *Social Philosophy* (1971), "like cases are to be treated alike and different cases differently". For example, equality with respect to law, and the universal right to vote in parliamentary elections, are applications of this principle.

However, equity should not be understood to require that all people ought to be like each other in all respects. Following Rawls (1971), procedural justice should be distinguished from the justice of outcomes or results. Thus, equal treatment may be taken to mean equal opportunities rather than actually even distribution of all good things. Moreover, equity is compatible with the rule that persons may be treated in different ways, if their positions differ in relevant ways. Substantial theories of justice should tell what a "relevant difference" may mean areas like wealth, health, and political power.

According to Feinberg, relevant differences include only such conditions for which their possessor can be held responsible, i.e. the person has had a fair opportunity to acquire or avoid them. Therefore, race, gender, age, I.Q., handicap, or social class cannot be used as a basis of negative discrimination. However, such properties may nevertheless be grounds for positive affirmative action which aims at the improvement of the opportunities of a group in a weak social position.

The most usual proposals for relevant differences include human needs (e.g. handicapped persons need more health care than others in order to have good life), merits and achievements (e.g. the most qualified applicant is appointed to a job), contributions and efforts (e.g. business profit is shared in proportion to the investment and amount of work by different parties).

It is clear that the application of these ideas to concrete situations may be very difficult, and have to be solved case by case.

In particular, it is notoriously difficult to analyze human needs - to distinguish subjective wishes from more objective preconditions of good life, or to weigh the basic physiological needs (food, shelter, health) against other more abstract "needs" for security, friendship, love, education, knowledge, and art. For this reason, the basic principle of classical utilitarianism (Bentham, Mill), viz. the greatest happiness for as many people as possible, is ambiguous. The same is true of Harsanyi's more recent formulation, where a good action should maximize average welfare.[6]

Utilitarian ethics is problematic for the reason that it does pay attention to the differences or variance in the distribution of social goods. But the attempt to minimize variance, i.e. to establish equal distribution, is likewise problematic. A middle way has been sought in John Rawls' famous treatment of justice as fairness in his *A Theory of Justice* (1971). Rawls does not demand that everything should be distributed to the citizens in equal portions. Rather, his ingenious device is the suggestion that the principles of society should be fixed behind a "veil of ignorance", without knowing in advance one's own position in the society. Even though everyone may pursue his or her interests after entering the society from behind the veil, initially the task is not to think what kind of society would be nice for me and my friends. Rawls thereby combines the condition of "the most extensive basic liberty" with a Difference Principle: "All social values - liberty and opportunity, income and wealth, and the bases of self-respect - are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone's advantage" (in particular, to the benefit of those people who are worse off). With some simplification, the justice of a society depends on the well-being of those who have the worst position.

By this standard, the Nordic model of welfare society is more just than its rivals. By the same measure, the attempt to open Russia to market economy did not improve social justice, in spite of the increase of average income, since the wealth concentrated in the hands and foreign bank accounts of the oligarchs and political elite.

---

[6] For utilitarianism, see Matti Häyry (1994).

The importance of the idea of positive freedom is even increased in the account of justice by Amartya Sen: instead of discussing the means of material wealth, his theory of justice is based on the distribution of human capabilities and abilities (cf. Chapter 18 above).

There is an important further challenge to the theories of justice. It concerns the rights and interests of future generations. Discussion of this difficult problem has been opened by David Gauthier (1986) in his account of morality by agreement. Welfare society has to be moderated to limits which make sustainable development possible.

Another special challenge, especially relevant in the development of the EU legislation, is to analyze the conditions of the economic market. Some authors (like Pentti Malaska) make a distinction between capitalism and market economy. Capitalism is the economic system that allows private ownership and fosters ruthless egoistic competition for one's own advantage. Market economy, on the other hand, should view economic competition as a means to common good (as in Adam Smith's assumption of the invisible hand). Free economic competition should mean freedom from manipulation (e.g. by big international companies), which may partly explain the paradox that free trade in Europe has to be guaranteed by an enormous amount of rules and directives. Further, market economy should be based upon principles concerning the ethical responsibilities of business enterprises with respect to the consumer and the natural and social environment.

Could the European Union be a form of co-operation that helps to bring about this ideal of market economy? Potentially misleading talk about "free trade" should then be replaced by the term *fair trade*, which has been used when Europe is trying to get rid of the "sins" of its colonial past. Fair trade has become an urgent topic of discussion also in connection with two recent developments, which show that under extraordinary circumstances the values of health and freedom may be primary to economy.

First, the national measures of protecting citizens from the covid-19 pandemic involved - besides medical means like vaccinations, masks, and security intervals - lockdowns, heavy restrictions on travel, work life, schools, cultural activities, and

restaurants. The policy "health first" had to be balanced with subsidies concerning the losses of the economic sector.

Secondly, Russia's unjustified and brutal attack on Ukraine generated a unified response from the EU and its member states: military aid to the Ukrainian army to defend the territorial integrity and sovereignty of the country, financial support to the suffering citizens and millions of refugees, sanctions against Russian state-owned enterprises and financial institutions. Russia is financing its war by selling energy to Europe, so that the European countries are now trying to loosen their dependence of Russian gas and oil – even though these moves at least in the short run slow down economic growth and accelerate inflation. But this is the price to be paid for the fact that Ukraine is defending its freedom and the whole of Western Europe against violent invasion by its aggressive neighbor.

Note. This updated chapter is based on two lectures and publications: "Active Citizens in Welfare Society", in *Euroforum 94 Seminar*, May 25, 1994, Helsinki; published by the European Cultural Foundation, Helsinki: Finnish National Committee, 1996, 7-17; "Principles of Justice, Law, and Welfare Society", in E. Paasivirta and Kirsti Rissanen (eds.), *Principles of Justice and the Law of the European Union*, Proceedings of the COST A7 Seminar, Hanasaari, Helsinki, Finland, October 17, 1994, published by KATTI B:4, Institute of International Economic Law, Helsinki: University of Helsinki, 1995, 5-22.

Literature

Aarnio, Aulis (1987). *The Rational as Reasonable*. Dordrecht: D. Reidel.

Aarnio, Aulis (2011). *Essays on the Doctrinal Study of Law*. Dordrecht: Springer.

Feinberg, Joel (1971). *Social Philosophy*. Englewood Cliffs, NJ: Prentice-Hall.

Gauthier, David (1986). *Morals by Agreement*. Oxford: Oxford University Press.

Häyry, Matti (1994). *Liberal Utilitarianism and Applied Ethics*. London: Routledge.

Niiniluoto, Ilkka (1985). "Truth and Legal Norms". In Neil MacCormick, Stavros Panov & Luigi Lombardo Vallauri (eds.), *Conditions of Validity and Cognition in Modern Legal Thought*, Archiv für Rechts- und

Sozialphilosophie (ARSP), Beiheft 25. Stuttgart: Franz Steiner Verlag Wiesbaden GMBH, 168-190.
Niiniluoto, Ilkka & Löppönen, Paavo (1994). *Suomen henkinen tila ja tulevaisuus*. Helsinki: WSOY.
Nozick, Robert (1974). *Anarchy, State, and Utopia*. New York: Basic Books.
Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Chapter 20: Technology Policy in a Liberal State

Technology policy is usually understood as a system of organizations and activities which allocate public funds to stimulate and support research and development in industrial companies and firms. It might seem that such a state-centered conception is outdated, belonging to the era of nation states with planning economies: perhaps a liberal state needs no technology policy at all? In this chapter, I argue that this would be a wrong conclusion. The development of technology is not only an economic phenomenon, a business enterprise in the free market, but rather its effects penetrate into the whole culture and society. Technology is a thoroughly political form of activity which both depends on human valuations and influences social values. A rational liberal state needs more, not less, technology policy, but its main emphasis should be in democratic methods and decision procedures that allow the citizens to participate in the assessment and control of technological change. But this requirement does not mean that truth claims in science should be placed under democratic steering. Technological progress should be based upon objective scientific truth and humanistic ethical values. In this way, technology might be able to fulfil its great promise to enhance human positive freedom.

Is Technology out of Control?

Technological progress is a central factor influencing industrial economies. Research and science-based development are productive forces. This situation is not changing in the "post-industrial" or "post-capitalist" society with Big Science, high tech, information technology, computers, automatization of production, mass media, and content industries. Technology continues to be the source of the wealth and competitiveness

of nations. It is, therefore, natural that most developed countries have established national systems of science and technology policy in order to promote and accelerate technological invention and innovation. And to gain more power in this task, small countries have joined their forces to build larger economic and political units like the European Union.

On the other hand, technology also has dramatic effects on the natural and social environment. Many of them have been unintended by-products of the employment of new technical tools and machines. No one (except some visionaries like Karl Marx and Charles Chaplin) could foresee the revolutionary social consequences of steam engine, spinning wheel, work line in factories, cars, and computers. The damaging effects of our technological form of life to the health of nature (such as waste of resources, pollution of air and water, greenhouse effect, etc.) have only gradually been realized in the last decades. The scientists in the Club of Rome have argued that there are limits to the growth, green movements have arisen to conserve nature and to protect life, and many governments have endorsed the goal of sustainable development (cf. Chapter 17 above).

It is thus clear that there is an urgent need to control technological change, both to promote and accelerate it in profitable directions and to assess its effects and to inhibit its dangerous growth.

Both of these tasks, acceleration and assessment, presuppose that technology is under human control. This view may be called *technological voluntarism*, as it claims that the emergence and change of technology depends on human will, i.e. can be influenced and directed by human evaluation and intervention.

There are powerful trends in the philosophy of technology which deny voluntarism. Usually, they are based on the idea that modern technology has become an independent system or monster which follows its own "inner logic" and directs "technological imperatives" to us. Such *technological determinism* has been formulated, as a horror picture, by Jacques Ellul (1964). In a milder form, Langdon Winner (1977) has defended the thesis that technology has become "autonomous", and today determines politics, rather than vice versa. But determinism has also been advocated by the spokesmen for *technocracy*, who suggest that for ordinary citizens it is simply the best to accept

and passively follow the advice of the technical experts who know where technology is going by its own inner necessities.

Besides technological voluntarism and determinism, the third main alternative is to claim that technological change is indeterministic, chaotic, and unpredictable. One formulation of this view is *decisionism*: value choices are arbitrary, subjective, and situational. This position is in fact an extreme version of voluntarism. Jürgen Habermas (1970) criticizes it by defending the possibility of rational discourse on values.

Another expression of the "contingent" and "heterogeneous" nature of technology is the *constructivist* approach in the STS-studies. Wiebe Bijker and John Law (1992) suggest that "sociotechnology" constitutes a "seamless web", where society and technology cannot (or should not) initially be distinguished, so that one of them does not dominate the other. In particular, this means that there are no predetermined trajectories for the evolution of technological projects.[1]

I think all the three approaches have their insights and merits, but they do not describe any inherent nature that technology has independently of us. Naive voluntarism is certainly an illusion. Technology is a powerful social system which, in Émile Durkheim's sense, has coercive power over us. Technological development does present its "imperatives" to us, but they are always conditional on some value premises (about what should be desired or avoided), and therefore we always have the option to disobey them.[2] Even though progress in technology may to some extent influence our goals and values, our choices need not be random, non-purposive, or externally determined, but also the social needs and ends have to be publicly and freely discussed in a democratic community. I think we have to recognize that it is a matter of a major political decision whether we *allow* technology to develop in the deterministic or chaotic manner or whether we wish to find democratic ways of assessing and controlling it.

---

[1] Cf. Bijker (1995).

[2] In Niiniluoto (1990), I argue that the logical form of the so called "technological imperatives" is what G. H. von Wright calls technical norms: "If you want A and believe that you are in situation B, then you ought to do X" (cf. Chapter 11 above). The value premise A is hidden, which gives the misguided impression that such command are unconditional imperatives. See also Niiniluoto (2017).

Democracy, Liberalism, and Freedom

The technocratic view surrenders the power in society to an elite which is assumed to guide others by their expertise without value choices and commitments. Another challenge to democracy arises from the position which admits that technological change is value-laden but claims that these values can be objectively known by a philosophical elite who can settle for others what is really good and right. This view, formulated already by Plato, has recently reappeared as value objectivism in environmental philosophy.[3]

It is conceivable that some day the "ecological imperatives" become so pressing that the Western model of democracy is unable to cope with them. This is not an issue to be discussed in this chapter, which is focused on the prospects of technology policy in a *liberal democratic state* (cf. Chapter 21 below). In such a state, the citizens enjoy personal freedom and rights. They are allowed to have different preferences, and there are systematic procedures for guaranteeing that different valuations are taken into account in social decision-making. In some form, market economy also is the cornerstone of a liberal state. But these conditions still leave open many alternative ways of building a democratic state.

Modern political debates have centered on the relations between the state and its individual citizens.[4] The Platonist, Hegelian, and Marxist traditions see the state as a higher-order entity which has primacy over its individual members. The modern individualist conception of human beings originated with the 14th century nominalists. It was a basis for the tradition of liberal democracy, from Locke to Mill, which led to ideas of special political rights (such as freedom of thinking, universal suffrage), expressed in declarations, laws, and constitutions of nation states. The Nordic model of welfare society is a sort of compromise between the Hegelian and liberalist traditions (cf. Chapter 19 above). It has involved a conception of a strong state, based upon democratic constitution, and large public sector which takes care of security networks, allowances, and many other welfare services to all citizens. But

---

[3] See Skolimowski (1981).

[4] In political science, the corresponding opposition is between "holists" and "methodological individualists".

the Nordic model has also given room for a large "third sector" with non-profit organizations and the private businesses. In contrast, the Catholic model of welfare has given more emphasis to the family and the church, while the American model relies largely on private insurance policies and the economic market.

Since the 1980s, many Western countries have been struggling with economic problems of welfare society, and *libertarian* views of the New Right have strongly influenced their politics. Deregulation, cuts in the public sector, commercialization, privatization, individual responsibility, initiative, and enterprising spirit are suggested as remedies against the nurturing and paternalizing tendencies of the welfare society. Libertarianism understands freedom in the negative sense as the absence of constraints. This view is often called "neo-liberalism", even though it seems to be basically an old doctrine, a repetition of the ideas of the 19th century social Darwinism: the whole of society is seen as a field of merciless struggle and selfish competition, where the strongest and fittest survive.[5] This view emphasizes man's will to power. The role of the state should be minimal: its task as a "night guard" (Robert Nozick) is to make sure that the conditions of free competition are maintained. But the libertarian criticism of welfare state for paternalism confuses two notions of freedom: *negative freedom* ("freedom from") as the absence of constraints and prohibitions, and *positive freedom* ("freedom to") as the possibility and capability of doing something. The basic rights of the citizens of liberal Western democracies include many negatively defined freedoms - such as the freedom of speech, thinking, trade. The view of a libertarian is restricted to this domain of rights. On the other hand, welfare society is characterized by principles of positive freedom as well - such as the rights to education, work, childcare, aid to the elderly, and health care. The existence of such rights gives opportunities and capabilities to the citizens (cf. Chapter 18 above). The democratic welfare society fares better than the libertarian which reduces questions of justice to negative freedom and efficiency in the market (cf. Chapter 21 above).

---

[5] This view belongs more to Herbert Spencer than Charles Darwin.

Science and Technology in the Free/Fair Market

Different conceptions of liberalism may lead to various solutions concerning the relationships between science, technology, and the market.

For a libertarian, technology should be seen as a way producing and distributing commodities in the free market. Firms, based on private ownership, produce goods with the aim of maximizing profit, and the consumers have complete freedom in making their own choices according to their own needs and preferences. In deciding to buy commodities, the consumers are as if "voting" for alternative technological products, so that the whole procedure appears to satisfy the ideal of democracy as well.

Science, both basic research and applied research, serves as basis of technological development. For the libertarian, this motivates the treatment of scientific research and higher education as "investments" which should yield economic profit in the short or at least not-too-long run. Science-based development brings about new *innovations*, i.e. products which as novel goods and services can be sold as commodities in the economic market.[6] This means that both science policy and technology policy are understood to be parts of the economic strategy concerning commerce and industry. For example, the Finnish Council of Science and Technology Policy, in its programmatic *Towards an Innovative Society* (1993), a development strategy for Finland as a "country of knowledge and competence", used the economic concept of "innovation system" to motivate the need of promoting R&D in Finland.[7]

When technology is seen as an effective instrument of national economy, a purely libertarian approach, with its skepticism about the state, does not work quite well in practice. It turns out that there is an optimum amount of state

---

[6] This the definition of "innovation" by the economist Joseph Schumpeter. Veugeler et al (2009) define innovation as "an exploited competence-based asset".

[7] Besides the Academy of Finland, which funds basic research, a new funding agency of technology, Tekes, was established in 1983, and in the next year the former Science Policy Council was changed to a new Science and Technology Policy Council. Today its name is Research and Innovation Council, and Tekes is Business Finland.

intervention, in the form of public subsidies and funding, which keeps technological invention and innovation in progress even in private firms. Similarly, scientific education and basic research are too important for modern nations to be left entirely to the mercy of the market. For these reasons, non-extremist libertarian thinkers like James Buchanan (1986) are willing to allow the state to finance "the enterprise of science" as a public good.[8]

Neo-liberal influences can also be seen in the currently fashionable competitive approach to scientific research and education. The success of science is not measured primarily by economic indicators, but in terms of quantities expressing numbers of publications and citations. Science is viewed as a kind of sport where the competition is between individual researchers, groups, departments, institutes, universities, and nations. Research and universities should be goal-directed, cost-benefit effective, managerially directed, and accountable in such terms. This leads to what might be called *Matthew strategy* in science policy: give more to those who already have, and take away from those who have less.[9]

John Ziman (1994) has described the recent trend in science policy as a change from the Mertonian ethos of Communalism, Universalism, Disinterestedness, Originality, and Skepticism (CUDOS) to Proprietary, Local, Authoritarian, Commissioned, and Expert work in science (PLACE). Ziman notes that the CUDOS norms have served to guarantee the progress and productivity of science by leaving space for personal creativity, openness to debate, and hospitability towards novelty. He illustrates vividly how the new PLACE principles are realized in the new global and national organization of Big Science, the research programs of the European Union, contract research in the universities and in industry, and the changing conditions of scientific careers. And he raises the important question, whether this reshaped system is still able to sustain the progress of scientific knowledge. Ziman expresses his strong misgivings about the new slogans and trends of science policy. Many of the new practices, he argues, are "so ill-judged that

---

[8] See Loukola (1995).

[9] I have used this term first in 1984 in analogy with Merton's (1973) descriptive term "Matthew effect" (see Niiniluoto, 1987).

they could do lasting damage to the health of science and its efficacy as a social institution" (p. 252). In particular, "accountability" lays stress on the narrowly instrumental aspects of science and is unfavorable to genuine scientific originality. "Evaluation" is counter-productive if carried out too frequently. "Selectivity" enforces the Matthew Effect and often leads to arbitrary decisions to "close down small, but beautiful research operations". The idea of producing "trained research workers" as a kind of "manpower" devalues the personal commitment that motivates good scientists. "Exploitation" of academic research leads to the "commercial patronage of university". Emphasis on "priorities" is a very serious threat to the integrity and credibility of the social sciences. Excessive "competition" undermines the communal spirit of inquiry. "Management" leads to bureaucratic top-down policies that are incompatible with personal and departmental autonomy.

On the whole, I think there is reason to agree with Ziman's analysis, but with three reservations. First, Ziman over-emphasizes the "individualism" of CUDOS against the "corporate" or "collectivized" spirit of PLACE. Science has been based on co-operation within the scientific community at least since the mid-seventeenth century. Secondly, creative work within Humboldtian universities may very well be highly successful in quantitative terms (cf. Chapter 13 above). Thirdly, the strategy of building "centers of excellence" may turn out to be useful for the whole research community, and thus justifiable by the Rawlsian difference principle.

In my view, the best way of philosophically defending the autonomy and integrity of scientific inquiry against reduction to the commercial principles of technology policy is provided by scientific realism and the conceptual distinctions between science and technology and between basic and applied research.[10] A step in the same direction is made by Timo Airaksinen's (1995) discussion of the role of "theoretical virtues" in the professional practice of science (such as wisdom, creativity, clarity, consistency, and criticism), and their difference to the "Franklinian virtues" in financing science and the "Machiavellian virtues" in profiting from science.

---

[10] See Niiniluoto (2017).

Support from other potential sources is not easily available. The communitarians, who pointedly criticize narrowly individualist and egoistic ethical principles, are overly skeptical about the prospects of science and Enlightenment.[11]

In the study of science, technology, and society (STS), there has been a popular view that "scientific knowledge has no privileged claim to truth", all theories and belief systems are "on a common epistemological footing", the scientists are not guardians of objective knowledge but "hired brains of special interests and lobbyists for their own".[12] The message of these STS-studies seems to be devastating: everything that you might have been afraid that could happen to science, is already true.

I agree that, in the study of the role of politics in shaping the scientific production of knowledge, it is important to be free from ideological prejudice in any direction. But the by now popular thesis that science has no rational or epistemological advantage over other practices of belief formation has itself become a prejudice in the STS-studies. If it were true, why should the society spend any funds at all to such an expensive activity as research - it would be cheaper and faster to consult fortune tellers using crystal balls and tarot cards.

Scientific research always takes place in a social context which influences its organization, orientation, and sometimes even the content of knowledge. The results of research are fallible and revisable, but - as long as the methods of scientific inquiry are employed - they are constrained by what is *true* and what kind of *evidence* is available. The basic "epistemic utilities" in the tentative acceptance of scientific hypotheses are such cognitive factors as truth, information, explanatory power, predictive power, accuracy, and simplicity. Such criteria are relevant also in applied science, which seeks instrumental knowledge that has practically significant applications and, hence, commercial value in the market (cf. Chapter 11 above).

The legitimate domain of technology policy is quite difference from that of science policy. Technology does not produce knowledge like science, but rather designs new artefacts, tools, and machines. They are not constrained by truth in the same

---

[11] Cf. Hellsten (1995).

[12] See Cozzens & Woodhouse (1995).

way as knowledge claims, but by what is physically and economically *possible*. This makes technology "value-laden" in a way which differs from science.[13] Therefore, the criteria for the acceptance of technological products are different from the epistemic utilities in science. Commercial value in the market is one of such factors, in our society even the primary criterion for the innovation and diffusion stages in the life-cycle of artefacts. But it need not be the only relevant factor. This is basic issue of technology assessment.

Facts and Values in Technology Assessment

Alex Michalos (1983) has argued that it is useless and even dangerous to appeal to the fact - value distinction in technology assessment. A deconstruction of such a distinction is also given by the constructivist approach. It denies the use of facts about artefacts as explanations of how they work: machines work because they have been chosen by relevant social groups, not vice versa.[14]

In my view, it is of utmost importance to make a clear distinction between the objective properties of an artefact and the value criteria in its assessment. For example, a car has a shape and color, and its engine has efficiency measurable in horse powers. The behavior of the car (e.g. its ability to carry passengers, its maximum speed) is a function of these facts.[15] To be sure, these properties are results of design, and in this sense depend on our decisions, but, when the car has been produced, they are as objective facts as the color of a bird and the material constitution of a tree. In the same way, the "technical norms" sought in applied research have a factual content about the relations between means and ends. The task of establishing facts about artefacts belongs to the scientific and technological experts.

The properties of artefacts make them *tools* which may have *instrumental value* relative to human purposes. Each technological artifact has an intended function, and its *effectiveness* or "instrumental goodness" depends on its ability to serve or fulfil

---

[13] I have elaborated this conception of technology in my Finnish book *Tekniikan filosofia* (2020). Cf. Bugliarello & Doner (1979).

[14] See Latour (1987), Bijker (1995).

[15] See Sahal (1987).

this function. Intended functions are sometimes specific (e.g. spade), sometimes open-ended (e.g. computer). They may be culturally and historically relative: a tool may change when it is transferred to another social context or when it is placed into an exhibition in a museum.

By their effectiveness, technological tools open up new resources and possibilities of human action. Thereby they increase the domain of human *positive freedom*. Such increase of freedom may be a good or bad thing relative to human values. Such possibilities are created by utilizing some resources, and the use of tools has also often unintended and unwanted by-products and side effects. Besides the misuse of tools and their harmful effects on nature and society, such by-products include increased risks.[16] This suggests a straightforward utilitarian calculus of evaluating technological projects by their costs and benefits.

Each tool and the related novel possibility can be evaluated by several criteria. One of them is the monetary exchange value of the product, or its *economic* profit, which is the domain of economic theories.[17] This is the only criterion which is standardly applied, when technology is viewed as a commodity in the economic market. But effectiveness and economy are not the only dimensions of technology assessment. As products of design, artefacts have *esthetic* qualities, studied today in applied esthetics. The relations of tools to the health of their users are studied in *ergonomics*. The relations of human technologies to the health of natural environment are treated in *ecology*. Technology has also an impact which is *social* in the broad sense (legal, institutional, political). Further, technical tools and their effects can always be evaluated by *ethical* standards which concern their moral worth.

In the 1970s, several Western industrial countries developed systems of *Technology Assessment* (TA) for the evaluation of large-scale technological projects.[18] In the United States, the Office of Technology Assessment (OTA) was established in 1973 - and closed in the 1990s. My remarks suggest that the

---

[16] See Shrader-Frechette (1991).

[17] Cf. Elster (1983).

[18] Cf. Durbin & Rapp (1993).

basic method of technology assessment can be expressed by the formula

TA = 6E + S,

where the six E's are: effectiveness, economy, esthetics, ergonomy, ecology, and ethics, and the S refers to the social impact and consequences of technology.

Technologists have sometimes misunderstood the nature of TA by believing that numerical indicators are always "objective" and "value-neutral". As noted by Shrader-Frechette (1983), they have also tended to ignore those dimensions that cannot be neatly quantified. The same problems are common in the use of numerical performance indicators of scientific progress: the numbers of exams and publications are easy to calculate, and questions about quality are easy to forget.[19] It is better to see such indicators as attempts to express and articulate values. There are specific issues in the "measurement" of each dimension of assessment (such as esthetic or ethical value), and it is ultimately a value question as well to balance or weight the dimensions relative to each other. The situation is basically the same in science and technology, even though the relevant epistemic and technological utilities differ from each other.

In a democratic state, the citizens should have the right to decide how much public funds are allocated in the state budget to research and education. However, in matters concerning scientific quality and the scholarly significance of research, the scientific community should have autonomy in the society at large, and it should function as an expert system rather than a democracy. When a scientific project or publication is evaluated, the members of the scientific community do not vote about the issue, but the best expert in the field is asked to do the assessment. This is the basic idea of *peer review* in science.[20]

Technological experts have also a special role in evaluating facts about technological constructs and rules, but their task is not to decide the value questions about the development of technological innovations. The decisionist strategy would lead

---

[19] See Niiniluoto (1987).

[20] See Merton (1973), Ch. 21, on the birth of the referee system in science.

to the libertarian market model of delegating decisions about the use of technological products to individual consumers. This populist idea in fact easily gives the power back to the technocrats, since human needs and hopes are constantly influenced and manipulated by marketing and advertising. Still, it may work well for some "innocent" artefacts: when a rational person chooses his or her toothbrush, we have no reason to exercise paternalism over such a decision.[21] And in some cases, market rationality may lead to socially desirable results, if the individuals have enlightened preferences (e.g. use electric cars and "eco-products").

But most commodities are not private affairs, as they may have harmful environmental and social effects. When it turns out that a deodorant affects the ozone layer, it becomes a dangerous tool. As Liisa Uusitalo (1986) has shown by applying game-theoretical models of "free riding" to consumer behavior, what seems individually rational may lead to collectively irrational consequences - unless some moral or legal constraints on our actions are accepted. Most of us approve some restrictions on the selling of drugs and guns. (Unfortunately, Archie Bunker's logic still has some popularity: he suggested that, in order to prevent hijacking, every passenger should carry a gun in a plane!) Some decisions about large-scale technological developments are socially so important and difficult that they are handled in democratically elected boards. For example, in Finland the Parliament has the right to make decisions about the program of building nuclear power plants.

Technologists should feel responsibility for the uses and effects of their inventions, and it is important that the engineers have agreed about codes of professional ethics. They should be ready to co-operate with other professions who have expertise in the different domains of valuations: medical doctors and psychologists in ergonomics, industrial designers and architects in esthetics, applied philosophers in ethics, sociologists and lawyers in the social studies, etc. These other professions should play an important role, when technology assessment is organized and administrated in a systematic way. But they in turn should be ready to work in co-operation with laymen, who are potential consumers of technological tools and

---

[21] See Häyry (1991).

methods. For example, consumer panels, including surgeons and women, have been used in the development of methods of treating breast cancer. In many issues about the quality of life, ordinary citizens are the best experts about their own attitudes and feelings.[22]

Stanley Carpenter (1983) has argued that technology assessment as a program is not sufficiently radical, since its own cost-benefit methods rely on technological and utilitarian ideas. He suggests that TA should be replaced by AT, i.e. Alternative Technology, advocated by "green" environmentalist movements and proponents of "appropriate", "soft", and "participatory" technology. I don't see that there is real contradiction between TA and AT. To promote democratic procedures in technology policy, we need participation of many interest groups on several different levels. We need state and city councils, political organizations, international co-operation, legislation on the treatment of wastes in industrial production, bureaus protecting consumers for unfair commerce, self-reflection by technological professions, teaching of engineering ethics, public debate on environmental issues, value discussion about good life, NGOs (like WWF), and active citizens working in free groups and societies.

In brief, rational technology policy in a democratic state should be based upon co-operation between the public, private, and "third" sectors. In this way, we might hope that technology would better redeem its promise to positively liberate human beings - in the terms of Macpherson (1983), not as Benthamite "infinite consumers" and "infinite appropriators" in the capitalist society, but as Millian "exerters and enjoyers" of their own unique "powers and capacities" in a liberal egalitarian democracy.

Note. This chapter is based on my article "Technology Policy in a Liberal State", in Sirkku Hellsten, Marjaana Kopperi & Olli

[22] See Schrader-Frechette (1985). In user-driven technology policy, the joint activity of technological designers and potential users is today called "co-design" or "co-creation" (see Veugelers et al., 2009). Its counterpart in science policy is the "mode 2 research", where researchers, engineers, designers, and potential customers work together to create "innovation cycles" (see Gibbons et al., 1994).

Loukola (eds.), *Taking the Liberal Challenge Seriously: Essays on Contemporary Liberalism at the Turn of the 21st Century.* Aldershot: Ashgate, 1997, 192-204.

Literature

Airaksinen, Timo (1995). "The Virtuous Face of the Ethics of Science". In Gasparski & Airaksinen, 13-26.

Bijker, Wiebe E. (1995). *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change.* Cambridge, MA: The MIT Press.

Bijker, Wiebe E. & Law, John (1992). *Shaping Technology/ Building Society: Studies in Sociotechnical Change.* Cambridge, MA: The MIT Press,

Buchanan, James M. (1986). *Liberty, Market and the State: Political Economy in the 1980s.* Oxford: Oxford University Press.

Bugliarello, G. & Doner, D. B. (eds.) (1979). *The History and Philosophy of Technology.* Urbana: University of Illinois Press.

Carpenter, Stanley (1983). "Technoaxiology: Appropriate Norms for Technology Assessment". In Durbin & Rapp, 115-136.

Cozzens, S. E. & Woodhouse, E. J. (1995). "Science, Government, and the Politics of Knowledge". In S. Jasanoff, G. E. Markle, J. C. Petersen & T. Pinch (eds.), *Handbook of Science and Technology Studies*. London: SAGE, 533-553.

Durbin, Paul & Rapp, Friedrich (eds.) (1983). *Philosophy and Technology*. Dordrecht: D. Reidel.

Ellul, Jacques (1964). *The Technological Society.* New York: Alfred A. Knopf.

Elster, Jon (1983). *Explaining Technical Change.* Cambridge: Cambridge University Press.

Gasparski, Wojciech & Airaksinen, Timo (eds.) (1995). *Science in Society*. Warsaw: IFIS Publishers.

Gibbons, Michael et al. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

Habermas, Jürgen (1970). *Toward a Rational Society*. Boston: Beacon Press.

Häyry, Heta (1991). *The Limits of Medical Paternalism*. London: Routledge.

Hellsten, Sirkku (1995). "The Communitarian Approach to Science in Society". In Gasparski & Airaksinen, 27-43.

Latour, Bruno (1987). *Science in Action*. Milton Keynes: Open University Press.

Loukola, Olli (1995). "On Libertarian Science Policy: Does Science Need Nurturing?". In Gasparski & Airaksinen, 51-64.

Macpherson, C.B. (1983). "Democratic Theory: Ontology and Technology". In Carl Mitcham & Ruth Mackey (eds.), *Philosophy and Technology*. New York: The Free Press, 161-170.

Merton, Robert S. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: The University of Chicago Press.

Michalos, Alex (1983). "Technology Assessment, Facts and Values". In Durbin & Rapp, 59-81.

Niiniluoto, Ilkka (1987). "Can Evaluation Be Applied in Universities?", *Tiedepolitiikka* No. 3, 11-18.

Niiniluoto, Ilkka (1990). "Should Technological Imperatives Be Obeyed?", *International Studies in the Philosophy of Science* 4, 181-189.

Niiniluoto, Ilkka (2016). "Science vs. Technology: Difference or Identity?". In Maarten Franssen, Pieter Vermaas, Peter Kroes & Anthonie Meijers (eds.), *Philosophy of Technology after the Empirical Turn*. Cham: Springer, 93-106.

Niiniluoto, Ilkka (2017). "G. H. von Wright on the Dangers of Technology". In Ilkka Niiniluoto & Thomas Wallgren (eds.), *On the Human Condition: Philosophical Essays in Honour of the Centennial Anniversary of Georg Henrik von Wright*. Acta Philosophica Fennica 93. Helsinki: Societas Philosophica Fennica, 421-431.

Niiniluoto, Ilkka (2020). *Tekniikan filosofia*. Helsinki: Gaudeamus.

Sahal, Devendra (1987). *Patterns of Technological Innovation*. Reading: Addison - Wesley.

Science and Technology Policy Council of Finland (1993). *Toward an Innovative Society; A Development Strategy for Finland*, Helsinki.

Shrader-Frechette, Kristin (1983). "Technology Assessment and the Problem of Quantification". In Durbin & Rapp, 151-164.

Shrader-Frechette, Kristin (1985). "Technology Assessment, Expert Disagreement and Democratic Procedures". In Paul Durbin (ed.), *Research in Philosophy and Technology*. Greenwich, CT: JAI Press, 103-129.

Shrader-Frechette, Kristin (1991). *Risk and Rationality: Philosophical Foundations for Populist Reforms*. Berkeley: University of California Press.

Skolimowski, Henryk (1981). *Eco-Philosophy*. Boston: Marion Bayars.

Uusitalo, Liisa (1986). *Environmental Impacts of Consumption Patterns*. Alderhot: Gover.

Veugelers, Reinhilde et al. (2009). *Evaluation of the Finnish National Innovation System – Policy Report*. Helsinki: Ministry of Employment and the Economy.

Winner, Langdon (1977). *Autonomous Technology*. Cambridge, MA: The MIT Press.

von Wright, G.H. (1963). *The Varieties of Goodness*. Routledge and Kegan Paul, London.

Ziman, John (1994). *Prometheus Bound: Science in a Dynamic Steady State.* Cambridge: Cambridge University Press.

Chapter 21: The Open Society and its New Enemies

Karl Popper's The Open Society

In the war time in 1944 Karl Popper, a Viennese philosopher in exile in New Zealand, wrote his celebrated and controversial magnum opus in political philosophy, *The Open Society and Its Enemies*. In attacking totalitarianism, the book identified three enemies of "the open society": Plato, Hegel, and Marx. When the work was published in 1945 some months after the end of World War II, and its author had taken up a new position at the London School of Economics in January 1946, a new ideological battle, the Cold War, was already beginning between the Western market economy democracies and the socialist block led by the Soviet Union. Popper's work, together with his economist friend Friedrich von Hayek's manifesto *The Road to Serfdom* (1944), became a standard text of postwar liberalism in its fight against communism. The 1957 book edition of Popper's *The Poverty of Historicism* (published originally in 1944-1945) was dedicated to the memory of "the victims to the fascist and communist belief in Inexorable Laws of Historical Destiny". When Sir Karl died in 1994, the Berlin Wall had already tumbled down and the Soviet Union had collapsed.

In *Conjectures and Refutations* (1963), Popper identified himself as a "liberal" in the classical sense: not a sympathizer of any political party, but "simply a man who values individual freedom and who is alive to the dangers inherent in all forms of power and authority" (p. viii). But soon this term gained special flavor when the British conservative party started to advocate what is usually called economic neoliberalism or

philosophical libertarism.[1] In the 1980s, during the Reagan-Thatcher era, Popper's and von Hayek's views were regularly linked to these political doctrines of the New Right. Thus, again Popper was used as a weapon in ideological debates, now against the Western models of welfare society.

But it is important to observe that some philosophers have interpreted Popper's position as "social democracy", as Brian Magee, later a MP of the British Labour Party, put it in his *Popper* (1973), or as an articulation of what the German sociologist Ralf Dahrendorf named "the Social Democratic Consensus" in 1979. It is also interesting to note that recent scholarship on the formation of Popper's views has emphasized his youthful socialist background in the Red Vienna. Malachi Haim Hacohen (2000) argues convincingly that while writing *The Open Society* during the war Popper knew nothing about the Soviet Union. He was launching a defense of democracy against fascism, not a "charter of cold war liberalism". Popper found the totalitarian fascist drive in Plato and Hegel, but Marx was treated as a progressive democrat whose main failure was historicism:[2] the belief in historical inevitability weakened the socialists' ability to confront fascism and left Central European democracies defenseless.[3]

Popper contrasted the open society with "the magical or tribal or collectivist" closed society whose laws and customs are felt to be unchanging and inevitable. In the open society, "individuals are confronted with personal decisions" and thereby they have responsibility in approving public policies.[4] Such a society tolerates and fosters the critical attitude that is characteristic to scientific thinking. It has to be democratic in the sense that the people in power may be criticised and replaced by others without violence. The open society is liberal and anti-totalitarian, as it involves a belief in "the possibility of

---

[1] It is interesting to note that an external observer, the Chinese philosopher Li Tieying (2002), mentions Keynes and Roosevelt as "neo-liberals" and classifies Hayek and Friedman as "neo-conservatives".

[2] See Popper (1957).

[3] Hacohen (2000), 383. For a view of Hegel different from Popper's, see Avineri (1972).

[4] Popper (1966), vol. 1, 173.

a rule of law, of equal justice, of fundamental rights, and a free society".[5]

Popper, who supported "piecemeal social engineering" with the principle of "minimizing avoidable suffering", did not propose the open society as an ideal or utopian goal. But it is still interesting to ask where we stand today with Popper's project of the open society. Has this form of society been realized or even approximated anywhere in the world? And who are currently its most dangerous enemies?

The Prospects of Democracy

No one can deny that democracy significantly improved its position in the last two decades of the twentieth century, especially after the fall of the Iron Curtain in 1989: according to David Held's (1999) classification, the number of authoritarian states dropped from 101 to 43, partial democracies (with institutionalised voting rights) increased from 11 to 43, and full liberal democracies (with basic civil rights and freedoms) from 35 to 78. But Freedom House reports in 2020 that global democracy has declined for the 16th consecutive year. In the Regimes of the Worlds (RoW) –classification in 2020, the world's 180 countries are divided about fifty –fifty between democratic and autocratic governments. According to the Democracy Index, published by the Economist Intelligence Unit in 2021, the number of full democracies is only 21, flawed democracies 53, hybrid regimes 34, and authoritarian regimes 59. The three highest scores are achieved by Norway, New Zealand, and Finland, and the lowest scores by North Korea, Myanmar, and Afghanistan.

One must admit that the progress of the Western model of democracy has met many difficulties.[6] Political dictators, military governments, and dogmatic religious leaders have not vanished from the earth. Democratic openness may still be limited in poor developing countries (several countries in Africa,

---

[5] Popper (1963), 5.

[6] I am not going to discuss in this chapter technical problems in implementing democratic elections (e.g. the famous voting machines in Florida) or conceptual difficulties in the analysis of democratic procedures, related to the social choice theory (e.g. Arrow's paradox, voting paradoxes).

Asia, and Latin America). Discipline rather than liberty is often taken to be the "Asian" tradition of government.[7] The new Russia with a privatized market economy has not been able to establish a stable democracy with civil security, free press, and political opposition. The Marxist theory, which contrasts "the capitalist individualistic democracy" or "the bourgeois democracy" with "people's democracy" or "people's democratic dictatorship under the leadership of the working class", is still alive in China[8] - indeed, in spite of the fact that China has cautiously opened the way to market economy, it suppressed the democracy of Hong Kong and watches and controls its citizens with new AI tools of face recognition.

In addition to such more or less familiar cases of closed societies, and the problems with building new sustainable democratic states, it is important for us also to exercise self-assessment and to critically evaluate our own nations. Perhaps we are able to locate *new* enemies of the open society? It is this issue that I take up in this chapter. I shall consider five relevant topics: (i) the welfare society and social justice, (ii) the information society, (iii) the market economy, multinational corporations, and globalization, (iv) populism, and (v) the world after 9/11 and other violent attacks.[9]

Welfare Society, Freedom, and Social Justice

A social organization may be evaluated on two grounds. First, is it *effective* in producing the desired results, and does it perform this task *efficiently* in the cost-benefit sense? Secondly, is it *just* or at least more just than its alternatives? I think the Nordic model of welfare society scores on both of these points (cf. Chapter 19 above). It is the latter aspect that is directly relevant to the question of openness, but it is relevant to consider the former as well. I will conclude that welfare society is not an

---

[7] Cf. Ishiguro (2003).

[8] See Li Tieying (2002).

[9] When the first version of this chapter was written in 2004, I failed to consider the significance of social media (Mark Zuckerberg founded Facebook in the same year). The section (iv) on populism is new, and the section (v) is much revised in the light of increasing violence in international affairs.

enemy of the open society, and the recent criticism against it is in many ways unfair.

The building of the Nordic model of welfare state was effectively started in the 1960s, even though its roots in Sweden date back to the early years of the 20th century. This model has accepted the basic principles of democracy and market economy, and its implementation in different countries has presupposed the growth of national economy. Thereby Finland successfully avoided the depressing lack of freedom in the Eastern socialism and the immense inequalities of wealth in the American capitalism. Nevertheless, the New Right movement reached Scandinavia in the early 1990s. In its ideological campaign against the large public sector and high taxes, it proposed privatization and reductions as cures to "save" the welfare society.[10]

A compromise, called "the third way", has been proposed by the UK sociologist Anthony Giddens (2000). It attempts to combine demands of global market economy, individual rights and responsibilities, and the social functions of nation states in a manner that is close to the political lines in the Nordic countries.

Today the welfare state is often blamed for excessive paternalism and domination which subjects its citizens to serfdom.[11] But this libertarian criticism confuses two notions of freedom: negative freedom ("freedom from") as the absence of constraints and prohibitions, and positive freedom ("freedom to") as the possibility and capability of doing something.[12] It is of

---

[10] In reading about such cuts, one is reminded of the old story about surgery: the operation was successful, but the patient died.

[11] This an echo from Hayek (1944).

[12] Quentin Skinner (2002), who defines "a third concept of freedom" as the autonomy of citizens with respect to the arbitrary will and mercy of the rulers, fails to appreciate the significance of positive freedom. Here is an example of a negative freedom without positive freedom: even if no one prevents me from buying a house in London or writing an essay on radio astronomy, I am not free to do so if I so decide, since I lack the relevant financial or educational resources. The importance of the idea of positive freedom is very prominent in the account of justice by Amartya Sen: instead of discussing the means of material wealth, his theory of justice is based on the distribution of human capabilities and abilities (see Chapter 18 above).

course an important issue how large domain of positive freedom a society can afford to guarantee, but it is clearly a mistake to urge that such freedoms or rights would be examples of domination and serfdom. In fact, the social security system satisfies the basic idea of John Rawls' account of justice as fairness, which recommends the distribution of all social values equally, unless an unequal distribution of any, or all, of these values is to the benefit of those people who are worse off.[13] By this standard, the Nordic model of welfare society is more just than its rivals.

Democracy, Corporations, and the Media

A widely accepted doctrine, formulated by Montesquieu in 1748, states that a good constitutional monarchy or democracy should separate its executive, legislative, and judicial powers. However, in modern societies these three powers (i.e. government, parliament, and courts of law) are not the only influential actors of the social system, and these other actors may turn out be threats to the open society.

One alternative to the Montesquieu-type of constitutional democracy is the *corporativist* model, where the society is divided into separately functioning orders - such as estates, classes, guilds, churches, professions, trade unions, and other corporations (in the broad sense) - and democratic participation is replaced by the representation of such corporations. In fascist politics, in Mussolini's Italy and Franco's Spain, such corporativism was implemented by the totalitarian one-party regime whereby the people lost their power to influence political decisions.

One of the corporations in modern society is the technological elite, often in co-operation with the military profession. Even though technology may be directed by democratic interventions and assessment to yield benefits to the quality of life and social progress (cf. Chapter 20 above), *technocracy* - or a society ruled by the technologists - is an enemy of the open society. In the technocratic perspective, the organization of the

---

[13] In my view, the Rawlsian Difference Principle should be amended by placing restrictions to unfairly large differences in resources.

society is seen merely as a kind of engineering problem and the issues about justice are reduced to mere efficiency. Thereby social and political problems are assigned to technical experts, and their relations to citizens' values and deliberation is suppressed. Here Popper's own style of speaking about "social engineering" may give a too technocratic impression of his views.

An example of an additional social force which upsets the democratic order is *organized crime*. The Sicilian mafia is a remnant of the premodern tribal or feudalist social order, but its extension to the American *cosa nostra* shows that such a threat to the open society may exist and grow in the modern capitalist society as well. It is clear that the normal operation of society is biased or suppressed if the gangsters "hold in their pockets" important but corrupted politicians, congress members, senators, or trade union leaders - by using, instead of public and rational argumentation, violence, bribery, corruption, and blackmail as their principal methods. Another recent example is the Russian mafia, created immediately with the collapse of the communist power and the opening of the private market. In the age of globalization, organized crime continues its operations effectively in the international scale.

The *press* has sometimes been called "the Fourth Estate". The significance of journalism has been great in modern states, but it has definitely further increased within the "information society", created through communication by the electronic media (telephones, radio, TV, cinema, videos, CD-ROMS, email, mobile phones, internet) and the processing of digital data by computers.

The *information society* has several different impacts to the conditions of democracy. On the positive side, following the noble goals of the enlightenment and its principles of the "freedom of information", the new media may help to disseminate genuine knowledge (i.e. true and justified information), and thereby to educate people and to improve their critical thinking (cf. Chapter 15 above). The development of ICT may also have dramatic political consequences: as Manuel Castells (2000) has argued, the citizens' access to non-censured information strengthened the position of liberal democracies and helped to overthrow the closed system of Soviet socialism.

Hence, the information society can be, in some important respects, a friend of the open society. But the new dictators of Russia and China have used their power to impose restrictions on the free press and internet.

More specifically, it has been proposed that computers could re-establish a form of direct democracy, instead of parliamentary or representative democracy, by allowing all citizens to immediately cast their votes on all interesting questions by using their mobile phones and web connections. I find these suggestions about "tele-democracy" largely unrealistic: democracy presupposes open critical discussion by well-informed participants, and this condition can be satisfied by all citizens (instead of elected representatives or MPs) only in relatively small communities.

On the other hand, Adolf Hitler's *Mein Kampf* in 1925-26 already realized that the media can be used for the purposes of totalitarian political propaganda. Thus, the media may become an enemy of the open society even in a civilized society. Still, the best "vaccination" against such ruthless propaganda is education. In the information society it is a crucial political question who owns and controls the media. In one extreme, the democratic control of the political leaders will be weak or even prevented, if the press is dominated by the existing power elite. This has been the case in Russia, where President Putin has gained a lot of centralized power, but also in another form in Italy, where the former Prime Minister Berlusconi is the owner of the largest TV channels. In a healthy multiparty democracy, which wishes to avoid "infocracy", there should be both private independent media channels and public service companies functioning upon parliamentary trust.

The fragmentary postmodern society, where the subjects are losing their critical abilities and autonomy in the flow of meaningless information, also suggests that the power of media may turn out to be a new kind of threat of the open society. The free market alone is not a cure for this syndrome. Warnings in this direction were already given in the criticism of mass media and "culture industry" by the Frankfurt School, in particular by Herbert Marcuse in *One-dimensional Society* (1964). The popular success of dull and passive TV programs (including the misnamed "reality tv") and the endless

labyrinth or mycelium of the internet has even intensified the problem that the liberal policy of allowing anything in the media market that can be sold may make the citizens helpless floaters in the seamless semiotic processes.

The problems of information society have been intensified with the development of *social media* in many ways. In spite of their unprecedented ability to connect people,[14] big social media companies like Facebook and Twitter collect enormous amounts of data about individuals (e.g. their shopping and Google searches) and use secret algorithms to influence their future choices. They have provided platforms for progressive citizens' projects,[15] but also become forums of hate messages and fake news, where people group themselves in closed "bubbles". In this way, they have paved way for the post-truth era (cf. Chapter 16 above) and opened new possibilities for the external influence of political elections by trollies. Hubert L. Dreyfus (2001), a well-known critic of AI, argued that the syntactical hyperlinks in the internet do not enable us to distinguish meaningful information from meaningless. Jaron Lanier (2018), a pioneer in digital technologies and virtual reality, has presented "ten arguments for deleting your social media accounts right row". Among them are losses in truth, meaning, free will, happiness, empathy, self-respect, and politics. In brief, by their negative effects, social media are a threat to the preconditions of well-functioning democracy. The best armament against such drawbacks is media literacy and critical thinking fostered by education.

Democracy, Economy, and Globalization

The gradual progress of free trade and enterprise, promoted by liberalist economic theories during two centuries, has mostly walked hand in hand with political democratization. It is evident that these two processes have mutually supported each other in the United States and in many European

[14] *Connecting people* was the well-known slogan of the Nokia company in its heydays.

[15] Examples include the democratic groups during the Arab spring in 2011 and the school children joining Greta Thunberg's critique of climate warming.

countries. Still, the relation of economy and politics is open to several alternatives.

According to the traditional view, economy is an instrumentally valuable tool in sustaining life, making a living in a household, guaranteeing work and employment, and producing common goods, commodities and services for consumption. National economy is important for states, as it promotes the well-being of individual citizens and allows a state to secure its sovereignty and political power.

The famous declaration of neo-liberalist economy is Milton Friedman's (1970) slogan: "The business of business is business". In other words, "the social responsibility of business is to increase its profits", and to speak about its other "social ends" is "preaching pure and unadulterated socialism". This view implies that economy should be left to operate upon its inner logic: the primary and only goal of business is to make more money for the owners. Thereby economic success becomes an intrinsic value, an end in itself. The role of the state in national economy is minimal: to establish and guarantee conditions for free competition between corporations, where freedom is understood in the negative sense as the absence of external constraints. When the restrictions on international trade and monetary transactions were abolished in the 1980s, and the capital was allowed to restlessly seek the best profits that can be found in the world-wide market, the internal logic of the free trade gained a new impetus. At the same time, the new information and communication technology has made a global information economy possible. The results of this process of globalization have brought profits to some regions and some groups, but disasters to some others. The new political role of international actors like the World Trade Organization (WTO) is to guarantee the conditions of free competition on a global scale. By neo-liberalist lights, the regulations of the welfare society concerning conditions of work and taxation are seen as hinders to the free movements of the homeless capital.

This growth of the market economy has encouraged an ideological credo that wishes to subsume all aspects of human and social life under the principles of cost-effect-efficiency and accountability. Terms borrowed from economy are now vital in all fields - examples include entrepreneurial university,

culture industry, content production, and social capital. This libertarian ideology is surprisingly similar to what Karl Marx called "historical materialism": the sphere of economy, the means and relations of production, determine the course of history and society.

The dominance of economy over politics has created serious problems for democratic governance. In the 1960s left-wing critics suspected that parliaments have lost their power to rich families that own the factories - and proposed the abolition of private property as a remedy. Today the states have a very limited choice, as they have to adjust their policies of corporate taxation and employment and to cut their welfare services to please the interests of the owners of trans-national companies and foreign investment banks. New coalitions like the European Union still try to save something of the social projects that are now too large for nation states to handle, and attempts are made to strengthen the position of the United Nations and to establish new kinds of institutions to implement democratic principles in the world government. In the UN Development Program there is still hope about "globalization with a human face".[16] But the loud critics of WTO negotiations see globalization as process where the world is ruled by multinational corporations:[17] more freedom for corporations means less freedom for nation states and less political power for their citizens. Besides active protests, low voting rates are also an indication of the feeling that ordinary citizens of formally democratic countries have lost their opportunities to influence political decisions.

An interesting criticism of laissez-faire capitalism has been presented by George Soros, a former student of Popper in the LSE and the founder of the Open Society Institute (OSI). Soros is himself a leading rich market capitalist, "the man who broke the bank of England" in 1992 and interfered and exploited the Asian financial crisis in 1997. The OSI and the Soros Fund are running philanthropic programs in Eastern Europe and Russia, Asia, South Africa, and Latin America to improve the civil society, human rights, education, media, and public health. In

---

[16] See Jolly et al. (1999).
[17] Korten (1995).

2004 Soros was personally active in a campaign against the re-election of George W. Bush.

Soros accepts a kind of double standard. In an interview in the State of the World Forum in 2000 he urged that "the businessmen need to separate their business interests from their interests as citizens", so that it is appropriate for people to pursue their profit motive in business. This sounds like Friedman's neo-liberalist slogan that we quoted above. But Soros also argues that the economic system does not satisfy the ideal theoretical assumptions of perfect knowledge and rationality, and therefore its stability can be preserved only by deliberative efforts of the state. This is urgent, as the breakdowns of economy may give rise to totalitarian regimes. In this sense, Soros (1997) argues, capitalism threatens to be a *new* enemy of the open society. Soros further opposes the libertarian assumption that the free market is always beneficial to society. Against such excessive individualism and "market fundamentalism" he advocates the idea that, outside the domain of business, the open society should respect principles of justice and allow the greatest possible freedom to all of their citizens.

The position of Soros can be compared to what Popper originally stated in *The Open Society*. Popper claimed that "the unrestrained capitalism" discussed by Marx has "completely withered away in the twentieth century".[18] It has given way to political and economic "interventionism", which - besides the collectivist "Russian" and fascist forms - is exemplified by the New Deal in America and has reached its highest democratic level so far in Sweden (*ibid*., 140). This is not only a historical remark, but Popper urges that "we must demand that unrestrained *capitalism* give way to an *economic interventionism*". The reason is what he calls "the paradox of freedom": freedom defeats itself, if it is unlimited, since "unlimited freedom means that a strong man is free to bully one who is weak and to rob him of his freedom", and therefore all citizens have the right to be protected by the state (*ibid*., 124-125).

Some authors have proposed a conceptual distinction between capitalism and the market economy. *Capitalism* is the economic system that allows private ownership and fosters

[18] See Popper (1966), vol. 2, 335.

ruthless egoistic competition for one's own advantage. It sounds like a revival of the 19th century (ill-named) "social Darwinism": society is seen as a jungle of egoistic beasts playing a zero-sum game of the survival of the fittest. The so-called casino economy, where corporations act like beasts, is uncontrolled capitalism in this sense.[19] It is this form of unrestrained capitalism that Popper clearly rejected in *The Open Society*.[20] The *market economy*, on the other hand, is based on free economic competition where some internal legal and ethical principles govern the fairness of competition: you are not allowed to cheat your partners and competitors, you should not take advantage of internal information for your personal benefit, etc. It is this minimal protective function that the neo-liberals allow for the state and for the entrepreneur's ethical commitment - and to this extent they are not advocating the crudest form of capitalism. Free economic competition should mean freedom from manipulation (e.g. by big international companies).

But, in spite of the wish of some Popperians to read their master in a libertarian way, it is also clear that Popper's "economic interventionism" allows the state to have a more significant role than the neo-liberalist "night guard". In the open society, for Popper and Soros, fair economic competition and co-operation should be viewed as means to common good and social justice - without assuming with Adam Smith that this important function is as if automatically realized by "the invisible hand".

To become an open society, a market economy has still to be enriched by further principles of fairness. The attempt to construct a fair market economy comes close to the communitarian ethics that seeks, against the individualist emphasis on struggle and competition, the advantages in mutual co-

---

[19] Capitalism based on pure egoism, under the strange cult name "objectivism", is cherished in the novel *The Fountainhead* (1957) by the Russian-born American writer Ayn Rand.

[20] For the tension between Popper's methodological individualism and his holistic account of the reality of human-made social constructions in the world 3, see Niiniluoto (2006).

operation and respect between citizens and nations. Economic agents enter voluntarily into competitive actions and take risks, but they also rely on the infrastructure provided by the society. Following the Rawlsian conception of justice, such economic activities should give advantages to all: by democratically designed means they should redistribute parts of the gains to all members in the society, so that they are defined as co-operative non-zero-sum games - and the same idea should be repeated on the international global level.

In addition to legislation and democratic procedures, ethical and social principles are also needed. As the failure of the Russian post-Soviet capitalism shows, the lack of the "social capital" and an underdeveloped civil sector can be a serious defect. The success of economy has to be based upon the mutual trust between the regime, the economic agents, and the citizens. *Pace* Friedman, the fair market economy in this sense should be based upon principles concerning the balanced ethical responsibilities of business enterprises with respect to their stakeholders, employers, consumers, and the natural and social environment. It acknowledges the fact that firms normally give profit to their owners, but regards the neo-liberalist overemphasis of the shareholder value as a one-sided and unhealthy feature of the economic system.

Sometimes the market economy is defended by claiming that its manner of operation in itself is "democratic". There is some truth in this thesis, if the citizens as customers are able to influence the behaviour of firms and companies. The enlightened customers "vote" by buying products of companies which behave morally with respect to environment and employees. For this reason, ethics and social responsibility have become urgent matters for corporations who wish to be successful. But the reality today is still that "money talks" and the shareholder values are stronger than the power of the customers. In order to be moral agents in a genuine sense, sheer utilitarian profit cannot be the ultimate motive, but corporations should adopt their ethical values on the conviction that we all are together seeking a world with a better quality of life for all.

Populism as a Challenge of Liberal Democracy

Populism has emerged as serious phenomenon in the political life of many Western countries. Political scientists define this concept in many ways, usually emphasizing the contrast between "the elite" and "the people". The rapid development of modernization and globalization has often led to circumstances where some groups feel that their concerns are disregarded by rich and powerful elites. This may make them vulnerable to emotional attachment to charismatic leaders and prone to simple answers to complex questions.

In addition to some left-wing populists (e.g. Syriza in Greece), most populist parties are right-wing conservatives (e.g. Le Pen in France, Fidesz in Hungary, Law and Justice in Poland, Lega Nord in Italy, Sweden Democrats, True Finns). They pose a challenge to liberal democracy, since their values are markedly non-liberal: in favor of law and order, nationalism, and traditional religions, in opposition to modern lifestyles, rights of ethnic minorities, the European Union, and permissive immigration policies.[21]

It is important that the voice of all people is heard in the political system, and there should be measures which improve equity among citizens. This is the ideal of democratic welfare states. So as long as a populist group functions as a party, which abides to normal political rules and practices, there no reason to complain about its activities – even though we liberals disagree with their values.

But sometimes populists, who gain power in society by elections, became threats to democracy. In Poland the Law and Justice party has acted against the rule of law by weakening the independence of the judicial system. The same has happened in Hungary, where Viktor Orbán's Fidesz started a conspiracy theory against George Soros, and its political campaign forced the Central European University (founded by Soros in 1991) to escape from Budapest to Vienna in 2018.

Even more dramatic violations of democratic principles were accomplished during Donald Trump's presidency in the

---

[21] Anne Applebaum (2020) gives a penetrating analysis of the lure of right-wing populism.

USA in 2016-20. As a ruthless businessman and tv-star, Trump is a typical populist leader, who has been able to convince his zealous supporters that his political power really means power of the people. His campaign was based on unwarranted self-praise and lies about the rival, and during his reign he sent about 15 false or misleading twitter-messages per day – until this account was closed. Trump is thus a central figure in the post-truth era (cf. Chapter 16 above). Trump refused to admit his defeat against Joe Biden, and the attack of his supporters to the Capitol Hill on January 6, 2021, was a disgrace of the American tradition of democracy. The Democracy Index has classified the USA since 2016 as "a flawed democracy" (with e.g. Hungary and Poland).

Historicism, 9/11, and Other Violent Attacks

Francis Fukuyama proclaimed in 1989 that history has come to a happy end with the permanent victory of liberal democracy (Fukuyama, 1992). By Popper's standards, Fukuyama's thesis is unfalsifiable due to its ambiguity: it may concern the actual historical position of political systems in the "material world" or the theory of such systems in the "ideal world".[22] To the horror of the Popperians, Fukuyama also revitalized Hegelian historicism by his belief that the world is destined to come to an end - even in a finite time.

A similar historicist view among the Marxists, with the belief that the proletariat with its developing class consciousness will actually reach the final stage of "self-identity", was defended by Georg Lukács in 1923.[23] He is mentioned neither in *The Open Society* nor by Fukuyama. But Fukuyama can be characterized as the Lukács of the bourgeoisie.

---

[22] Fukuyama's claim was not initially very plausible, and the problems in establishing democratic regimes in many countries count against his optimist thesis, if it is understood as a historical claim about the actual victory of democracy among nations. The problems in the theory of democracy (see note 6), and the need to develop its new forms for international politics, disprove Fukuyama's thesis, if it is interpreted as the claim that our ideal of liberal democracy has reached its final form.

[23] See Lukács (1971).

The dramatic events in the world after the end of the Cold War – from a crisis to another crisis - cannot be understood as mere side-tracks and disturbances within the victorious march of liberal democracy. Ulrich Beck's (1992) "risk society" and Samuel Huntington's (1996) "clash of civilizations" seem to characterize the new world order after the tragic terrorist attack of 9/11 and later violent attacks.

After the collapse of the World Trade Towers in New York by the aircraft hijacking attack of the islamist Al-Qaeda in 2001, President George W. Bush's declaration of a war against terrorism and "the axis of evil" led to the bombing of the rocks of Afghanistan and the invasion of Iraq. Even Bush himself suspected that this new kind of war cannot ever be brought to an end. It took ten years before Osama bin Laden was killed in a hit by the CIA, and the Guantanamo Bay prison is still operating after twenty years. In this sense, terrorism threatens to create an atmosphere where the leading democratic power in the world violates its own ideals about freedom and human rights.[24]

Another act of violence, this time by the order of President Vladimir Putin of Russia, was conducted by the brutal attack of the Russian army against the sovereign state of Ukraine on February 24, 2022. With the background as a KGB officer, Putin is certainly well aware of the 1832 principle of Carl von Clausewitz that "war is a mere continuation of policy by other means". This is what Anatol Rapoport calls the "political philosophy of war": war is a rational instrument of national policy, with the aim of "compelling our enemy to fulfil our will".[25] Putin's wars in Georgia, Chechenia, Syria, and his occupation of Crimea prepared the invasion of the Ukraine. Some commentators believe that Putin is wholly pragmatic and tactically chooses the tone and content of his speech – including lies and threats - relative to different audiences at home and abroad. But I venture to suggest an alternative hypothesis: it may be a mistake to think that Putin's obsession with Ukraine could be

---

[24] The NATO bombing of Belgrade in 1999 was against international law, but it helped to stop the ethnic conflicts and genocide among the countries of Yugoslavia.

[25] See the introduction by Anatol Rapoport to Clausewitz (1968).

understood in terms of instrumental rationality (cf. Chapter 8 above) or balances of costs and benefits. In appealing to the historical idea of a Holy Russia he ignores that Russia received its orthodox religion via Ukraine, and Kyiv is older than Moscow.[26] Putin's illegal invasion of Ukraine thus seems to exhibit what Rapoport calls the "eschatological philosophy of war", which aims to unfold some grand design – in his case, the Pan-Slavic revival of the old imperial Russia and the Soviet Union before its collapse in 1991.[27]

In the same way, the terrorists or their leaders are not typically acting on what we would regard as instrumentally rational goal-directed strategies. It is more plausible to see them acting on what Weber called "value rationality", even though their values are not reasonable: they are willing to sacrifice themselves for a higher purpose which is defined by their religious attachment to a historical destiny. In Popper's terms, their approach is based upon historicism: their belief in the objective value of their actions is justified by a belief in the predetermined end of the history. In the 1961 Addenda to *The Open Society*, Popper warned against the dangerous assumption that someone knows objective values and standards: such "moral historicism" leads to "the identification of standards either with established might or with future might".[28]

---

[26] One of Putin's advisers about Ukraine has been the self-learned philosopher Alexandr Dugin, whose openly fascist politics and ultranationalist "neo-Eurasianist" ideology has been inspired by Heidegger and Hitler.

[27] Rapoport gives a fierce critique of the American Neo-Clausewitzians (Raymond Aron, Herman Kahn), who treated wars as a normal practice of nation states. But in my view both the political and eschatological conceptions of war are dangerous. The former has the advantage that it is more predictable (the "balance of terror" prevailed during the cold war), while the latter may be accidental and irrational. The best philosophy is what Rapoport calls "cataclysmal": wars are catastrophes to be avoided by diplomacy, peace research, and pacifist policies.

[28] See Popper (1966), vol. 2, 393. After G. W. Bush's re-election, it was argued by David Klinghoffer, a columnist in *Jewish Forward*, that Bush was supported by a majority of Americans, Christians and Jews, who believe in objective values defined by God's will. Donald Trump, in spite of known moral aberrations, was likewise strongly supported by the Christian conservatives – and some his followers, blinded by conspiracy theories like

As a peace project after World War II, the European Union adopted from Immanuel Kant's *Zum ewigen Frieden* (1795) the idea that the international law of nations should be based on a federation of free nations which are republican by their civil constitution.[29] The ethical demand of "perpetual peace" is sustained by the fact that democratic nations which exchange trade and ideas between them have no reasons for wars. But even a pacifist has to acknowledge the legitimacy of defence wars: to protect liberal democracy against authoritarian opponents who rely on moral historicism, it is unfortunate that swords and bombs are sometimes needed for this task. This was the only way of stopping Adolf Hitler's eschatology with the Master Race. But terrorism and attack wars should not be combatted by bringing in another kind of historicism - belief in the historical destiny of Western democracy. The result is a clash between an established might and a projected future might. The alternative is to promote the open society by education, rational argumentation, freedom of thought, recognition and self-respect, economic wealth, and social justice.

Note. This chapter is an updated version of a lecture in the The Tampere Club, September 20, 2004. It was published as "The Open Society and its New Enemies: Critical Reflections on Democracy and Market Society", in Geoffrey Brennan (ed.), *Preconditions of Democracy*, The Tampere Club Series vol 2. Tampere: Tampere University Press, 2006, 171-186.

Literature

Applebaum, Anne (2020). *Twilight of Democracy: The Seductive Lure of Authoritarianism*. New York: Doubleday.

Avineri, Shlomo (1972). *Hegel's Theory of the Modern State*. Cambridge: Cambridge University Press.

Beck, Ulrich (1992). *Risk Society: Towards a New Modernity*. London: SAGE.

Qanon, even hailed him as a new Messiah. The head of the Russian Orthodox church, patriarch Kirill, regards Putin as "God's miracle" and supports the Russian war against "the Pride parades" of Ukraine.

[29] See Kant (2003).

Castells, Manuel (2000). *The Information Age: Economy, Society and Culture*. Oxford: Blackwell.

Clausewitz, Carl von (1968). *On War*. Harmondsworth: Penguin.

Dreyfus, Hubert L. (2001). *On the Internet*. London: Routledge.

Friedman, Milton (1970). "The Social Responsibility of Business is to Increase its Profits", *The New York Times Magazine*, September 13, 1970. (Reprinted in M. D. Bayles & K. Henley (eds.), *Right Conduct: Theories and Applications*, 2nd ed., New York: Random House, 1989, 316-322.)

Fukuyama, Francis (1992). *The End of History and the Last Man*. London: Penguin.

Giddens, Anthony (2000). *The Third Way and Its Critics.* Cambridge: Polity Press.

Hacohen, Malachi Haim (2000). *Karl Popper: The Formative Years, 1902-1945.* Cambridge: Cambridge University Press.

Hayek, Friedrich von (1944). *The Road to Serfdom*. Chicago: The University of Chicago Press.

Held, David et al. (1999). *Global Transformations, Politics, Economics and Culture*. Cambridge: Polity Press.

Huntington, Samuel (1996). *The Clash of Civilizations and the Remaking of World Order*. New York: Simon & Schuster.

Ishiguro, Hidé (2003). "The Future of Democracy". In *The Future of Democracy: Essays of The Tampere Club*. Tampere: Aamulehti, 146-156.

Jolly, R. et al. (1999). *Human Development Report 1999.* Oxford: Oxford University Press.

Kant, Immanuel (2003). *Perpetual Peace: A Philosophical Sketch*. London: Hackett. (Original in German 1795.)

Korten, David C. (1995). *When Corporations Rule the World*. West Hartford, CT: Kumarian Press.

Lanier, Jaron (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. New York: Henry Hult and Co.

Li Tieying (2002). *On Democracy*. Beijing: Social Sciences Documentation House.

Lukács, Georg (1971). *History and Class Consciousness*. London: Merlin Press.

Magee, Brian (1973). *Popper*. London: Fontana.

Marcuse, Herbert (1964). *One-Dimensional Society: Studies in the Ideology of Advanced Industrial Society*. Boston: Beacon Press.

Niiniluoto, Ilkka (2006). "World 3: A Critical Defence". In Ian Jarvie, Karl Milford & David Miller (eds.), *Karl Popper: A Centenary Assessment vol II*, Aldershot: Ashgate, 59-69.

Popper, Karl R. (1945). *The Open Society and Its Enemies*, vols. 1-2. London: Routledge and Kegan Paul. (The fifth revised edition, 1966.)
Popper, Karl R. (1957). *The Poverty of Historicism*. London: Routledge.
Popper, Karl R. (1963). *Conjectures and Refutations*. London: Routledge and Kegan Paul.
Skinner, Quentin (2002). "A Third Concept of Liberty", *Proceedings of the British Academy* 117: *Lectures 2001*, 237-268.
Soros, George (1997). "The Capitalist Threat", *The Atlantic Monthly* 279, 45-58.

Index of Names