

ETHICS, RISKS AND DIFFICULT DECISION- MAKING

EXAMPLES AND EPISTEMIC FOUNDATIONS

*Göran Hermerén, Johannes Persson, Eva Sjöstrand
& Robin Stenwall (eds)*

KVHAA Handlingar
Filologisk-filosofiska serien 28

Ethics, Risks and Difficult Decision-making

Examples and Epistemic Foundations

*Göran Hermerén, Johannes Persson,
Eva Sjöstrand & Robin Stenwall (eds)*



KUNGL. VITTERHETS HISTORIE OCH ANTIKVITETS AKADEMIEN
HANDLINGAR FILOLOGISK-FILOSOFISKA SERIEN 28

Göran Hermerén, Johannes Persson, Eva Sjöstrand & Robin Stenwall eds. *Ethics, Risks and Difficult Decision-making: Examples and Epistemic Foundations*. Kungl. Vitterhets Historie och Antikvitets Akademien (KVHAA), Handlingar, Filologisk-filosofiska serien 28. Stockholm 2026. 148 pp.

Abstract

This book begins with case studies of difficult decision-making in areas such as nuclear strategy, the climate crisis, the Flint water crisis, care of extremely preterm-born children, solar radiation management and local knowledge. The difficulties involved are due to uncertainties, knowledge gaps, unstable values and controversies over the ranking orders of the values at hand. The importance of how decision problems are presented and communicated is demonstrated. These case studies constitute the first theme of this book.

The second theme explores philosophical ideas raised by questions about the relations between knowledge, truth, belief and probability. Should decisions be based on knowledge, defined—as Plato proposed—as true and justified beliefs? An important point of departure for expanding this discussion is Frank Ramsey's essay 'Truth and probability' and its central idea that the rational decision-maker maximizes (subjective) expected utility. Can this idea be applied to the examples of difficult decision-making presented in the case studies? If not, why is that, and what are the alternatives?

Keywords

belief, decision-making, expected utility, knowledge, Plato, probability, Ramsey, truth

Kungl. Vitterhets Historie och Antikvitets Akademien
(KVHAA, The Royal Swedish Academy of Letters,
History and Antiquities), Stockholm
vitterhetsakademien.se

© 2026 The authors and KVHAA

ISBN 978-91-88763-78-5

ISBN (pdf) 978-91-88763-79-2

ISSN 0083-677X

Available with open access, <https://doi.org/10.62077/3qhje>

Cover design and typography by Johan Laserna

Printed by Bulls Graphics, Halmstad 2026



Contents

Preface	7
Introduction	9
GÖRAN HERMERÉN	
Rethinking rationality in nuclear strategy: Psychological biases and policy implications	13
PAUL SLOVIC	
Addressing the climate crisis: The power of language and framing	19
WÄNDI BRUINE DE BRUIN	
The precautionary principle has no exceptions	29
MARTIN PETERSON	
Care of extremely preterm-born children: Ethics, risks and difficult decisions	45
GÖRAN HERMERÉN	
Local knowledge, proven experience and scientific evidence	65
JOHANNES PERSSON	
What's action got to do with it?	79
MONIKA GRUBER	

On objections to Ramsey on belief	91
KEVIN MULLIGAN	
Can there be probabilistic knowledge?	109
PASCAL ENGEL	
A hundred years of rational decision-making — and then what?	127
NILS-ERIC SAHLIN	
The contributors	141
Index	145

Preface

This publication is based on presentations at a symposium on ethics, risks and decision-making, held on 17–18 October 2024 at the Royal Swedish Academy of Letters, History and Antiquities in Stockholm.

The papers focus on two themes, the first being difficult decision-making, with examples from several areas, and the second being philosophical issues raised by making (rational) decisions, with special emphasis on epistemic justifications of such decisions. These themes are related in ways that will be presented and explained in the introduction.

The authors are internationally renowned researchers, all of whom have collaborated with Nils-Eric Sahlin, whose contributions to research we want to honour. They all accepted with great enthusiasm the invitation to speak at the symposium and contribute with papers. Short biographies of the authors are presented at the end of this book.

The Royal Swedish Academy of Letters, History and Antiquities in Stockholm funded the symposium and, together with the Royal Society of Letters at Lund, the editing and printing of this volume. We are grateful for their generous support.

The editors

Introduction

GÖRAN HERMERÉN

The purpose of this introduction is to present the two themes discussed in this book, their relations to each other and to the book's title, as well as to explain the reasons for the symposium and this publication. Decision-making raises both empirical and normative issues. How are decisions made? On what grounds? How ought decisions to be made? What are the difficulties? How can or should they be avoided or minimized?

Theme 1:

Examples of difficult decision-making

We make decisions every day, sometimes based on rules of thumb and how the situations at hand are usually dealt with, at other times on the basis of scientific evidence. Some of these decisions concern simple issues, others can lead to radical changes in our lives, and several can even be a threat to the existence of humanity. Some decisions can concern a choice between small risks for a major disaster and big risks for a minor disaster.

“Maximize expected utility” is a well-known slogan in decision theory and a key to rational decision-making, applied and developed by Frank Ramsey—and, in this volume, discussed by Nils-Eric Sahlin. But there are obstacles to rational decision-making which need to be understood. If we want to be able to predict preferences, we need to understand the importance of how the problems are presented and how our minds work, in particular the psychological principles governing perception of decision problems. A single child in need will generate and receive a lot of

compassion but the compassion will not increase proportionately when the number of children in need increases.

In a multicultural society there are frequently different and diverging views about what we know, the scientific evidence, as well as concerning what practical experience is relevant in a particular situation. Moreover, there is not always agreement about basic values and how short- and long-term interests of those involved are to be balanced against each other.

There is no shortage of challenges at present. They include the climate crisis, pandemics such as Covid-19, fake news and economic crises, the use of artificial intelligence in warfare and to manipulate political processes, as well as the current situation in the Middle East. Attempts to deal with such issues illustrate the interplay between scientific and other evidence, practical experience and values in decision-making.

The relations between scientific evidence and practical experience in decision-making need to be clarified. This also holds for the relations of these concepts to ethics and to different kinds of values. Values that are generally accepted in a culture tend to be invisible for people in that culture; it is easy to overlook assumptions that are commonly taken for granted. Ideas that are generally accepted tend to be self-evident points of departure for making decisions. What is morally problematic for some is, as we know, not always problematic for others.

Thus, we have good reasons not only to distinguish between different traditions in ethics, such as Kantian, utilitarian or human rights-based approaches. Our decisions and choices are influenced by the alternatives we see, how they are described, and how we suppose that their consequences will affect us, our children and grandchildren—and their children and grandchildren. What are the risks? What are the possibilities? What risks are acceptable, and on what grounds are such decisions taken?

In the discussions of such issues, we need to distinguish between known probabilities for a certain outcome and risks due to knowledge gaps and uncertainties; the latter are often referred to as epistemic risks. The distinction between these kinds of risks, and particularly between epistemic risks and others, has been explored, analysed and applied in different contexts by Nils-Eric Sahlin.

The case studies, in the contributions by Paul Slovic, Wändi Bruine de Bruin, Martin Peterson, Göran Hermerén, Johannes Persson and Nils-Eric Sahlin, present examples of difficult decision-making in a variety of

areas, such as (respectively) rationality in nuclear strategy, the climate crisis, the Flint water crisis, care of extremely preterm-born children, local knowledge and solar radiation management.

There are several explanations of the difficulties for (rational) decision-making suggested by the case studies presented in the first five papers. They include uncertainties of several kinds, knowledge gaps, unstable values and value conflicts, including controversies over the ranking orders of values.

Theme 2:

Epistemic justifications of difficult decisions

The second theme, presented in the contributions by Monika Gruber, Kevin Mulligan and Pascal Engel, as well as by Nils-Eric Sahlin, explores some fundamental philosophical ideas raised by questions such as: Should decisions be based on knowledge, defined—as already proposed by Plato in the dialogue *Theaetetus*—as true and justified beliefs?

This raises fundamental questions concerning how to understand beliefs and the relations between beliefs, knowledge and truth, as well as their relations to actions. Other questions are also suggested and discussed, including: What are the objects of beliefs? How are degrees of belief to be determined? Can there be probabilistic knowledge? Can local knowledge be analysed in terms of justified true beliefs?

Moreover, what role could or should various ideas about rational decision-making or different interpretations of the precautionary principle play in this context? What are Ramsey's views on these matters, and what can we learn from them? What are the objections to Ramsey's views, and how can they be dealt with?

A particularly important point of departure for the analysis of decision-making and the problems discussed in the second theme of this book is Ramsey's famous essay 'Truth and probability', written in 1926 and published in 1931.¹ This essay is generally regarded as the first complete theory of rational decision-making. It provides an answer to the question of how we should make decisions. The central idea is that the rational decision-maker maximizes (subjective) expected utility.

Can this theory be applied to the examples of difficult decision-making presented in this book? If the answer is negative, as suggested in some of

INTRODUCTION

the contributions, why is that, and what are the alternatives? If this book can stimulate discussions of these complex and difficult issues, it has served an important function.

Note

1. The essay was first published in Ramsey 1931 and republished in Ramsey 1990.

References

- Ramsey, Frank P. 1931. *The Foundations of Mathematics and Other Logical Essays*, ed. Richard Bevan Braithwaite, London: Kegan Paul, pp. 156–198.
- Ramsey, Frank P. 1990. *F. P. Ramsey: Philosophical Papers*, ed. David Hugh Mellor, Cambridge: Cambridge University Press, pp. 52–109.

Rethinking rationality in nuclear strategy

Psychological biases and policy implications

PAUL SLOVIC

Here we stand in the middle of this new world with our primitive brain, attuned to the simple cave life, with terrific forces at our disposal, which we are clever enough to release, but whose consequences we cannot comprehend.

Albert Szent-Gyorgyi

Nils-Eric Sahlin has always thought deeply about rationality and ethics in the context of difficult decisions in medicine and other important domains pertaining to valuing lives and protecting them. In keeping with his contributions to ethical dilemmas in lifesaving, I shall briefly describe recent psychological studies pertaining to one of the most difficult, ethically challenging, and consequential decisions any human being would ever face: Whether to order the use of nuclear weapons to vanquish an enemy.

The question of nuclear weapons management remains one of the most critical issues of the 21st century. While nuclear deterrence has seemingly prevented large-scale wars since 1945, recent advancements in nuclear technology, increased global proliferation, and shifting political dynamics pose significant risks. Despite six decades of research into decision-making, nuclear policy remains largely insulated from the many important findings from this research. Research seeks to bridge that gap by examining the psychological mechanisms that influence nuclear decision-making, the cognitive biases that may lead to miscalculations, and the policy interventions required to mitigate existential risks.

The assumption of rationality in nuclear deterrence

Traditional nuclear deterrence theory is built on the assumption that state actors are rational decision-makers who, when provided with accurate information, will make choices that maximize national security. This perspective holds that nuclear weapons serve primarily as tools for war prevention rather than warfighting. As Admiral Richard Mies, former head of the United States Strategic Command, has asserted, nuclear weapons act as a deterrent against major wars and ensure strategic stability. However, research in psychology challenges this notion by demonstrating that decision-making is often subject to cognitive and emotional limitations that significantly alter and destabilize decisions.

Deterrence assumes that all actors involved will act predictably and in their best interest. But history has shown that world leaders are not immune to misjudgments. The Cuban Missile Crisis of 1962 exemplifies how close the world came to nuclear annihilation due to miscalculations and misunderstandings between the United States and the Soviet Union. Modern nuclear strategy continues to rest on the assumption that rational leaders will never deliberately initiate a nuclear war, yet this assumption disregards the emotional and psychological pressures that can drive decision-making in times of crisis.

Psychological challenges to rationality in nuclear decision-making

Decision-making operates through two primary cognitive systems: System 1 (experiential) and System 2 (analytical). System 1 relies on intuition, emotions, and narrative reasoning, while System 2 engages in deliberative, logical processing. Although both systems play a role in decision-making, high-stakes military conflicts, forcing decisions to be made in novel situations under emotional stress and severe time pressure, are likely to favor System 1, leading to irrational choices driven by affective responses rather than careful, reasoned analysis.

One such affective response is *psychic numbing*, wherein individuals struggle to emotionally appreciate large-scale human suffering. Research has demonstrated that as the number of victims in a crisis increases, emotional engagement decreases, leading to diminished motivation to prevent

mass casualties.¹ Studies have found that people tend to respond emotionally to singular, identifiable victims, yet struggle to grasp the gravity of mass casualties.² A large body of research can be summarized by the troubling conclusion that “the more who die, the less we care”.³ This cognitive shortcoming has profound implications for nuclear decision-making. It means that leaders making nuclear-related decisions will certainly fail to appreciate the full scale of destruction that can be caused by a nuclear strike.

Another cognitive bias relevant to nuclear policy is the *prominence effect*, which describes the tendency to prioritize a single dominant factor while neglecting other critical considerations. In nuclear strategy, this often manifests as an overriding emphasis on national security and military advantage, while ethical considerations regarding civilian casualties and environmental devastation receive little or no attention. This can lead to disproportionate military responses, including the use of nuclear weapons against civilian populations.

A notable example is the 1945 decision to drop atomic bombs on Hiroshima and Nagasaki. While military leaders justified the bombings as a means to end World War II swiftly and minimize further troop losses on both sides, ethical considerations about civilian casualties were largely overlooked. The prominence effect explains why leaders may prioritize immediate security concerns over long-term humanitarian and civilizational impacts.

Findings from online survey experiments

Our understanding of the ways cognitive limitations and emotional biases may affect decisions about the use of nuclear weapons has been greatly informed by online survey experiments originally designed by political scientists Scott Sagan and Ben Valentino and replicated and extended by my colleagues and me.⁴ One recent study, involving nearly 3,000 American participants, presented respondents with a grim choice: Continue a difficult but winnable ground war with escalating American troop casualties or launch a nuclear strike to end the war quickly and save the lives of the troops but kill up to 2 million enemy civilians. The findings were as startling as they were unsettling. Even when faced with the prospect of losing as few as 100 American lives—a number that seems trivial against

the backdrop of national security—over 20% of respondents opted for the nuclear strike. As the projected casualties increased to 40,000, support for the nuclear option surged to 50%.

Support for nuclear strikes correlated strongly with conservative political ideologies, pro-death penalty attitudes, and restrictive immigration and abortion policies. In other words, we found that individuals with punitive worldviews were far more likely to endorse extreme military actions, including nuclear strikes.

Survey participants also demonstrated *scope insensitivity*, a form of *psychic numbing*, when they exhibited similar support for nuclear strikes regardless of whether the projected death toll was 100,000 or 2 million civilians.

Another disturbing finding was that adding a second nuclear strike option to the choice set, one much better than the other, greatly increased the percentage of respondents who chose to bomb civilians rather than continue the ground war. For example, when presented with two nuclear strike options—one causing 100,000 deaths and another causing 2 million, along with the option of continuing the ground war—many respondents, even those who had preferred the ground war at an earlier occasion, now chose the 100,000-death bomb, calling it the “better option”. Their explanations for their preference reversal shows they deceive themselves into feeling virtuous by saving the troops and expecting to kill 100,000 civilians with a better bomb, rather than rejecting nuclear use altogether. We have named this “the curse of betterness”, a phenomenon where decision-makers rationalize the lesser of two evils as an acceptable course of action. It is related to a deception well known to marketers and consumer psychologists, who have named it the *decoy effect*. It is used to make one’s product sell better against a competitor’s product by creating another product clearly inferior to one’s favored product, thus enhancing the attractiveness of the favored product against the competition.

The implications of these findings are both profound and disconcerting. In a nuclear age marked by the constant threat of miscalculation, it is chilling to consider that the decisions made by our leaders—decisions that are supposed to be guided by sober, strategic calculation—might instead be swayed by the same cognitive biases that lead us to overvalue a better bargain at the supermarket. When the stakes involve not just financial loss but the potential extinction of millions, such biases

are not merely inconvenient quirks of human psychology: they are existential risks.

Current nuclear deterrence strategy rests on the assumption of rational decision-making, a presumption that this research calls into question. If top military and political leaders are susceptible to the influence of affect, security prominence, comparison effects, and psychic numbing, then the stability of deterrence could be more precarious than we imagine. The possibility that a punitive leader might one day be swayed by the “better bomb” effect—a cognitive bias that makes a catastrophic but less devastating strike appear more acceptable—should send shockwaves through policy circles and the public alike.

In confronting these unsettling realities, we are compelled to rethink how we manage nuclear arsenals and design the decision-making processes that govern their use. The research suggests an urgent need for institutional safeguards that account for the inherent limitations of human cognition. Training programs, decision support systems, rigorous checks and balances, attention to the character of the president and our top military advisors, and renewed and enforceable arms-control treaties, must be reexamined in light of these psychological vulnerabilities.

As policymakers and citizens, we must confront the uncomfortable truth that our decision-making processes, however evolved, are ill-equipped to handle the moral and strategic complexities of nuclear war. We must strive for a new paradigm of nuclear policy—one that recognizes and mitigates the cognitive biases that could, in a moment of crisis, propel us toward irreversible harm to civilization.

Notes

1. Slovic 2007.
2. Västfjäll *et al.* 2014; see also Slovic 2007.
3. Slovic & Västfjäll 2015.
4. Slovic *et al.* 2020; Slovic *et al.* forthcoming (2026).

References

- Slovic, Paul 2007. ‘If I look at the mass, I will never act: Psychic numbing and genocide’, *Judgment and Decision Making* 2:2, pp. 79–95. <http://journal.sjdm.org/vol2.2.htm>

- Slovic, Paul & Daniel Västfjäll 2015. 'The more who die, the less we care: Psychic numbing and genocide', in Susanne Kaul & David Kim eds, *Imagining Human Rights*, Berlin: De Gruyter, pp. 55–68.
- Slovic, Paul, C. K. Mertz, David M. Markowitz, Andrew Quist & Daniel Västfjäll 2020. 'Virtuous violence from the war room to death row', *PNAS* 117:34, pp. 20474–20482. <https://doi.org/10.1073/pnas.2001583117>
- Slovic, Paul, Melissa Peterson, Rose McDermott, Daniel R. Post & Daniel Västfjäll, forthcoming (2026). 'When our minds go nuclear: Rethinking nuclear strategy and policy through the psychology of risk and decision making', *Texas National Security Review*.
- Västfjäll, Daniel, Paul Slovic, Marcus Mayorga & Ellen Peters 2014. 'Compassion fade: Affect and charity are greatest for a single child in need', *PLoS One* 9:6, e100115. <https://doi.org/10.1371/journal.pone.0100115>

Addressing the climate crisis

The power of language and framing

WÄNDI BRUINE DE BRUIN

Introduction

Around the world, most people are now worried about climate change. The 2019 Lloyd's Register Foundation World Risk Poll found that 69% of people worldwide were “somewhat worried” or “very worried”.¹ These numbers have remained relatively stable over the years.² The worldwide recognition of the threat of climate change signals a shift from 2007–2008, when there was still worldwide apathy or lack of awareness about climate change.³

The Intergovernmental Panel on Climate Change (IPCC) has long been warning that climate change will bring more extreme weather to countries around the world.⁴ Worldwide, the weather has been getting hotter, and there have been more heat waves, droughts, and wildfires. The IPCC has released reports every 7 years. These reports reflect insights from climate science for policymakers, practitioners, and members of the public. Among other things, the reports discuss two types of behavior change that are needed in the face of climate change: “mitigation” and “adaptation”. “Mitigation” means stopping climate change from getting worse by reducing the amount of greenhouse gases we put into the atmosphere. “Adaptation” means protecting ourselves against the climate impacts that are already happening.

IPCC reports have been criticized for using jargon and university-level language.⁵ As a result, communications about climate can be hard to understand for people who do not have a degree in climate science. Climate jargon can be confusing, off-putting, and lead people to go elsewhere for their information.⁶ Everyday language is even preferred by

people who read well.⁷ Many adults in the United States, Canada, United Kingdom, and Europe read at the level of 12- or 13-year-olds.⁸ Readability studies have found that text meant for the public is better understood if it is written for 12–13 year-olds.⁹ Sentences of 16–20 words and words with no more than 2 syllables are easier to read.¹⁰

Of course, there are always exceptions. Some short words can be confusing, and some long words can be easy to understand. To test how well the wording works, it is therefore important to conduct interviews with members of the intended audience.¹¹ In these interviews, they should be presented with words and phrases. They should also be asked to interpret the words and phrases, and to give suggestions for improvement.

Do people understand climate terminology?

Out of concerns that climate communications may be too hard for people to read, we teamed up with the United Nations Foundation to test how well climate change terminology is understood among the US general public.¹² In our study, we asked people from across the US to interpret words and phrases about climate change. The terms and phrases were selected by climate scientists who served as authors of IPCC reports. The terms included “mitigation” and “adaptation”, which I introduced at the beginning of my chapter.

Overall, we found that people were confused by climate change terminology. Even those who were concerned about climate change often did not know how to interpret climate jargon. One person we interviewed said: “it sounds like you’re talking over people”.¹³

When we asked interviewees to define “mitigation”, we found that it was thought of as a difficult term. When asked to define it, we heard definitions such as “resolving a conflict”. Thus, mitigation was confused with mediation. Interviewees also had trouble coming up with definitions that linked the term “mitigation” to climate change.¹⁴ Regarding “adaptation”, we found that people thought of it as an easy term. But when asked to define it, many said it meant “turning a book into a movie”, reflecting its use in everyday language, where it typically refers to the movie version of a book. However, that is not what “adaptation” means in the context of climate change.¹⁵

We also asked interviewees who did understand the terms to provide

descriptions in everyday language. We found that it is possible to talk about climate terms in a simple way. It was suggested that “mitigation” can be described as stopping climate change from getting worse,¹⁶ and that “adaptation” can be described as protecting people against the climate change that is already happening.¹⁷

Our findings are not limited to the US. In a follow-up study we did in Germany, we found similar confusion about climate jargon.¹⁸ Even though the German language has a lot of long sentences, one interviewee said: “[By the time I reached the end of the sentence,] I forgot how the sentence started”.¹⁹ In a study that was conducted in the UK, one participant referred to climate jargon as “science babble”.²⁰

Using everyday language is not only helpful for members of the general public. A policymaker we interviewed about IPCC reports said: “It’s never good if you feel lost [during] a first look because you lose interest, and if you lose interest you just skip to the next page”.²¹

I have been sharing these findings with climate scientists, climate journalists, and others who communicate about climate change. When I talk to these audiences, they often have questions about specific terms they use. A common question I have received is whether it might be a good idea to change the term for climate change.

Should we change the term we use for climate change?

“Climate change” used to be more commonly referred to as “global warming”. Google Trends between 2004–2014 found that, as an internet search term, “global warming” was less common than “climate change”.²² However, “climate change” and “global warming” were searched equally often in 2014, and “climate change” overtook “global warming” in 2016.²³ “Global warming” refers to increasing temperatures and “climate change” refers to broader changes in the climate.²⁴ An advisor to the US government under President George W. Bush popularized the term “climate change” to de-emphasize the role of human activity and sound less alarming.²⁵ In the past, Democratic-leaning individuals also used to think of “global warming” as more serious than “climate change” while Republican-leaning individuals used to think of “climate change” as more serious than “global warming”. But a recent review found that

these partisan differences have now faded and a majority of individuals from both political groups tend to express concern about both terms. Many people currently treat “global warming” and “climate change” as if they have the same meaning.²⁶

Terms like “climate crisis” and “climate emergency” have been used to stress the urgency with which climate action is needed. In 2019, *The Guardian* newspaper switched to using these terms instead of “climate change”.²⁷ The term “climate justice” has been introduced by grassroots organizations to draw attention to climate change as a human rights challenge:²⁸ that is, low-income communities will suffer the most from the effects of climate change despite being the least responsible for causing it.²⁹

Psychologists have long shown that wording matters.³⁰ Words that draw attention to different aspects of a topic can sway public opinion about that topic.³¹ Seemingly slight variations in how options are framed can affect people’s judgments and decisions.³² For example, we have found that food without meat and dairy products is more likely to be chosen when it is labeled “plant-based” rather than “vegan”—and becomes even more popular when labels say “healthy” or “sustainable” to emphasize the benefits of eating food without meat and dairy products.³³ Hence, there is reason to suspect that the terms that have been proposed to emphasize different aspects of “climate change” may affect people’s responses, including their concern about the topic, and their willingness to do something about it.³⁴

We therefore conducted a national US wording experiment that examined people’s responses to the terms “climate change”, “global warming”, “climate crisis”, “climate emergency”, and “climate justice”.³⁵ The study was conducted with a nationally representative sample recruited through the University of Southern California’s Understanding America Study. Overall, “climate change” and “global warming” were rated as most familiar and most concerning, and “climate justice” the least, with ratings for “climate crisis” and “climate emergency” falling in between. Moreover, we found no evidence for “climate crisis” or “climate emergency” eliciting more perceived urgency than “climate change” or “global warming”. Rated willingness to support climate-friendly policies and eat less red meat were less affected by presented terms, while these ratings were lowest for “climate justice”. Although effects of terms on rated familiarity,

concern, and perceived urgency varied by political leaning, “climate justice” generally received the lowest ratings on these variables among Democrats, Republicans, and Independents/others. Auxiliary analyses found that when terms were unfamiliar, participants were generally less likely to express concern, urgency, policy support or willingness to eat less red meat.

Thus, we found no evidence that switching to newer terms will actually raise concern, feelings of urgency or willingness to act.³⁶ That is likely a result of these terms not being very familiar. Even when the alternative terms do become familiar, there is no guarantee that they will raise concern. Indeed, it has been suggested that saying “climate crisis” can backfire due to sounding alarmist.³⁷ A study in the UK found that narratives that focus on “climate justice” are politically polarizing because they resonate with left-leaning audiences and are off-putting to right-leaning audiences.³⁸ Therefore, we recommend sticking with the familiar terms “climate change” and “global warming”.

Conclusion

Climate change affects people around the world. Communities are increasingly being hit by severe weather, including heat waves, droughts, and wildfires. Climate communications aim to inform people about expected climate impacts and what to do about it. But these communications may fail if people have trouble understanding what they are saying. In the words of Pete Ogden, Vice President for Energy, Climate, and the Environment at the United Nations Foundation 2021: “We have to get better at communicating the dire threat posed by climate change if we want to bring more people into the fight for faster and greater action on climate. That starts by being able to translate the best science into language that all of us can understand.”³⁹

Acknowledgements

The work presented here was funded in part by the Lloyd’s Register Foundation, the University of Southern California’s Wrigley Institute for Environment and Sustainability, and the University of Southern California’s Schaeffer Institute for Public Policy and Government Service. I am grateful for my collaborators, includ-

ing Pete Ogden and his team at the United Nations Foundation, and Kate Weber and her team at the University of Southern California's Dornsife Public Exchange. I also want to thank Jill Darling and her team at the University of Southern California's Center for Economic and Social Research for helping us to collect data through the Understanding America Study.

Notes

1. Bruine de Bruin & Dugan 2022.
2. Lloyd's Register Foundation 2024; Bruine de Bruin *et al.* 2025.
3. Lee *et al.* 2015.
4. IPCC 2022.
5. Barkemeyer *et al.* 2015.
6. Bruine de Bruin *et al.* 2021; Bruine de Bruin *et al.* 2024a.
7. Wong-Parodi *et al.* 2013.
8. Paasche-Orlow *et al.* 2003; Davis *et al.* 2006; Daraz *et al.* 2011; Neuhauser & Paul 2011; OECD 2013.
9. Wong-Parodi *et al.* 2013.
10. McLaughlin 1969; Cutts 2013.
11. Bruine de Bruin & Bostrom 2013.
12. Bruine de Bruin *et al.* 2021.
13. Bruine de Bruin *et al.* 2021.
14. Bruine de Bruin *et al.* 2021.
15. Bruine de Bruin *et al.* 2021.
16. Bruine de Bruin *et al.* 2021.
17. Bruine de Bruin *et al.* 2021.
18. Wege *et al.* 2024.
19. Wege *et al.* 2024.
20. Brook Lyndhurst Ltd 2007.
21. Bruine de Bruin *et al.* 2024a.
22. Lineman *et al.* 2015.
23. Penz 2017.
24. Penz 2017.
25. Penz 2017.
26. Benjamin *et al.* 2017.
27. Schäfer *et al.* 2023.
28. Dutta 2019–2020.
29. Schlosberg & Collins 2014; Saraswat & Kumar 2016; Dutta 2019–2020.
30. Sudman *et al.* 1996; Schwarz 2007.
31. Cantril 1944; Payne 1951.
32. Tversky & Kahneman 1981; Levin *et al.* 1988.
33. Sleboda *et al.* 2024.

34. E.g., Schuldt *et al.* 2011.
35. Bruine de Bruin *et al.* 2024b.
36. Bruine de Bruin *et al.* 2024b.
37. Feldman & Hart 2021.
38. Whitmarsh & Corner 2017.
39. United Nations Foundation 2021.

References

- Barkemeyer, Ralf, Suraje Dessai, Beatriz Monge-Sanz, Barbara Gabriella Renzi & Giulio Napolitano 2015. 'Linguistic analysis of IPCC summaries for policymakers and associated coverage', *Nature Climate Change* 6, pp. 311–316.
- Benjamin, Daniel M., Han-Hui Por & David V. Budescu 2017. 'Climate change versus global warming: Who is susceptible to the framing of climate change?', *Environment and Behavior* 49:7, pp. 745–770.
- Brook Lyndhurst Ltd 2007. 'Public understanding of sustainable energy consumption in the home: Final report to the Department for Environment, Food and Rural Affairs', London: Department for Environment, Food and Rural Affairs (DEFRA).
- Bruine de Bruin, Wändi & Ann Bostrom 2013. 'Assessing what to address in science communication', *Proceedings of the National Academies of Sciences of the United States of America* 110, supplement 3, pp. 14062–14068.
- Bruine de Bruin, Wändi & Andrew Dugan 2022. 'On the differential correlates of climate change concerns and severe weather concerns: Evidence from the World Risk Poll', *Climatic Change* 171, 33.
- Bruine de Bruin, Wändi, Lila Rabinovich, Kate Weber, Marianna Babboni, Monica Dean & Lance Ignon 2021. 'Public understanding of climate change terminology', *Climatic Change* 167, 37.
- Bruine de Bruin, Wändi, Lila Rabinovich, Kate Weber, Marianna Babboni, Lance Ignon, Rachel Wald, Monica Dean, Alix Kashdan & Sigourney Luz 2024a. 'Improving figures for climate change communications: Insights from interviews with international policymakers and practitioners', *Climatic Change* 177, 57.
- Bruine de Bruin, Wändi, Laurel Kruke, Gale M. Sinatra & Norbert Schwarz 2024b. 'Should we change the term we use for "climate change"? Evidence from a national U.S. terminology experiment', *Climatic Change* 177, 129.
- Bruine de Bruin, Wändi, Patrycja Sloboda & Tsegaye Ginbo Gatiso 2025. 'Global public concerns about climate change: The role of education, direct experience, and indirect experience', *Journal of Risk Research* 28:3–4, pp. 298–312.
- Cantril, Hadley 1944. *Gauging Public Opinion*, Princeton: Princeton University Press.
- Cutts, Martin 2013. *Oxford Guide to Plain English*, Oxford: Oxford University Press.
- Daraz, Lubna, Joy C. Macdermid, Seanne Wilkins, Jane Gibson & Lynn Shaw 2011.

- ‘The quality of websites addressing fibromyalgia: An assessment of quality and readability using standardised tools’, *BMJ Open* 1:1, pp. 1–10.
- Davis, Terry C., Michael S. Wolf, Pat F. Bass, Mark Middlebrooks, Estela Kennen, David W. Baker, Charles L. Bennett, Ramon Durazo-Arvizu, Anna Bocchini, Stephanie Savory & Ruth M. Parker 2006. ‘Low literacy impairs comprehension of prescription drug warning labels’, *Journal of General Internal Medicine* 21:8, pp. 847–851.
- Dutta, Soumya 2019–2020. ‘From “climate change” to “climate justice”’, *India International Centre Quarterly* 46:3/4, pp. 285–301.
- Feldman, Lauren & P. Sol Hart 2021. ‘Upping the ante? The effects of “emergency” and “crisis” framing in climate change news’, *Climatic Change* 169, 10.
- IPCC/The Intergovernmental Panel on Climate Change 2022. *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.
- Lee, Tien Ming, Ezra M. Markowitz, Peter D. Howe, Chia-Ying Ko & Anthony A. Leiserowitz 2015. ‘Predictors of public climate change awareness and risk perception around the world’, *Nature Climate Change* 5, pp. 1014–1020.
- Levin, Irwin P., Sara K. Schnittjer & Shannon L. Thee 1988. ‘Information framing effects in social and personal decisions’, *Journal of Experimental Social Psychology* 24:6, pp. 520–529.
- Lineman, Maurice, Yuno Do, Ji Yoon Kim & Gea-Jae Joo 2015. ‘Talking about climate change and global warming’, *PloS One* 10, pp. 1–12.
- Lloyd’s Register Foundation 2024. *What the World Worries About: Global Perceptions and Experiences of Risk and Harm*. World Risk Poll 2024 report. <https://wrp.lrfoundation.org.uk/sites/default/files/2024-11/World%20Risk%20Poll%202024%20Report%20%20-%20What%20the%20world%20worries%20about%20.pdf>, accessed 5 December 2024.
- McLaughlin, G. Harry 1969. ‘SMOG grading: A new readability formula’, *Journal of Reading* 12:8, pp. 639–646.
- Neuhauser, Linda & Kala Paul 2011. ‘Readability, comprehension, and usability’, in Baruch Fischhoff, Noel T. Brewer & Julie S. Downs eds, *Communicating Risks and Benefits: An Evidence-based User Guide*, Silver Spring, MD: United States Department of Health and Human Services, Food and Drug Administration.
- OECD/Organisation for Economic Co-operation and Development 2013. ‘Country note. Survey of adult skills: United States’, <https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/piaac/country-specific-material/cycle-1/USA-Country-Note-Cycle1-Round1-2013.pdf>, accessed 12 March 2026.
- Paasche-Orlow, Michael K., Holly A. Taylor & Frederick L. Brancati 2003. ‘Readability standards for informed-consent forms as compared with actual readability’, *The New England Journal of Medicine* 348:8, pp. 721–726.
- Payne, Stanley L. 1951. *The Art of Asking Questions*, Princeton: Princeton University Press.

- Penz, Hermine 2017. “Global warming” or “climate change”? in Alwin F. Fill & Hermine Penz eds, *The Routledge Handbook of Ecolinguistics*, New York: Routledge, pp. 277–292.
- Saraswat, Chitresh & Pankaj Kumar 2016. ‘Climate justice in lieu of climate change: A sustainable approach to respond to the climate change injustice and an awakening of the environmental movement’, *Energy, Ecology and Environment* 1, pp. 67–74.
- Schäfer, Mike, Valerie Hase, Daniela Mahl & Xeno Krayss 2023. ‘From “climate change” to “climate crisis”? Analyzing changes in global news nomenclature from 1996 to 2021’, *Bergen Language and Linguistics Studies* 13:1.
- Schlosberg, David & Lisette B. Collins 2014. ‘From environmental to climate justice: Climate change and the discourse of environmental justice’, *Wiley Interdisciplinary Reviews: Climate Change* 5:3, pp. 359–374.
- Schuldt, Jonathon P., Sara H. Konrath & Norbert Schwarz 2011. “Global warming” or “climate change”? Whether the planet is warming depends on question wording’, *Public Opinion Quarterly* 75:1, pp. 115–124.
- Schwarz, Norbert 2007. ‘Attitude construction: Evaluation in context’, *Social Cognition* 25, pp. 638–656.
- Sleboda, Patrycja, Wändi Bruine de Bruin, Tania Gutsche & Joseph Arvai 2024. ‘Don’t say “vegan” or “plant-based”: Food without meat and dairy is more likely to be chosen when labeled as “healthy” and “sustainable”’, *Journal of Environmental Psychology* 93, pp. 1–9.
- Sudman, Seymour, Norman M. Bradburn & Norbert Schwarz 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*, San Francisco: Jossey-Bass.
- Tversky, Amos & Daniel Kahneman 1981. ‘The framing of decisions and the psychology of choice’, *Science*, New Series, 211:4481, pp. 453–458.
- United Nations Foundation 2021. ‘Climate change communications study finds that common scientific terminology confounds public’, <https://unfoundation.org/media/climate-change-communications-study-finds-that-common-scientific-terminology-confounds-public/>, accessed 5 December 2024.
- Wege, Lena, Wändi Bruine de Bruin & Astrid Kause 2024. ‘Public understanding of climate change terminology in Germany’, *Climatic Change* 177, 81.
- Whitmarsh, Lorraine & Adam Corner 2017. ‘Tools for a new climate conversation: A mixed-methods study of language for public engagement across the political spectrum’, *Global Environmental Change* 42, pp. 122–135.
- Wong-Parodi, Gabrielle, Wändi Bruine de Bruin & Casey Canfield 2013. ‘Effects of simplifying outreach materials for energy conservation programs that target low-income consumers’, *Energy Policy* 62, pp. 1157–1164.

The precautionary principle has no exceptions

MARTIN PETERSON

1. Introduction

The *absolutist* interpretation of the precautionary principle states that any action with a nonnegligible potential for a catastrophic outcome should be avoided, regardless of its benefits, even in the presence of considerable scientific uncertainty about the magnitude and likelihood of the catastrophe.¹ In contrast, *non-absolutist* interpretations take the precautionary principle to be one of several *prima facie* principles that must be balanced against others.²

In an earlier paper, written when I was a young postdoctoral researcher benefiting from extensive and fruitful discussions with Nils-Eric Sahlin, I demonstrated that a wide range of absolutist interpretations of the precautionary principle are incompatible with a set of basic conditions for rational choice.³ Other scholars later expanded on this criticism, and the most comprehensive and technically sophisticated contribution I am aware of is by Orri Stefánsson.⁴ In the current contribution, I will present two new impossibility results for the precautionary principle that are independent of my earlier work.⁵

The first impossibility result explores whether the precautionary principle could be made less demanding and, therefore, more plausible by allowing for exceptions. Consider, for instance, the suggestion that acts carrying a non-negligible potential for a catastrophic outcome should always be avoided whenever possible, *except* when doing so would be extremely costly. I will demonstrate that all absolute versions of the precautionary principle that allow for exceptions are incompatible with certain fundamental conditions for reasonable decision rules. This is a more

general point than the one I made in my original paper because it applies to a wider class of formulations of the precautionary principle.

The second impossibility result builds on Kenneth Arrow's well-known impossibility theorem in social choice theory.⁶ My point is that if we interpret the precautionary principle as a *prima facie* principle to be balanced against other *prima facie* principles, then no method of aggregating such principles can meet a set of plausible conditions that mimic Arrow's conditions for social welfare functions. Because this result does not assume that it is sometimes appropriate to make exceptions to the precautionary principle, it is quite general in scope.

2. An example: The water crisis in Flint, Michigan

Why might it be appealing to allow exceptions to the precautionary principle? The Flint water crisis serves as a compelling illustration, even though the precautionary principle does not play any official role in US drinking water regulations.⁷

In April 2014, the City of Flint ceased purchasing drinking water from the water plant operated by the nearby City of Detroit.⁸ To save money, Flint restarted its own water plant, built in 1954, which draws raw water from the heavily polluted Flint River.⁹ Soon after the switch, residents began to complain about the smell and taste of the water. To address these concerns, the City invited a large engineering firm to review the city's water treatment process in February and March 2015. The company sent two consulting engineers to Flint, who spent about two weeks in the city. In the final report submitted in March 2015, the consulting engineers made ten recommendations, including a recommendation to "Contract with your engineer and initiate discussions with the State [of Michigan] on the addition of a corrosion control chemical."¹⁰

A few months later, concerns were raised about high lead levels in the drinking water. Corrosion initiated by the change in water chemistry was believed to be the root cause, but experts disagreed on how severe the lead issue was. Flint switched back to Detroit water in October 2016, and the Governor of Michigan apologized to the residents. In one of the many lawsuits that followed, the government was ordered to pay \$626 million in damages to 90,000 Flint residents. Five government officials resigned or were fired, and two were charged with involuntary man-

slaughter. In a separate class action lawsuit, plaintiffs alleged that the consulting engineers had failed to give appropriate professional advice to the City of Flint.

Although the precautionary principle played no role in the Flint water crisis, and no regulatory document required the City to apply it, advocates of the precautionary principle will likely argue that it would have been appropriate to invoke it: Was the switch of raw water source compatible with the precautionary principle given the uncertainties about the new water chemistry? Should the City have immediately switched back to Detroit water when the problems occurred? Should all lead pipes have been replaced when the water chemistry changed, including all privately owned service lines? And so on.

By permitting the precautionary principle to have exceptions, we could make it less demanding and less vulnerable to counterexamples. For example, rather than concluding that a change in raw water source is *never* permitted unless *all* uncertainties are eliminated, it might be reasonable to permit exceptions if the cost savings are substantial. Similarly, instead of interpreting the precautionary principle to require the replacement of virtually *all* lead pipes nationwide, including privately owned service lines, exceptions could be made for well-functioning water systems where such replacements are prohibitively expensive.

All these examples suggest that it is worthwhile to explore whether the precautionary principle could reasonably be allowed to have exceptions. In the next couple of sections, I will use formal methods to substantiate a negative answer.

3. The absolutist interpretation

Numerous precisifications of the precautionary principle have been proposed in the literature. If we interpret the precautionary principle as an absolutist principle, then a risky option y that meets the applicable criterion for invoking the precautionary principle will always be ranked below some other option x if the latter does not meet said criterion. This is not the only possible definition of the term “absolutism”. For instance, André Nollkaemper writes that:

In several treaties, the precautionary principle is formulated in absolutist terms. It stipulates that once a risk of a certain magnitude is *identified*, preventive measures to erase that risk are *mandatory*.¹¹

I agree with Nollkaemper that many policymakers formulate the precautionary principle in this manner. However, Nollkaemper's notion of absolutism is much stronger than the one I propose. I do not take absolutism to entail that once a risk of a certain magnitude has been *identified*, then preventive measures to erase the risk are *mandatory*. My definition leaves it open what level of uncertainty and magnitude of catastrophic outcome is required for invoking the precautionary principle.

We can model the notion of absolutism I propose by assuming that the pairwise ranking of the relevant options *remains the same in all choice situations, regardless of the other principles* included in the analysis. This means that if y (an option associated with a risk that warrants precautionary measures) is ranked below x (an option that does not do so) in some option set, then x is ranked above y in *all* option sets.

For the purposes of this paper, choice situations can be identified with the options available to the agent: $A = \{a, b, \dots\}$ is the choice situation in which $a, b, \dots$ are available, and so on. A decision rule is a function that takes the set of all choice situations $S = \{A, B, \dots\}$ as its argument and ranks the options in each choice situation. The formula $x >_P^A y$ means that x is ranked as strictly better than y according to principle P in choice situation A . There is no need to assume that the ranking is complete.

Let $>^A$ (without lower index) denote the all-things-considered ranking in choice situation A derived from some absolutist or non-absolutist principles $x >_P^A, x >_Q^A, \dots$. Because absolutist principles do not have to be balanced against other principles, if P is an absolutist principle, the formula $x >_P^A y$ guarantees that $x >^A y$, and in every other option set B that includes x and y as elements, it also holds that $x >^B y$. If P had been a mid-level or prima facie principle, none of this would have followed. Non-absolutist principles take the ranking of an option to depend on additional principles, meaning that nothing can be inferred about the all-things-considered ranking from a single mid-level or prima facie principle. Formally, if P is a mid-level or prima facie principle, the formula $x >_P^A y$ does not reveal any information about the all-things-considered ranking in any option set.

Consider the following formal definition:

Definition of Absolutist Precautionary Principle: P is an absolutist precautionary principle just in case, for every option set A and every x and y , if $x >_P^A y$ then $x >^B y$ for every B that includes x and y as elements.

This notion of absolutism is inspired by how moral philosophers use the term. As explained by Frank Jackson and Michael Smith, the key characteristic of an absolutist moral principle is that it absolutely prohibits actions of a special kind K , where “ K may be promise breaking (as in Kant), intentionally killing the innocent (as in some Catholic ethical theory)”, and so on.¹² If act x is, say, the breaking of a promise and y is not, then Kantians will always rank y above x regardless of what other options are available. The same goes for the choice between intentionally killing the innocent and not doing so, and my proposal is that we think of absolutism in discussions of the precautionary principle in the same way.

4. Are exceptions possible?

Few people believe that action-guiding principles ought to be absolutist in the sense outlined in the preceding section. Promises must not *always* be kept, and lying is not *always* wrong. And in exceptional cases, it may even be permissible to kill an innocent person if this is necessary to save many more innocent people. Similar exceptions may also apply to the precautionary principle. Suppose the precautionary principle requires us to, for example, replace old lead pipes or lower the action level of the United States Environmental Protection Agency (EPA) for lead in drinking water. In that case, a reasonable version of the precautionary principle must not entail that we are required to do so *without exception*. If the cost of a precautionary measure is exceptionally high or some other relevant circumstances obtain, exceptions would surely be appropriate. For the purposes of this paper, there is no need to specify precisely under what circumstances such exceptions apply; all that matters is that some acts are, under some suitably specified conditions, excepted.

The claim that the precautionary principle has exceptions is best expressed by restricting the set of options to which it is applicable.

Let $>_p^{A-X}$ denote the ranking of principle P applied to all options in the option set A except the subset X . If the precautionary principle has exceptions, those exceptions have to be evaluated by some other principle, such as the cost-benefit principle or the ALARA (As Low As Reasonably Achievable) principle. A system of policy principles that includes an absolutist version of the precautionary principle that permits exceptions will thus have at least two elements: the precautionary principle, which does not apply to some non-empty set of exceptions X , and some other principle Q that ranks the elements of X . Consider the following formal definition:

Exceptions to the Precautionary Principle (EPP): There exists an absolute principle P with exceptions X such that if $a, b \notin X$ and $a >_p^{A-X} b$, then (i) for every choice situation B that includes a and b as elements it holds that $a >^B b$, and (ii) there exists some other (absolute or non-absolute) principle Q that ranks the elements of X in A .

This seemingly plausible version of the precautionary principle is incompatible with three fundamental conditions on rational choice:

Strict Dominance (SD): If every principle R applicable to A ranks $a >_R^A b$, then $a >^A b$.

The SD condition is required for handling cases with three or more policy principles. For systems with just two principles, it can be dropped. The remaining pair of conditions assert that the all-things-considered ranking must be transitive, and that every logically possible choice situation can be evaluated:

Transitivity (TR): For every choice situation A , if $a >^A b$ and $b >^A c$, then $a >^A c$.

Unrestricted Domain (UD): All logically possible choice situations can be evaluated by the policy principles adopted by the decision-maker.

The conditions stated above entail the following impossibility result:

The First Impossibility Result: Conditions *EPP*, *SD*, *TR*, and *UD* are incompatible in every choice situation that has at least three differently ranked options.

The proof is straightforward:¹³ Consider an absolute principle *P* that has some exceptions *X* and another principle *Q* that is applicable to *X*. Condition *UD* guarantees that there exists a choice situation $A = \{a, b, c, d\}$ in which all applicable principles *R* (including *P* and *Q*, if applicable) rank $d >_R^A a$ and $b >_R^A c$. We assume that the choice situation is chosen such that *P*'s ranking is $a >_P^{A-X} b$, but the exceptions *c*, $d \notin X$ are handled by *Q*, which ranks $c >_Q^A d$. *EPP* now guarantees that because $a >_P^{A-X} b$, it holds that $a >^A b$. Moreover, because of clause (ii) of *EPP*, $c >_Q^X d$ guarantees that $c >^A d$, and *SD* entails that $d >^A a$ and $b >^A c$. By putting these rankings together, we find that $d >^A a >^A b >^A c >^A d$, which violates *TR*.

So far, we have assumed that *a*, *b*, *c*, and *d* are distinct options, but it was noted earlier that only three options are required. To see why, consider the case in which *a*, *b* and *c*, *d* have one element in common. Suppose, for instance, that $a = c$ and imagine that the choice situation *A* is such that all applicable principles *R* rank $b >_R^A d$. Let *P*'s ranking include $a >_P^A b$ but let *c* and *d* be excepted from the ranking. The exception is handled by *Q*, which ranks $d >_Q^A c (= a)$. *EPP* entails that $a >^A b$ and $d >^A c$. Because $c = a$, we have $d >^A a$. However, *SD* entails that $b >^A d$. It follows that $a >^A b >^A d >^A a$, which contradicts *TR*. The other cases, in which $a = d$, or $b = c$ or *d*, are analogous.

5. No exceptions in Flint

This section discusses the first impossibility result within the context of the Flint water crisis. As emphasized in the preceding section, this result hinges on the assumption that the precautionary principle is interpreted as an absolutist principle. The second impossibility result, applicable to non-absolutist interpretations, will be explored in the next section.

In October 2024, the United States government under President Joe Biden announced that the EPA would lower the “action level” for lead in drinking water—the threshold triggering mandatory safety measures—from 15 to 10 parts per billion (ppb). A suggestion to lower the action level to 0 ppb, which would have squared well with the precautionary

principle, was dismissed as excessively expensive. It was furthermore announced that lead pipes owned by water utilities but not privately owned service lines will have to be replaced in the next ten years at a cost of approximately \$20 billion.

The Biden administration's decision was based on a cost-benefit analysis that indicated the benefits outweighed the costs.¹⁴ However, the scientific uncertainties are significant. The benefits listed in the analysis include the prevention of ADHD in 2,600 children and the annual loss of 200,000 IQ points in children. These numbers are rough estimates, as is the attempt to translate them into monetary benefits. According to the EPA, the estimated total benefits are worth \$130 to \$250 billion over 10 years, and the estimated cost \$15 to \$20 billion. Critics, however, argue that the cost will be significantly higher and the associated benefits will be substantially smaller.¹⁵

The upside of relying on the precautionary principle instead of cost-benefit analysis is that the former does not force us to quantify costs and benefits, meaning that the conclusion would be less uncertain. As long as the potential for a sufficiently bad outcome is nonnegligible, eliminating lead pipes and reducing the action level it is always better. The following statement summarizes the policy endorsed by the EPA in 2024:

a: Replace the vast majority of service lines and lower the EPA threshold to 10 ppb for \$20 billion.

This formulation does not explicitly mention the benefits; however, if we seek to apply the precautionary principle, there is no need to weigh costs against benefits. It suffices to identify a nonnegligible, potentially catastrophic risk that could be eliminated by taking some precautionary measure—in this example, replacing the vast majority of service lines and lowering the EPA threshold to 10 ppb.

It is instructive to compare option *a* to three further options that EPA could have chosen but did not actually choose. The costs listed below are rough estimates included for illustrative purposes:

b: Replace the vast majority of service lines but keep the EPA threshold at 15 ppb for \$15 billion.

c: Do nothing: keep all lead pipes and keep the EPA threshold at 15 ppb for \$ 0.

d: Replace all lead pipes and lower the EPA threshold to 0 ppb for \$ 50 billion.

Suppose the precautionary principle is applicable to every choice between two or more policies that impact the safety of drinking water, *except* when the difference in cost is at least \$ 50 billion. Clearly, *a* is better than *b* according to the precautionary principle because of the lower action level, and *b* is likewise better than *c* according to the precautionary principle. However, the precautionary principle does not apply to the choice between *c* and *d* because the cost of *d* is \$ 50 billion. Here, we need to apply some other principle. Suppose we choose the cost-benefit principle and conclude that *c* is better than *d*. (The *UD* condition guarantees that some choice situation of this type is possible; the example considered here is just an illustration.) Finally, we apply the precautionary principle to the choice between *d* and *a*, which we can do because the difference in cost is less than \$ 50 billion, and conclude that *d* is better than *a*. Hence, the all-things-considered ranking is $d \succ^A a \succ^A b \succ^A c \succ^A d$, which contradicts TR.

6. The second impossibility result

The second impossibility result targets the idea of interpreting the precautionary principle as a *prima facie* principle. According to W. D. Ross, “right acts can be distinguished from wrong acts only as being those which, of all those possible for the agent in the circumstances, have the greatest balance of *prima facie* rightness”.¹⁶

My goal in this section is to show that a version of Arrow’s well-known impossibility theorem is applicable to the aggregation of *prima facie* principles: no rule for calculating the overall balance of *prima facie* rightness can satisfy a set of conditions stated below, provided that the number of *prima facie* principles is finite but greater than one, and if at least three options are available to the agent. This point is applicable to *every* set of *prima facie* principles; it does not only target the precautionary principle.

Each prima facie principle, including the precautionary principle, ranks a finite set of options $A = \{a, b, \dots\}$ with respect to prima facie rightness. The formula $a >_i b$ means that a is more prima facie right than b with respect to prima facie principle i , and $a > b$ (without index) means that the overall balance of prima facie rightness of a is greater than that of b . An option is right all things considered (or *sans phrase*, as Ross put it) if no other option has a higher balance of prima facie rightness. Consider the following conditions:

Ordering: The overall ranking of prima facie rightness is complete and transitive.

Independence of Irrelevant Alternatives: The pairwise ranking of a and b must not depend on whether some third option c is available.¹⁷

In order to state the remaining conditions, it is helpful to use a technical term invented by Amartya Sen:¹⁸

Definition: A set of prima facie principles D (it can be a singleton set) that is a subset of the set of all prima facie principles G is *decisive with respect to* the ordered pair (a, b) if and only if $a > b$ whenever every prima facie principle i in D ranks $a >_i b$. A set of prima facie principles is *decisive* just in case it is decisive with respect to all pairs of options.

The remaining conditions and the impossibility result can now be stated as follows:

Dominance: The set of all prima facie principles G is decisive.

No Dictatorship: No single prima facie principle (i.e., no singleton set D) in G is decisive.

The second impossibility result: If the number of prima facie principles is finite but greater than one, and the number of distinct options is at least three, then no rule for aggregating prima facie principles satisfies the following conditions: Ordering, Independence of Irrelevant Alternatives, Dominance, and No Dictatorship.

The proof is analogous to Sen's simplified proof of the original version of Arrow's theorem and will not be reproduced here.¹⁹ It is merely the interpretation that is novel. The relevance of this result depends on what other *prima facie* principles we wish to combine with the precautionary principle. Note that *all* *prima facie* principles must satisfy the conditions for the result to hold.

Independence of Irrelevant Alternatives is often considered controversial in social choice theory. Consider the following counterexample discussed by Sen:²⁰ There are two oranges and a kiwi left in a fruit bowl. If I take one of the oranges, you will get to choose between two different types of fruits, an orange, and a kiwi, once it is your turn to pick a fruit. So, because I am a polite person, it seems rationally permissible for me to pick an orange if all three fruits are available, but if there is only a single orange and a kiwi left, the only rationally permissible option is to pick the kiwi, because that is the fruit I like the most. In that case, my desire to leave you with a choice is no longer relevant. Sen's point is that these preferences violate Arrow's original version of Independence of Irrelevant Alternatives because the preference between the kiwi and the first orange depends on an "irrelevant" alternative: the availability of the second orange.

Is this counterexample applicable, *mutatis mutandis*, to *prima facie* principles? For many plausible *prima facie* principles, the answer seems to be no. Consider, for instance, a *prima facie* principle requiring policymakers to be truthful and honest. That principle is not sensitive to "irrelevant" options because the agent is either truthful and honest or not. The requirement to be truthful and honest does not depend on the availability of other options. A similar point is applicable to the *prima facie* precautionary principle: if some option is impermissible to choose, at least *prima facie*, because it is too risky and uncertain, then this is so regardless of what other options are available.

Note that Independence of Irrelevant Alternatives is not a claim about individual *prima facie* principles but about the aggregation of *prima facie* principles. My suggestion is that if no individual *prima facie* principle is sensitive to relational properties (such as being the only option of a certain type), then a plausible aggregation rule should also be insensitive to such properties. This suggests that for many sets of policy principles holding *prima facie*, including the *prima facie* precautionary principle, Arrow's cannot be easily refuted.

7. Conclusion

The impossibility results outlined in this paper present a challenge for defenders of the precautionary principle. The first impossibility result shows that it is not attractive to interpret the precautionary principle as an absolute principle that allows for exceptions. The second result demonstrates that we cannot regard it as a *prima facie* principle. The only viable action guiding interpretation seems to be to interpret it as an absolute principle that does not permit any exceptions—but that would reintroduce the impossibility results that motivated me to write this paper. My overall conclusion is that it is more attractive to interpret the precautionary principle as a belief-guiding principle for addressing epistemic risks.²¹

Acknowledgements

I would like to thank Nils-Eric Sahlin and all other participants in the symposium on ethics, risks and decision-making, held at the Royal Swedish Academy of Letters, History and Antiquities in Stockholm on 17–18 October 2024, for their stimulating discussions and helpful feedback. I would also like to thank Orri Stefánsson for extremely helpful written comments on an earlier draft.

Notes

1. See Steel 2015 for an overview. See also Sandin *et al.* 2002.
2. A third, in my opinion, highly plausible interpretation is to read the precautionary principle as an epistemic principle designed to guide our beliefs; see Peterson 2007. This interpretation is inspired by N.-E. Sahlin's notion of epistemic risk; see, e.g., Sahlin 1983; Sahlin & Persson 1994.
3. Peterson 2006; Peterson 2017a.
4. Stefánsson 2019; see also Boyer-Kassem 2017.
5. The first formal result is structurally similar to Amartya Sen's observation that a particular way of defining rights in social choice theory is incompatible with the Pareto principle; see Sen 1970. The formal analysis in sections 3 and 4 of this paper draws heavily on my working paper 'Can absolute moral principles have exceptions?' Peterson 2024a. The second formal result is structurally similar to Kenneth J. Arrow's (1951) famous impossibility result. My adaptation draws partly on Peterson 2024b.
6. Arrow 1951.
7. Full disclosure: I have served as an expert witness for defendants in several lawsuits related to the Flint water crisis, for which I have been financially compensated.

8. For approximately four years, until spring 2024, I was retained as an expert witness on engineering ethics in a class action lawsuit related to the Flint water crisis. The suit ended with a \$25 million settlement in February 2024 and I was then relieved of my duties. The opinions expressed in this paper are my own and are not intended to have any bearing on legal issues.

9. The National Resource Defense Council (2025) writes that “For more than a century, the Flint River, which flows through the heart of town, has served as an unofficial waste disposal site for treated and untreated refuse from the many local industries that have sprouted along its shores, from carriage and car factories to meatpacking plants and lumber and paper mills.”

10. Veolia North America 2015.

11. Nollkaemper 1996, p. 73.

12. Jackson and Smith 2006 (p. 267) discuss whether absolutist principles can be applied to risky actions in a reasonable manner. This is a topic I will not discuss here.

13. As noted earlier, the proof is similar to Amartya Sen’s (1970) observation that a certain way of defining rights in social choice theory is incompatible with the Pareto principle; see also Peterson 2017b, ch. 13.

14. EPA 2024.

15. Samson *et al.* 2024.

16. Ross 1930, p. 45; italics removed. Ross is careful to point out that the term “prima facie” is somewhat misleading because it suggests that “one is speaking only of an appearance which a moral situation presents at first sight, and which may turn out to be illusory; whereas what I am speaking of is an objective fact involved in the nature of the situation” (1930, p. 20).

17. See Ray 1973 for a discussion of different technical versions of this condition.

18. Sen 1995.

19. See Sen 1995, p. 4, n. 9.

20. Sen 1995.

21. Peterson 2007. See also Sahlin 1983 and Sahlin & Peterson 1994 for thoughtful discussions of the notion of epistemic risk.

References

- Arrow, Kenneth J. 1951. *Social Choice and Individual Values*, 2nd edn 1963, New York: Wiley.
- Boyer-Kassem, Thomas 2017. ‘Is the precautionary principle really incoherent?’, *Risk Analysis* 37:11, pp. 2026–2034.
- EPA/United States Environmental Protection Agency 2024. ‘Lead and copper rule improvements’, https://www.epa.gov/system/files/documents/2024-10/final_lcrl_fact-sheet_general_public.pdf, accessed 25 October 2024.
- Jackson, Frank & Michael Smith 2006. ‘Absolutist moral theories and uncertainty’, *The Journal of Philosophy* 103:6, pp. 267–283.

- National Resource Defense Council [Melissa Denchak] 2025. 'Flint water crisis: Everything you need to know', <https://www.nrdc.org/stories/flint-water-crisis-everything-you-need-know>, accessed 3 March 2026.
- Nollkaemper, André 1996. "'What you risk reveals what you value", and other dilemmas encountered in the legal assaults on risks', in David Freestone & Ellen Hey eds, *The Precautionary Principle and International Law*, Dordrecht: Kluwer Law International, pp. 73–94.
- Peterson, Martin 2006. 'The precautionary principle is incoherent', *Risk Analysis* 26:3, pp. 595–601.
- Peterson, Martin 2007. 'Should the precautionary principle guide our actions or our beliefs?', *Journal of Medical Ethics* 33:1, pp. 5–10.
- Peterson, Martin 2017a. 'Yes, the precautionary principle is incoherent', *Risk Analysis* 37:11, pp. 2035–2038.
- Peterson, Martin 2017b. *An Introduction to Decision Theory*, Cambridge: Cambridge University Press.
- Peterson, Martin 2024a. 'Can absolute moral principles have exceptions?', unpublished manuscript.
- Peterson, Martin 2024b. 'The greatest balance of prima facie rightness', unpublished manuscript.
- Ray, Paramesh 1973. 'Independence of irrelevant alternatives', *Econometrica* 41, pp. 987–991.
- Ross, W. David 1930. *The Right and the Good*, ed. Philip Stratton-Lake, Oxford: Oxford University Press.
- Sahlin, Nils-Eric 1983. 'On second-order probabilities and the notion of epistemic risk', in Bernt P. Stigum & Fred Wenstøp eds, *Foundations of Utility and Risk Theory with Applications*, Dordrecht: Springer Netherlands, pp. 95–104.
- Sahlin, Nils-Eric & Johannes Persson 1994. 'Epistemic risk: The significance of knowing what one does not know', in Berndt Brehmer & Nils-Eric Sahlin eds, *Future Risks and Risk Management* (Risk, Governance and Society 9), Dordrecht: Springer Netherlands, pp. 37–62.
- Samson, Carleigh C., Sheldon V. Masters, Christian Ley Mathews & Chad J. Seidel 2024. 'Estimating the impact of the proposed lead and copper rule improvements', *AWWA Water Science* 6:5, e70004.
- Sandin, Per, Martin Peterson, Sven Ove Hansson, Christina Rudén & André Juthe 2002. 'Five charges against the precautionary principle', *Journal of Risk Research* 5:4, pp. 287–299.
- Sen, Amartya 1970. 'The impossibility of a Paretian liberal', *Journal of Political Economy* 78:1, pp. 152–157.
- Sen, Amartya 1995. 'Rationality and social choice', *The American Economic Review* 85:1, pp. 1–26.
- Steel, Daniel 2015. *Philosophy and the Precautionary Principle*, Cambridge: Cambridge University Press.

Stefánsson, H. Orri 2019. 'On the limits of the precautionary principle', *Risk Analysis* 39:6, pp. 1204–1222.

Veolia North America 2015. 'Flint, Michigan: Water quality report', <https://www.cityofflint.com/wp-content/uploads/Veolia-REPORT-Flint-Water-Quality-201503121.pdf>, accessed 3 March 2026.

Care of extremely preterm-born children

Ethics, risks and difficult decisions

GÖRAN HERMERÉN

Introduction

Practice and policy in Sweden concerning care of extremely preterm-born children is different from policy and practice in all other European countries. There are also some differences between clinics within Sweden concerning the care of these children. In the Netherlands the line for attempts to save extremely preterm-born children is drawn at gestation week 24.¹ In Sweden attempts have been made to save preterm-born infants not only in week 23 and 22 but also in week 21.

Uncertain outcomes, unknown risks and value conflicts make decisions in the care of extremely preterm-born infants difficult. By way of introduction, I mention some general difficulties as well as the importance of how the decision problem is framed and the obvious need to clarify the values at stake. Key issues include the ethical costs of treating these children with intensive care, and the role of the mother/family in the decision-making process.

I will also comment on some of the arguments for and against national guidelines for the care of these children, as well as on the apparent lack of discussion of the costs for this care and for adequate support to those who need it. Finally, some recommendations for debates on the controversies raised by the care of these children are suggested, based on my experiences in various advisory groups.

Some general difficulties

The discussion of these issues is sensitive because it involves the rights and interests of persons with severe functional handicaps. What life is worth living and who is to decide about this? But there are also some other difficulties, related to the theme of this book: knowledge gaps, uncertainties and conflicting values.

“Maximize expected utility” is a well-known slogan in decision theory. Decisions are made on the basis of values and scientific and other evidence, as is often pointed out. Both premises raise problems in this particular context. “What if there is no science to follow, or the science we have points in many directions and with a trembling hand?” as Nils-Eric Sahlin once put it.² We know a fair amount about the health problems in the group of extremely preterm-born children. Nevertheless, the prognosis in the individual case is often difficult.

What do we know? Earlier studies include the EXPRESS study.³ According to a recent study led by Ann Hellström (the TINY study) a great number of children born before gestation week 24 suffer from impaired functions of vital organs such as brain, eyes, lungs, kidneys and bowels.⁴ Many suffer from neuropsychological problems. One out of four children develop autism spectrum condition. 88% suffer from asthma, kidney and bowel or gut-related problems. Only 15 out of 383 children born after on average 23.3 weeks gestation and with a birth weight of 565 grams did not have other documented diagnoses than “normal”; thus they can be considered to be developed normally.

There is also a wide spectrum of relevant values. These values may be interpreted in several ways, and they conflict with each other on many interpretations, as in the debates on euthanasia or abortion. The utilities may be difficult to estimate. The problems of adding and subtracting utilities and their opposites related to happiness and suffering experienced by individuals are well known.

So, what is needed? In general terms, we need more and better knowledge and clarification of the values at stake, including long-term follow-up studies. “Long-term” means more than the first years: we need to know more about the intellectual, emotional and medical development of the extremely preterm-born during their school years and later in life. Such studies will provide relevant information, though they will not

solve the difficult problem: What life is worth living and who is to decide about this?

It should be noted that value issues are already raised at this point: decisions have to be made about what variables are important and should be studied in the follow-up investigations.

Values at stake

To begin with, the stakeholders, their interests and values, as well as conflicts between these values, have to be identified, clarified and examined. The relevant values include health, life, quality of life, participation, effectiveness (in terms of saving lives, improving the quality of life for the newborn and the mother/family), fairness (for instance, non-discrimination on the basis of religious persuasions, political views or ethnic origin), as well as a reasonable relation between the costs and effects of medical interventions. The relations between these values also need to be examined: for instance, life is a condition for health—but also for pain and suffering.

States of affairs can be valued for their own sake or for the sake of something else, usually as a means to a desired end. There is a considerable literature on the conceptual framework used to characterize values. We need to know more than what the values of the stakeholders are and how they are interpreted and characterized. We also need to know how the values are ranked in relative importance by the stakeholders and how they judge distance between the values. The value distance can be exemplified by the question: Is a particular value, such as minimizing the suffering of the child, only a little or much better than another value, for instance, improving the quality of life of the mother?

Moreover, we also need to pay attention to value uncertainty and value instability. “Value uncertainty” here means that the decision-maker values several different outcomes or states of affairs but cannot decide how these values are to be ranked in cases of conflict or specify the value distance between them. By “value instability” in this context is meant that the decision-maker cannot predict when and how the ranking order or the distance between his or her values will change.⁵

In fact, two sets of values may be relevant in this context. When an advisory ethics council such as the European Group on Ethics (EGE) or

the Nuffield Council makes an ethical analysis of current issues, they may use general values such as human rights, democracy and the rule of law as points of departure, and/or distinguish between procedural principles like transparency and accountability and substantive principles such as justice and inclusion.⁶ The values of the stakeholders in a particular case are also identified and examined; they are often more specific and diversified.

Due to knowledge gaps and various kinds of uncertainties, serious mistakes can be made in the care of extremely preterm-born children, which later will be regretted by all stakeholders.

The framing of decision problems

Debates on ethical issues raised by the care of extremely preterm-born children can be carried out in different contexts: for instance, they may be informed by (i) personal experiences of those who have given birth to infants in week 22 or 23, (ii) various external factors, such as the age of the woman, whether she has a partner, whether support from society is available, (iii) an explicit moral framework, such as different versions of utilitarianism or a human rights-based approach,⁷ or (iv) take place in a religious context where different religions provide the context and slogans, such as the “sanctity of life”, play a prominent role. One of the problems will then be that the “sanctity of life” doctrine is ambiguous and can be interpreted in several ways.⁸

A specific difficulty has to do with the framing of decision problems. In their seminal 1981 paper, Amos Tversky and Daniel Kahneman showed that the preferences of options in a decision situation can be changed depending on variations in the framing of acts, contingencies or outcomes. The frame “that a decision-maker adopts is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision-maker”.⁹

The same problem can often be framed in several different ways. For instance, Tversky and Kahneman demonstrate that it makes a difference if the problem is presented in terms of gains or losses of human lives: “choices involving gains are often risk averse and choices involving losses are often risk taking”.¹⁰ This observation is relevant to the two alternatives discussed in the next section. Understanding the psychological principles

that govern the perception of decision problems and the evaluation of probabilities and outcomes is essential for those who want to be able to predict preferences.

Summing up, Tversky and Kahneman state that “we have presented a series of demonstrations in which seemingly inconsequential changes in the formulation of the choice problems caused significant shifts of preference”.¹¹ They conclude the discussion of their results as follows:

Individuals who face a decision problem and have a definite preference (i) might have a different preference in a different framing of the same problem, (ii) are normally unaware of alternative frames and of their potential effects on the relative attractiveness of options, (iii) would wish their preferences to be independent of frame, but (iv) are often uncertain how to resolve detected inconsistencies.¹²

Risks and ethical issues

Not all extremely preterm-born infants survive. But survival has been improved by intensive care and other active efforts at neonatal clinics. This has been demonstrated both by earlier studies (the EXPRESS study) and more recent ones (the TINY study).

However, many of those extremely preterm-born infants who survive will suffer from often life-long and serious functional handicaps because their brains, eyes, kidneys and other vital organs were not developed enough when they were born, as already mentioned.

What risks are worth taking? What risks is it ethically acceptable to expose others to? On what grounds? As is well known, there are two dimensions of risks that both can be graded: the probability dimension (how likely is the outcome?) and the harm or negative value dimension (how bad is the outcome—and for whom?). They can be used to characterize the alternatives outlined below.

The ethical problems raised by care of extremely preterm-born children include the ethical costs of treating a large number of these children with intensive care; ethical costs defined in terms of values that cannot at all, or only to a limited extent, be achieved. Decisions have to be taken, sometimes in haste, and they will have consequences for a number of people concerned: the newborn, the mother, the family, possible siblings, the clinic, other patients and society at large.

Here there is no alternative without ethical costs. These costs have to be weighed against gains. But the classification of outcomes as costs or gains (or advantages and disadvantages) is bound to be controversial. Obviously, what is an advantage or a chance for one stakeholder could be a disadvantage or a risk for others.

Having said this, let us consider the following two (of several possible) alternatives.

Alternative 1: No life-sustaining intensive care is provided to extremely preterm-born infants before gestation week 24.

Advantages: A clear criterion is provided.¹³ Many children with severe functional handicaps are not born.

Disadvantage: Some children that could have been born and would have developed normally are not born, but they are not many.

Alternative 2: Active efforts, including providing intensive care, are made to save the life of every child born before week 24 and showing signs of life.

Advantages: Life-sustaining intensive care of these children will increase the number of survivals. Many children that would not have been born in the Netherlands are born without other diagnoses than “normal” and can be considered to be normally developed, according to the earlier mentioned Hellström study, 15 out of almost 400.

Disadvantages: Many of these children born extremely preterm will suffer from severe functional handicaps, often including life-long functional problems with their brain, eyes, lungs, kidneys and bowels, as shown in the earlier mentioned TINY study led by Hellström. No attention is paid to the situation and wishes of the mother/family. The quality of life for mother and family may be negatively affected, especially in situations when they do not get the support from society they need; single mothers is a particularly vulnerable group.

There is no simple arithmetic that will solve this kind of problem. To save about 15 extremely preterm-born children to a life without other diagnoses than “normal”, close to 400 extremely preterm-born children need to

be treated with intensive care and other interventions. These children will in many cases suffer from often life-long and severe functional handicaps, as indicated by the TINY study. This thought is disturbing or at least problematic, especially if the life and suffering of everyone is considered to have the same value.

It may be tempting to back up this view by referring to a principle of equal value from the beginning—proposed by Ingmar Persson, and stated as follows: no human being is from the beginning worth living a life that for him or her is (subjectively) better or worse than the life any other human being is worth living.¹⁴

This idea can be developed in different ways. But it is important to note that this is not how we think. Experiments and historic observations by Paul Slovic and others show that we do not think that every human life is of equal value.¹⁵ The choices we make are not consistent with this idea. What, if any, are the normative implications of these findings? Do they show that we should abandon the principle of equal value from the beginning, or any other principle of equal value? Not necessarily. It depends on what value premises are added to the description of the results of these empirical studies.

The conclusion might instead be this: if we want to promote principles suggesting or implying that every human life is of equal value, there are a number of challenges we need to address. They are related to how our minds work and that the ranking order of our values that we have stated is not always identical with the ranking order of our values that is revealed by our actions.

Will an artificial uterus be a solution?

Many extremely preterm-born children die early or experience severe life-long health problems since several of their vital organs (brain, eyes, lungs, kidneys and bowels) are not developed enough, as already mentioned. It has been suggested that these problems could be eliminated or reduced if the foetus was put in an artificial uterus. The hypothesis is then that these vital organs would continue to develop and mature during the time in the artificial uterus.¹⁶

Survival of animal embryos for a number of weeks in an artificial uterus has been demonstrated. But “survival” is not very precise. “Survival”—

with how many vital organs functioning normally? Even if the answer is “with all of them”, we know from other cases—the Neurosedyn/Thalidomide scandal,¹⁷ for instance—that we cannot always draw valid conclusions from what seems to be safe for animals to what will be safe for humans. Thus, we must acknowledge that there are uncertainties and partly unknown risks here.¹⁸ For instance, what are the long-term consequences of being in the uterus of the mother vs. being in an artificial uterus outside the body of the mother? At present, we do not know, as Nils-Eric Sahlin has pointed out in an interview.¹⁹ Thus we must conclude that given what we know today, it is difficult to make an informed decision, and easy to make a decision that will be regretted later.

The ethical aspects need to be discussed early, including the consequences of the various alternatives for important values in our culture, such as human dignity, integrity and human rights. Building in ethical concerns from the start in the planning and production process would prevent ethical concerns being added too late to make a difference. Ethics by design is the idea.²⁰ The discussion of ethical aspects should begin before any investments in research, production and marketing have been made.

Moreover, ethical analyses should inform the regulation. There is obviously a need to consider the legal situation. Both hard and soft law is required in this rapidly developing field. If a newborn outside the uterus shows signs of life, then according to at least Swedish law, it has the rights of a child.²¹ But it does not violate the law to provide palliative care to such a child, if the prognosis is poor enough. In other words, the law does not require that life-sustaining intensive care is automatically used to save every child born before gestation week 24 and showing signs of life.

The role of the family

Against this background, let us consider the following ethical problem: To what extent should the mother/family be involved in the decision-making? There are four main alternatives with variations and somewhat different pros and cons:

The doctor decides. According to this alternative, the mother/parents should be consulted and listened to. They should be allowed to ex-

press their fears, hopes, wishes and beliefs, and receive the information relevant to them. But what they say should only be part of the basis for the decision. The decision should be made by the doctor, who obviously should be guided by what he or she has good reasons to believe will be in the best interest of the patient. Thus, the decision should not be left to the mother or family. It is a difficult decision for the mother or family to make, and one that can give rise to regret and blame in the future. This alternative is in line with the established view that the doctor has the medical responsibility for the treatment.

The mother/family decides. The general rule at least in Sweden is that a patient cannot demand a certain healthcare or medical intervention. But the patient is entitled to refuse a treatment proposed by the doctor. In line with this, there is the possibility that the mother/family is entitled to decline a treatment proposed by the doctor. But what does the mother/family say no to? It becomes important how the proposal of the doctor is stated: as a proposal to begin life-sustaining intensive care, to interrupt or stop such care, or not to begin such treatment.

Doctors and parents decide together. The ideal in this alternative is to reach an agreement based on an individualized assessment of each situation. This approach ("shared decision-making") is an essential aspect of person-centred care. The process is important. It is time-consuming but has advantages, though issues of responsibility can become somewhat unclear. Time is needed but is not always available. Besides, the doctor can influence the views of the mother in ways not intended and not realized by the family (or the doctor).

The courts decide. This has happened in the United Kingdom and perhaps elsewhere as well. The Manchester conjoined twins is an example,²² which led to reports and headlines in media similar to these:

"Siamese twins born in England can be separated—against their parents' wishes—to save the life of the stronger baby, even though the operation will kill the weaker twin, a High Court judge ruled last week."

“In the first case of its kind in Britain, Mr Justice Johnson gave the go-ahead at the High Court in London for surgeons to operate on the twin girls, born on 8 August 2000 at St Mary’s Hospital, Manchester.”

A problem is that these decisions are based on underlying values ranked in a certain way and/or on ethical principles, even if the decision is made by a judge. The ranking order of these values and principles should be clarified and argued for. Here is an area where a moral framework based on human rights may clash with utilitarian principles, particularly if the rights at stake includes the right to life.²³

In the UK, there is a tradition for utilitarian thinking; utilitarianism since the days of Bentham has played and plays an important role informing public policy. Is it likely that a German court would arrive at the same decision? Perhaps not. This only shows that in situations of this kind, involving life-and-death issues at the beginning and end of life, legal decisions are based on tacit ethical principles which need to be identified, made explicit and examined.

There are three particular challenges that have to be met in this context:

Two patients. One obvious challenge is that at least in the situation of delivery and during a considerable time afterwards, there are two patients, the mother and the newborn, the latter of whom cannot speak for himself or herself, nor make any decisions. Their interests may not always be identical.

Informed consent. Informed decisions are what is wanted. But it is easy to underestimate the difficulty of providing information that is understood by persons who may be grieving and perhaps also are in a state of shock, particularly when time is short—as it is at the clinic when certain decisions have to be made.²⁴

Lack of self-knowledge. It is also easy to underestimate the risk that the doctor’s own values will influence the selection of information and the order in which pieces of information are given. It can make a difference if the doctor starts with positive or with negative news.

Should one of these models of who should decide be used in all situations, or are there good reasons for using one model in certain situations and other models in other situations? For instance, on the basis of the prognosis, the time of the gestation (week 23 plus 1, or week 23 plus 6), or the resources and willingness of the mother/family to care for a child with possibly severe functional handicaps? It has been indicated that in Belgium parents are more involved in the decision-making during the antenatal phase than during the postnatal phase.²⁵

If different models are to be used, the reasons for choosing a particular model in a specific situation need to be made explicit. Should this be made clear in national or professional guidelines?

Guidelines?

General rules can be counterproductive, as the prognosis in the individual case may be difficult and the situation of the mother or family often varies. Families have different hopes, fears, resources and cultural backgrounds. Moreover, the mother may be young, say 21, with many chances to become pregnant again, or she may be 43 and has tried IVF unsuccessfully several times; she may be single or have a partner; she may already have one or more other children with or without functional handicaps; she may or may not be a drug addict. In addition, society can or cannot provide adequate support to those who need it.

It may be argued that these variations should not play a major role in the decision-making in this context; the best interest of the child should be in focus, as stated in the United Nations' Convention on the Rights of the Child. But this suggestion is of little help as long as there is disagreement both about the goal—What exactly is in the best interest of the child in these situations?—and about the methods—How should this goal be achieved? Certainly, all life-prolonging intensive care is not in the best interest of the child. These disagreements need to be discussed and the reasons for and against different positions examined.

I belong to those who are in favour of national guidelines provided that they take the context into account. My general reason for this position is that what healthcare one receives should be based on medical needs and should not depend on where one is living, what hospital one visits or what doctor one happens to meet.

In addition to this general reason, there are at least two concerns raised by national guidelines: (a) concerns related to the risks if there are no national guidelines and (b) concerns related to the risks if existing guidelines are too generalized.

Related to (a). The absence of national guidelines makes it possible for some clinics to routinely provide life-sustaining intensive care to every child born before gestation week 24 and showing weak signs of life, disregarding the wishes as well as the situation of the mother/family; and this also creates the peculiar situation that neighbouring hospitals can have different policies.

Related to (b). To be useful the guidelines must take the context into account. The situation of the mother can vary considerably. Families are different in a number of ways as already mentioned. For these reasons I am critical of guidelines such as “In week 22 life-sustaining intensive care can be considered”. It is both too vague and too generalized. The expression “can be considered” covers a variety of processes and efforts—from a hasty thought to a careful analysis and consideration of arguments and conditions in the situation at hand.

What is needed is a clause specifying *under what circumstances* such life-sustaining intensive care in week 22 can be considered. This requires research, thinking and well-considered judgment. Ethical issues must first be identified and analysed.

The elephant in the room

Instead of pressing the frontiers and continuing the race to try to save preterm-born children in week 21—incidentally close to the limit of the time when abortion can be permitted in Sweden—a case could be made for improving the health-care of extremely preterm-born children and increasing the support to those families and single mothers that need it. But this also raises the question about costs: Who is to pay, and where should the money come from?

The costs are considerable. For instance, the cost of treating necrotizing enterocolitis—which about 20% of the preterm-born children

in the Hellström study suffered from—has been estimated in the United States to \$500,000 per child.²⁶ Moreover, the costs are not just carried by the neonatal clinics in the healthcare system: municipalities have to pay to support families and single mothers. This raises several issues, and attempts to eliminate the variations between the support available in different municipalities may clash with their right to govern themselves.

The elephant in the room is then this: everyone realizes that more money is needed to improve the care and to provide more and better support to those who need it. But nobody seems keen on initiating a debate on alternative uses of resources and on priority-setting in this sensitive area, though there is a general and legitimate interest in maintaining a reasonable relation between costs and effects of medical interventions.

Such a debate raises formidable issues: How are the needs of the various stakeholders to be interpreted and weighed against each other? On the assumption that no patient needs a treatment that has no (positive) effect on his or her condition, how are the effects of life-sustaining intensive care provided to those born before gestation week 24 to be characterized and ranked, given the results of the Hellström study? What costs, and whose costs, are to be taken into account? And how are the costs to be related to the effects? Should the costs only be considered in the last analysis, when effect and need have been clarified, or should we follow the suggestions of some health economists in treating cost-effectiveness as an independent principle as important as other ethical principles?²⁷

Conclusions and recommendations

Let us return to the starting point: decisions are made on the basis of values and scientific and other evidence. Up to now I have drawn attention to problems related to both these premises. Do we have to end on these difficulties, or can something constructive be said about ways to deal with them?

To begin with, two situations have to be separated:

First, there is the situation when a group of experts from different disciplines, such as brain research, intensive care, law, ethics and communication, together with representatives from patient organizations, convene to draft policy documents, for instance, national guidelines of the

sort referred to earlier. This way there is little time pressure; the group can work for months to arrive at their recommendations.

Second, there is the situation at a neonatal unit when a child is born extremely preterm. Then some decisions need to be made quickly. Those immediately involved include the newborn child, the mother/family, the doctors and the nurses. Other patients, the clinic and society at large are also indirectly affected by the decisions made.

There is some overlap between what is required in these two situations. Science communication is important in both, and here there is more work to be done. Obviously, medical jargon should be avoided, and opportunities to ask questions need to be provided. The process is important in both situations. But there are also considerable differences, mainly related to the constraints of time at the clinic.

Drawing on experiences during the recent pandemic, Nils-Eric Sahlin and Sylvia Schwaag Serger underline the importance of science communication and propose recommendations for researchers, universities, media and decision-makers.²⁸ They stress the need to think ahead, to use multidisciplinary perspectives and to justify the recommendations. There are some similarities between their recommendations and the ones proposed below. But there are also differences related to the fact that the problems and the stakeholders are not the same in challenges raised by pandemics and the care of extremely preterm-born infants.

For decision-making at the neonatal unit, the following recommendations may be suggested:

- Stakeholders have to be respected, heard and listened to.
- Let all stakeholders involved express their concerns, hopes and fears.
- Provide basic and relevant medical information to the mother/family.
- Avoid medical jargon and provide opportunity for the mother/family to ask questions.
- Provide opportunity for the health-care staff, doctors and nurses to meet regularly and discuss difficult cases.
- Create and maintain conditions for constructive discussions.

In the policy group, efforts have to be made to understand the way those with different views make sense of the world, particularly in cases when their views and values are not shared by others.

For decision-making in the policy group, the following recommendations may be suggested:

- Begin with what is comparatively simple: identify the stakeholders and those who should be heard.
- Ensure there is agreement concerning the procedure.
- Identify knowledge gaps and various kinds of uncertainties.
- Strive for transparency—make points of departure explicit, particularly about values.
- Equals should be treated equally; avoid double standards and discrimination on medically irrelevant grounds.
- Keep the dialogue going—even if a certain vagueness may be required to achieve this.
- Explain the reasons for the recommendations.

If robust decisions are wanted, it is important to (i) listen to different views and check the evidence on which they are based, as well as to (ii) anticipate objections to the proposals made and discuss how these objections can be dealt with.

The debate should go beyond arguing in terms of slogans like “Every newborn child has a right to adequate and appropriate treatment”. What is the purpose of the treatment and what methods are to be used to achieve that purpose? This needs to be clarified and argued for.

The recommendations made above should only be considered as a starting point. They can be developed further by distinguishing between different situations and challenges. In particular, it needs to be clarified who is responsible for the various actions recommended here (provide..., ensure..., explain..., treat..., create..., maintain...).

Forward-looking responsibility issues include: (1) What should society, authorities and agencies do to create conditions, including regulations, information and economic support? (2) What should various involved individuals (doctors, nurses, heads of clinics, patients, etc.) do or avoid doing in order to minimize harm and to increase benefits for patients?

A diagram indicating the various tasks mentioned above can be used to specify the responsibilities of the various authorities, agents and persons involved. To create and complete a diagram of this kind could be the starting point of a constructive discussion.

The message in this section of my paper is mainly to stress that there are ways of dealing with difficult situations other than relying on one's gut feeling. There is more work to be done. Progress is possible, but time and patience are required.

Acknowledgement

I want to thank Nils-Eric Sahlin for helpful comments on an earlier version of this text.

Notes

1. For overviews, see Zeitlin *et al.* 2013; Aujolat *et al.* 2018.
2. Sahlin 2021, p. 1. For a perspective on scientific research emphasizing the importance of knowing what one does not know, see Persson & Sahlin 2013.
3. See, e.g., Serenius *et al.* 2015; Serenius *et al.* 2016; Hellgren *et al.* 2020.
4. For an overview, see Morsing *et al.* 2023. For more details, see Lundgren *et al.* 2021; Lundgren *et al.* 2022; Hellström *et al.* 2022; Morsing *et al.* 2022.
5. For a discussion of these definitions and some applications, see Hermerén *et al.* 2014.
6. See, for instance, EGE 2021, pp. 4–11; EGE 2024, pp. 29–35.
7. See, e.g., Gewirth 1986; Gewirth 1996; Beyleveld & Brownsword 2001; Singer 2011; and Kagan 2018 for such frameworks.
8. See, e.g., Kuhse 1987; Weingarten 2007.
9. Tversky & Kahneman 1981, p. 453.
10. Tversky & Kahneman 1981, p. 453.
11. Tversky & Kahneman 1981, p. 457.
12. Tversky & Kahneman 1981, pp. 457–458.
13. This criterion in terms of gestational age can be supplemented by other criteria in terms of body weight.
14. Persson 2004, p. 41; my translation.
15. Slovic 2020. See also www.arithmeticofcompassion.org
16. See, for instance, Solerte 2020; Bulletti *et al.* 2023.
17. See Wikipedia's article 'Thalidomide scandal' for background and references.
18. This is also stressed by Nils-Eric Sahlin, interviewed by Maria Backman; see Backman 2024.
19. Backman 2024.

20. See, e.g., Cummings 2006; EGE 2021, pp. 14–17.
21. Rynning 2004.
22. Dyer 2000.
23. Harris 2001; Holm & Erin 2001.
24. See Johansson *et al.* 2020 for a discussion of some relevant issues concerning the relations between foetus–mother, child–mother.
25. Aujolat *et al.* 2018, p. 3.
26. Ley *et al.* 2019, p. 4.
27. In Sweden these issues have been discussed, e.g., in the medical journal *Läkartidningen*, a debate to which Nils-Eric Sahlin made important contributions. See, for instance, Sahlin 2008.
28. Sahlin & Schwaag Serger 2022, pp. 241–243.

References

- Aujoulat, Isabelle, Séverine Henrard, Anne Charon, Ann-Britt Johansson, Jean-Paul Langhendries, Anne Mostaert, Danièle Vermeulen & Gaston Verellen 2018. ‘End-of-life decisions and practices for very preterm infants in the Wallonia–Brussels Federation of Belgium’, *BMC Pediatrics*, 18(1):206. <https://doi.org/10.1186/s12887-018-1168-x>
- Backman, Maria 2024. ‘Science fiction-livmodrar kan börja användas på människor redan i höst’, *Göteborgs-Posten*, 20 August, <https://www.gp.se/livsstil/science-fiction-livmodrar-kan-borja-anvandas-pa-manniskor-redan-i-host.303de0cf-12e7-4a9f-bd72-5916b253dc8a>, accessed 27 February 2026 (interview with Nils-Eric Sahlin).
- Beyleveld, Deryck & Roger Brownsword 2001. *Human Dignity in Bioethics and Biolaw*, Oxford: Oxford University Press.
- Bulletti, Francesco Maria, Romualdo Sciorio, Antonio Palagiano & Carlo Bulletti 2023. ‘The artificial uterus: On the way to ectogenesis’, *Zygote* 31:5, pp. 457–467. <https://doi.org/10.1017/S0967199423000175>
- Cummings, Mary L. 2006. ‘Integrating ethics in design through the value-sensitive design approach’, *Science and Engineering Ethics* 12, pp. 701–715. <https://doi.org/10.1007/s11948-006-0065-0>
- Dyer, Clare 2000. ‘Conjoined twins separated after long legal battle’, *BMJ* 2000;321:1175.
- EGE/European Group on Ethics 2021. *Values for the Future: The Role of Ethics in European and Global Governance*, Brussels: European Commission.
- EGE/European Group on Ethics 2024. *Solar Radiation Modification: Ethical Perspectives*, Brussels: European Commission.
- Gewirth, Alan 1986. ‘Why rights are indispensable’, *Mind: New Series* 95:379, pp. 329–344.
- Gewirth, Alan 1996. *The Community of Rights*, Chicago: University of Chicago Press.

- Harris, John 2001. 'Human beings, persons and conjoined twins: An ethical analysis of the judgement in *Re A*', *Medical Law Review* 9:3, pp. 221–236. <https://doi.org/10.1093/medlaw/9.3.221>
- Hellgren, Kerstin, Lena Jacobson, Paolo Frumento, Jenny Bolk, Ulrika Ådén, Melissa E. Libertus & Mariagrazia Benassi 2020. 'Cerebral visual impairment captured with a structured history inventory in extremely preterm born children aged 6.5 years', *Journal of the American Association for Pediatric Ophthalmology and Strabismus* 2020;24(1):28.e1–28.e8. <https://doi.org/10.1016/j.jaapos.2019.11.011>
- Hellström, Ann, Lena Jacobson, Abbas Al-Hawasi, Lena Hellström-Westas, Alexander Rakow, Mats Johnson, Karin Sävman, Gerd Holmstrom, Eva Larsson, Lotta Gränse, Marie Saric, Birgitta Sunnqvist, Lois Smith, Anna-Lena Hård, Eva Morsing & Pia Lundgren 2022. 'Retrospective evaluation of ophthalmological and neurological outcomes for infants born before 24 weeks gestational age in a Swedish cohort', *BMJ Open* 12(8):e055567. <https://doi.org/10.1136/bmjopen-2021-055567>
- Hermerén, Göran, Ingar Brinck, Johannes Persson & Nils-Eric Sahlin 2014. 'Value uncertainty and value instability in decision-making', in Julien Dutant, Davide Fassio & Anne Meylan eds, *Liber Amicorum Pascal Engel*, Geneva: University of Geneva, pp. 100–110.
- Holm, Søren & Charles A. Erin 2001. 'Deciding on life: An ethical analysis of the Manchester conjoined twins case', in Ludger Honnefelder, Dieter Sturmer, Jan P. Beckmann, Christophe Horn & Ludwig Siep eds, *Jahrbuch für Wissenschaft und Ethik*, vol. 6, Berlin: Walter de Gruyter, pp. 67–88.
- Johansson, Mats, Göran Hermerén & Nils-Eric Sahlin 2020. 'Paternal consent in prenatal research: Ethical aspects', *Medicine, Health Care and Philosophy* 23:2, pp. 325–331. <https://doi.org/10.1007/s11019-019-09919-1>
- Kagan, Shelly 2018 [1998]. *Normative Ethics*, New York: Routledge.
- Kuhse, Helga 1987. *The Sanctity of Life Doctrine in Medicine: A Critique*, New York: Oxford University Press.
- Ley, David, Tomas Abrahamsson, Magnus Domellöf, Henrik Hagberg, Baldvin Jonsson & Ann Hellström 2019. 'Extremt förtidig födelse—svåra utmaningar trots stora framsteg', *Läkartidningen*, vol. 116.
- Lundgren, Pia, Lena Jacobson, Anna-Lena Hård, Abbas Al-Hawasi, Eva Larsson, Lotta Gränse, Marie Saric, Birgitta Sunnqvist, Kristina Tornqvist, Agneta Wallin, Gerd E. Holmstrom, Lois L.E. Smith, Eva Morsing & Ann Hellström 2021. 'High rate and large intercentre variability in retreatment of retinopathy of prematurity in infants born <24 gestational weeks', *BMJ Open Ophthalmology* 6:e000695. <https://doi.org/10.1136/bmjophth-2020-000695>
- Lundgren, Pia, Eva Morsing,, Anna-Lena Hård, Alexander Rakow, Lena Hellström-Westas, Lena Jacobson, Mats Johnson, Gerd Holmström, Staffan Nilsson, Lois E. Smith, Karin Sävman & Ann Hellström 2022. 'National cohort of infants born before 24 gestational weeks showed increased survival rates but no improve-

- ment in neonatal morbidity', *Acta Paediatrica* 111:8, pp. 1515–1525. <https://doi.org/10.1111/apa.16354>
- Morsing, Eva, Pia Lundgren, Anna-Lena Hård, Alexander Rakow, Lena Hellström-Westas, Lena Jacobson, Mats Johnson, Staffan Nilsson, Lois E. H. Smith, Karin Sävman & Ann Hellström 2022. 'Neurodevelopmental disorders and somatic diagnoses in a national cohort of children born before 24 weeks of gestation', *Acta Paediatrica* 111:6, pp. 1167–1175. <https://doi.org/10.1111/apa.16316>
- Morsing, Eva, Pia Lundgren, Anna-Lena Hård, Alexander Rakow, Lena Hellström-Westas, Lena Jacobson, Mats Johnson, Staffan Nilsson, David Ley, Karin Sävman & Ann Hellström 2023. 'Allt fler barn födda före vecka 24 överlever, men sjukligheten är hög', *Läkartidningen*, vol. 120.
- Persson, Ingemar 2004. *Jämlikhet från början*, Nora: Nya Doxa.
- Persson, Johannes & Nils-Eric Sahlin 2013. *Vetenskapsteori för sanningssökare*, Stockholm: Fri Tanke Förlag.
- Rynning, Elisabeth 2004. 'Åldersgräns för mänskliga rättigheter? Om rätten till hälso- och sjukvård vid livets början', in Anna Hollander, Rolf Nygren & Lena Olsen eds, *Barn och rätt*, Uppsala: Justus, pp. 149–180.
- Sahlin, Nils-Eric 2008. 'Från allas lika värde till hälsoekonomisk kalkyl', *Läkartidningen*, vol. 105.
- Sahlin, Nils-Eric 2021. 'Foreword', in Nils-Eric Sahlin, *Vetenskap och beprövad erfarenhet: Slutrapport från ett forskningsprogram* [Science and Proven Experience: Final report from a research programme], Lund: Media-Tryck, pp. 9–14.
- Sahlin, Nils-Eric & Sylvia Schwaag Serger 2022. 'Decision-making in a time of spin and unspoken values', in Mats Benner, Göran Marklund & Sylvia Schwaag Serger eds, *Smart Policies for Societies in Transition: The Innovation Challenge of Inclusion, Resilience and Sustainability*, Cheltenham: Edward Elgar Publishing, pp. 228–244.
- Serenius, Fredrik, Mats Blennow, Karel Marsál, Gunnar Sjös & Karin Källen 2015. 'Intensity of perinatal care for extremely preterm infants: Outcomes at 2.5 years', *Pediatrics* 135(5):e1163–72. <https://doi.org/10.1542/peds.2014-2988>
- Serenius, Fredrik, Uwe Ewald, Aijaz Farooqi, Vineta Fellman, Maria Hafström, Kerstin Hellgren, Karel Maršál, Andreas Ohlin, Elisabeth Olhager, Karin Stjernqvist, Bo Strömberg, Ulrika Ådén & Karin Källén 2016. 'Neurodevelopmental outcomes among extremely preterm infants 6.5 years after active perinatal care in Sweden', *JAMA Pediatrics* 170:10, pp. 954–963. <https://doi.org/10.1001/jamapediatrics.2016.1210>
- Singer, Peter 2011 [1980]. *Practical Ethics*, 3rd edn, Cambridge: Cambridge University Press.
- Slovic, Paul 2020. 'The more who die, the less we care: Confronting the deadly arithmetic of compassion [Keynote address at the 41st Annual North American Meeting of the Society for Medical Decision Making, 21 October 2019, Portland, Oregon]', *Medical Decision Making* 40:4, pp. 407–415. <https://doi.org/10.1177/0272989X20919294>

- Solerte, Maria Laura 2020. 'Artificial uterus—Research background to improve survival and outcome of extremely low birth weight newborns', *Journal of Gynecological Research and Obstetrics* 6:3, pp. 67–71. <https://doi.org/10.17352/jgro.0000909-28>
- Tversky, Amos & Daniel Kahneman 1981. 'The framing of decisions and the psychology of choice', *Science*, New Series, 211:4481, pp. 453–458. <https://doi.org/10.1126/science.7455683>
- Weingarten, Michael 2007. 'On the sanctity of life', *British Journal of General Practice* 57:537, pp. 332–333.
- Zeitlin, J., K. Szamotulska, N. Drewniak, A. D. Mohangoo, J. Chalmers, L. Sakkeus, L. Irgens, M. Gatt, M. Gissler & B. Blondel 2013. 'Preterm birth time trends in Europe: A study of 19 countries', *BJOG: An International Journal of Obstetrics & Gynaecology* 120:11, pp. 1356–1365. <https://doi.org/10.1111/1471-0528.12281>

Local knowledge, proven experience and scientific evidence

JOHANNES PERSSON

Local knowledge and proven experience

It is common to clearly differentiate between local and scientific knowledge. An early example is Stephen Toulmin and June Goodfield's comparison between scientific knowledge and knowledge generated in the "basic crafts".¹ According to the authors, the crafts—such as agriculture, medicine, metalworking, weaving, dyeing, perfumery and glass-making—are the oldest elements in human society: "There one can still hear the voice of the authentic practical man, suspicious of upstart theories only a few centuries old".²

A more recent example has emerged from the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). A controversy between participants using two distinct framings, "ecosystem services" and "Mother Earth", led to a decision on a conceptual framework³ that could be interpreted as dualistic. On the one hand, it contains the main elements and relationships for the conservation and sustainable use of biodiversity and ecosystem services, human well-being and sustainable development. These expressions are highlighted in green in the figure accompanying the decision. On the other hand, the figure contains expressions, in blue colour, such as "systems of life", "living in harmony with nature" and "Mother Earth".⁴ In the IPBES-2 final report, this was transferred to the differentiation we are interested in in this paper: "text in green denotes the concepts of science; and text in blue denotes those of other knowledge systems".⁵ It should be noted that the authors remark that this is a simplification, but it is nevertheless an example of a possible and rather sharp differentiation between scientific knowledge and local knowledge.

Our understanding of what local knowledge constitutes is far from perfect. The crafting of a theoretical term often requires another theoretical expression against which it can be developed and in relation to which it makes immediate sense. “Local knowledge” and its conceptual cousins are very much like that. One gets the feeling that local knowledge is only labelled as “local” when pitted against “scientific knowledge”. Even major characterizations of local knowledge build on the contrast with scientific knowledge. Firket Berkes’ criteria provide one example (although it should be noted that he in fact uses the expression “traditional ecological knowledge”). According to Berkes, local knowledge

1. is mainly qualitative (as opposed to quantitative)
2. has an intuitive component (as opposed to being purely rational)
3. is holistic (as opposed to reductionist)
4. considers mind and matter as united (as opposed to being separated)
5. is moral (as opposed to supposedly value-free)
6. is spiritual (as opposed to mechanistic)
7. is based on empirical observations and accumulation of facts by trial-and-error (as opposed to experimentation and systematic, deliberate accumulation of fact)
8. is based on data generated by resource users themselves (as opposed to that by a specialized cadre of researchers)
9. is based on diachronic data, i.e., long time-series on information on one locality (as opposed to synchronic data, i.e., short time-series over a large area).⁶

There are pros and cons of highlighting the science–local knowledge distinction. Among the advantages are that it clearly points to the fact that there is knowledge outside academic science and what is published in peer-reviewed papers. Among the disadvantages are that neither science nor knowledge outside academia are particularly homogenous. It might well be that intra-level variation exceeds inter-level differences. My conjecture is that this is the case with local knowledge.

I want to point to another possible approach to local knowledge. There is an interesting and relevant Swedish conceptual pair, *vetenskap och beprövad erfarenhet*, or “science and proven experience”, which is useful to

compare with the internationally much more well-known distinction between scientific knowledge and local knowledge. In Sweden, clearly the notion of “science and proven experience” has been emphasized in various contexts. It has influenced the medical and health professions since the late 19th century, and arguably longer, and more recently the educational professions. The “proven experience” part of the pair is one variety of local knowledge, I would argue. What is special with the Swedish notion is that the relations between science and proven experience have been in focus, rather than the relata themselves. Swedish legislation requires that the decision-maker pays attention to both:

Each physician, whether he be an employee at an institution or an independent practitioner of the medical profession, is obliged, as his chief duty, to deliver such counsel, and, as far as circumstances permit, to extend such therapeutic endeavours, to every patient under his care as are necessitated by the patient's condition and as are consonant with science and proven experience.⁷

However, the “proven experience” notion was relatively uncharted before a large research program funded by Riksbankens Jubileumsfond and led by Nils-Eric Sahlin was undertaken.⁸ Johannes Persson, Annika Wallin, Barry Dewitt and Lena Wahlberg have reported that proven experience has an evidentiary role in clinical decision-making, and this role depends in part on its certainty and systematicity.⁹ Three dimensions of proven experience that were found salient in the research program are that it can function as test or evidence, that professionals acquire it (someone becomes experienced/competent) and that it originates from or is accepted in practice. This conceptualization was shared by both physicians and nurses.¹⁰

In comparison, there has been plentiful interest in the notion of local knowledge from anthropology, development studies, conservation biology, etc. Several, although surprisingly similar, definitions have been proposed. So, how can the comparison be useful? One reason, I think, is precisely that the relations between science and proven experience are placed at the centre stage, and that this has not really been the case when local knowledge has been discussed. True, local knowledge has been important for turning our attention to questions of rights, legitimacy, etc., but more seldom—as far as I can judge—has the focus been on the

question about the relative strengths and weaknesses of local knowledge compared with scientific knowledge in decision-making in the real world. Exactly this question is often on the agenda when science and proven experience are discussed. This paper will clearly not contribute to the study of the history of the Swedish notion (or to the history of the notion of local knowledge in general), but I want to note that there might be an interesting connection between the fact that proven experience can apply to an individual who has experience of an activity and not only to the activity itself, i.e., you can be a person of proven experience and there can be proven experience of an activity (for instance, an intervention). For the Swedish reader I add a paragraph from 1720:

Så bör ock vid alla embetens besättning förnämligast anses personernes bepröfvade erfarenhet och förtjänst, hvilka igenom studier, krigsförrättningar och flere nyttige vetenskaper gjordt sig till befordran värdige, men särdeles iakttages, att inge till alldes ovane och med deras ämnen ej öfverensstämmande sysslor måge antagas, utan hvar och en befordras efter dess naturlige böjelse och erfarenhet, så att dugelige personer till tjänsterne förordnas; ej heller må någons förvärfvade höghet, gunst och anseende gifva en eller annan insteg i sådana sysslor och embeten, hvartill han ingen skickelighet hafver, mindre må någon för ringare stånd eller härkomst förskjutas, när han eljest till tjänsten bepröfvas vara skickelig och värdig.¹¹

Putting a substantial value on diachronic, well-tried experience seems to yield Berkes' points 8 and 9. There is arguably an intimate connection between this feature of proven experience and the way he characterizes local knowledge. In addition, whereas it might be that today's ecologists favour synchronic data gathering by conducting experiments with short duration in many places at once rather than studying one ecosystem for a longer time,¹² by requiring that decision-makers should be individuals of proven experience the gap becomes reduced in size.

The growing importance of local knowledge: IPCC and IPBES

In modern times governmental bureaucracies have not always treated local knowledge as knowledge *per se*.¹³ Sometimes local knowledge has been perceived primarily as an obstacle to progress in presence of scientific

knowledge. Arun Agrawal mentions the 1950s and 1960s, in particular.¹⁴ Farmers' traditions have been described as "'backward', conservative, inefficient, inferior, and based on 'ignorance' or myths".¹⁵

However, this tendency has changed, if not everywhere then at least in important policy contexts, such as the intergovernmental panels on biodiversity and climate change. The reasons for this are interesting in their own right, but a subject for another paper. Part of the answer might be that scientific consensus is no longer enough for decision-making. The Intergovernmental Panel on Climate Change, IPCC, is recognized as a pioneer in providing policy-relevant science. It has orchestrated scientific knowledge on climate change: "In doing so it has spoken on behalf of global science with one voice, thereby acquiring a reputation as the epistemic authority in matters of climate policy. It was jointly awarded the 2007 Nobel Peace Prize along with former United States vice president Al Gore."¹⁶ Almost from the start, IPCC utilized a strategy aiming to produce standardized assessments similar to the evidence-based medicine movement. It is widely known that IPCC has experienced challenges, and the critique in 2009–2010, when among other things it was communicated that grey literature was uncritically cited, probably strengthened the standardization strategy. Scientific, peer-reviewed knowledge became the gold standard.¹⁷ Within evidence-based approaches, the integration of evidence, interpreted in this more rigorous way, with practical experience becomes a challenge, and the risk is that local knowledge will have no real place as evidence.¹⁸ By implication IPCC has struggled with alternative forms of expertise, such as local knowledge¹⁹—knowledge that can be notoriously difficult to assign evidentiary value to within such a model. Then came the Paris Agreement in 2015 stating that climate change adaptation "should be based on and guided by the best available science and, as appropriate, traditional knowledge, knowledge of indigenous peoples and local knowledge systems, with a view to integrating adaptation into relevant socioeconomic and environmental policies and actions, where appropriate".²⁰ This ambition is interesting from the perspective of the interaction between science and local knowledge: mentioning is of course one thing, integrating is something very different. IPCC is rather distant from projects carried out locally (where it is more evident that and how local knowledge can play a role). Given its history, how could IPCC accomplish the task set out in the Paris Agree-

ment? The Technical Summary (TS) that complements and expands key findings up to 2021 provides an update. It clearly states:

Compared to earlier IPCC assessments, this report integrates more strongly across the natural, social and economic sciences, highlighting the role of social justice and diverse forms of knowledge, such as Indigenous knowledge and local knowledge, and reflects the increasing importance of urgent and immediate action to address climate risk.²¹

Below are two examples from the Technical Summary:

Indigenous knowledge and local knowledge can contribute to reducing the vulnerability of communities to climate change (medium to high confidence). Supporting Indigenous self-determination, recognising Indigenous Peoples' rights and supporting Indigenous knowledge-based adaptation are critical to reducing climate change risks and effective adaptation (very high confidence).²²

Taking account of interdisciplinary scientific information, Indigenous knowledge and local knowledge and practical expertise is essential to effective ecosystem-based adaptation (high confidence). There is a large risk of maladaptation where this does not happen (medium confidence).²³

There is thus clear recognition in IPCC of the growing importance of local knowledge. Indigenous and local knowledge is increasingly referred to in IPCC reports.²⁴

The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, IPBES, accorded greater value to regional and local scales from the start. To one extent this is explained by the more local nature of what affects biodiversity. Moreover, an initial gap analysis indicated the need to include other, non-formal types of knowledge. IPBES started as a multistakeholder initiative involving actors from conservation and Indigenous peoples' organizations, prominent scientists, environmental agreements and state representatives.²⁵ An interesting debate on what counts as legitimate knowledge, how it is generated and how this influences its practical effectiveness took place in the formative period of IPBES:

The knowledge of traditional and “ordinary” citizens might not meet scientific criteria or be amenable to standardization, but ignoring or misappropriating such experience undermines the possibilities for innovation.²⁶

Thus, in the case of IPBES too, the establishment of the panel contributed to the growing recognition of the importance of local knowledge in decision-making.

In addition, there are clear indications that the use of local knowledge in scientific enterprises increases. For instance, Ryan Brook and Stéphane McLachlan found that, for journals indexed in the Biological Abstracts database, studies using experience-based local knowledge increased from less than 1% in 1980 to 8% in 2004.²⁷

Scientific and local knowledge:

A well-known but problematic distinction

The existing scientific literature on local knowledge uses the expression in rather similar ways. Three themes dominate the characterization of local knowledge: relations between living organisms and the local environment, the interaction between individuals and societies with the local environment, and the transfer of knowledge from generation to generation:

Indigenous and local knowledge (ILK) refers to dynamic bodies of integrated, holistic, social and ecological knowledge, practices and beliefs pertaining to the relationship of living beings, including people, with one another and with their environments.²⁸

Traditional ecological knowledge is a cumulative body of knowledge and beliefs, handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and with their environment.²⁹

Indigenous knowledge is local knowledge—knowledge that is unique to a given culture or society. [...] Such knowledge is passed down from generation to generation, in many societies by word of mouth.³⁰

It might well be that these uses fit certain phenomena that we would agree constitute local knowledge. However, from an epistemological perspective there might be reason to be hesitant. It is clear that we want local knowledge and scientific knowledge to be used together in decision-making. When this happens it is likely that at least for the decision-maker in question their local knowledge (and perhaps their understanding of the scientific knowledge as well) is affected to some degree.

Now, this is something that research on local knowledge nowadays agrees with. The IPBES glossary speaks of “dynamic” bodies of knowledge, and others go further. Intepe and colleagues state that local knowledge is “a highly dynamic and diverse knowledge, constantly shifting with changing circumstances, environments, new layers of understanding, and through interactions with ‘scientific’ or ‘Western’ knowledge”.³¹

However, the question is whether this dynamic character is compatible with the characterizations above. To be both dynamic and integrated is not easy, especially so if we are looking at a body of knowledge that is “handed down through generations”. My suspicion is that Berkes’ conceptualization cannot easily harness evidence and practice that evolve in the intersection between research and development and professional practice, at least not right away. Neither can it easily account for local knowledge that travels between local communities. It follows, as far as I can understand, that such a variety of local knowledge is difficult to combine with the idea of import of scientific evidence and other varieties of evidence into local knowledge. If this is true, it entails that Berkes’ conception of local knowledge might lead to more isolated versions of local knowledge rather than more integrated ones.

An improved understanding should be open to the generation of new local knowledge (and new scientific knowledge) through the feeding of problems between science and local knowledge, much like the interaction between science and proven experience.

Three reasons against a sharp divide:

Local knowledge in science

I want to conclude by briefly touching on three reasons that I believe tell against a sharp demarcation. They all have to do with the presence of local knowledge in science. (I agree that they should be presented

together with cases of scientific knowledge in local knowledge, but that would require a more careful analysis of what local knowledge is, and so has to await a later paper.)

First, it is evident that several scientific disciplines have imported much of their content from local knowledge throughout history. That local biological knowledge has been an important source in the early formation of biological science is clear. For instance, Rumphius' 17th-century work played a role in Linnaeus' classificatory work, and Rumphius certainly benefited from local biological knowledge in producing his catalogue, *Herbarium Amboinense*. He, his wife and his daughters took advantage of local knowledge in the production of the catalogue. His wife, Susanna, had access to traditional gynaecological knowledge of the local women that was not disseminated outside their circles. There are recordings of how one of his daughters "cunningly followed the teacher when she collected her mind-enhancing plant on a rubbish dump [...] and brought a specimen home".³²

Second, science contains a great deal of essential knowledge that closely resembles local knowledge. The roles of skill and personal experience are often understated in the philosophy of science. Experiment, observation, speculation, calculation and measurement are essential components of science.³³ Setting up an experiment that works, designing a survey or measuring a phenomenon requires skill and experience. "Local knowledge is based on data generated by resource users themselves (as opposed to that by a specialized cadre of researchers)", says Berkes,³⁴ but in the field, or laboratory, there is no difference between resource users and scientists. There is, if you prefer, a vast domain of such local knowledge in science. It is usually not communicated to other kinds of decision-makers than scientists themselves. You need to be a skilled experimentalist in order to generate scientific evidence.

Third, local knowledge is relatively often imported into science, both as data and as hypotheses or results to be validated scientifically. For instance, local knowledge is imported into scientific assessments of the abundance and distribution of forest elephants in Cameroon's timber-production forests,³⁵ as there is not enough scientific data to build on in those remote areas.

Notes

1. Toulmin & Goodfield 1965.
2. Toulmin & Goodfield 1965.
3. IPBES n.d.
4. Díaz *et al.* 2015.
5. IPBES n.d., p. 2.
6. Berkes 1993.
7. Pontin 1891, § 59; my translation.
8. Persson *et al.* 2019.
9. Persson *et al.* 2022.
10. Dewitt *et al.* 2021.
11. Regeringsformen 1720, § 40.
12. Vucetich *et al.* 2025.
13. Norgaard 1984.
14. Agrawal 1995.
15. Thrupp 1989.
16. Beck *et al.* 2014.
17. Turnhout *et al.* 2012.
18. Persson *et al.* 2018.
19. Beck *et al.* 2014.
20. United Nations 2015.
21. IPCC 2023, TS A.1, Background.
22. IPCC 2023, TS B.7.2.
23. IPCC 2023, TS D.4.7.
24. Carmona *et al.* 2023.
25. Esguerra & van der Hel 2021.
26. Turnhout *et al.* 2012.
27. Brook & McLachlan 2008.
28. IPBES, glossary.
29. Berkes 1993.
30. Warren 1991.
31. Intepe *et al.* 2023.
32. Baas & Veldkamp 2014.
33. Hacking 1983.
34. Berkes 1993.
35. Brittain *et al.* 2020.

References

- Agrawal, Arun 1995. 'Dismantling the divide between indigenous and scientific knowledge', *Development and Change* 26:3, pp. 413–439. <https://www.doi.org/10.1111/j.1467-7660.1995.tb00560.x>
- Baas, Pieter & Jan F. Veldkamp 2014. 'Dutch pre-colonial botany and Rumphius's Ambonese Herbal', *Allertonia* 13, pp. 9–19.
- Beck, Silke, Maud Borie, Jason Chilvers, Alejandro Esguerra, Katja Heubach, Mike Hulme, Rolf Lidskog, Eva Lövbrand, Elisabeth Marquard, Clark Miller, Tahani Nadim, Carsten Nesshöver, Josef Settele, Esther Turnhout, Elefteria Vasileiadou & Christoph Görg 2014. 'Towards a reflexive turn in the governance of global environmental expertise: The cases of the IPCC and the IPBES', *GAIA: Ecological Perspectives for Science and Society* 23:2, pp. 80–87. <https://www.doi.org/10.14512/gaia.23.2.4>
- Berkes, Fikret 1993. 'Traditional ecological knowledge in perspective', in Julian T. Inglis ed., *Traditional Ecological Knowledge Concepts and Cases*, Ottawa: International Program on Traditional Ecological Knowledge, International Development Research Centre.
- Brittain, Stephanie, Madeleine Ngo Bata, Paul De Ornellas, E. J. Milner-Gulland & Marcus Rowcliffe 2020. 'Combining local knowledge and occupancy analysis for a rapid assessment of the forest elephant *Loxodonta cyclotis* in Cameroon's timber production forests', *Oryx* 54:1, pp. 90–100. <https://www.doi.org/10.1017/S0030605317001569>
- Brook, Ryan K. & Stéphane M. McLachlan 2008. 'Trends and prospects for local knowledge in ecological and conservation research and monitoring', *Biodiversity and Conservation* 17:14, pp. 3501–3512. <https://www.doi.org/10.1007/s10531-008-9445-x>
- Carmona, Rosario, Graeme Reed, Stefan Thorsell, Dalee Sambo Dorough, Joanna Petrsek MacDonald, Tunga Bhadra Rai & Gideon Abraham Sanago 2023. 'Analysing engagement with Indigenous peoples in the Intergovernmental Panel on Climate Change's Sixth Assessment Report', *NPJ Climate Action* 2, 29. <https://www.doi.org/10.1038/s44168-023-00048-3>
- Dewitt, Barry, Johannes Persson, Lena Wahlberg & Annika Wallin 2021. 'The epistemic roles of clinical expertise: An empirical study of how Swedish healthcare professionals understand proven experience', *PLoS One* 16:6, e0252160. <https://www.doi.org/10.1371/journal.pone.0252160>
- Díaz, Sandra *et al.* 2015. 'The IPBES Conceptual Framework—connecting nature and people', *Current Opinion in Environmental Sustainability* 14, pp. 1–16. <https://www.doi.org/10.1016/j.cosust.2014.11.002>
- Esguerra, Alejandro & Sandra van der Hel 2021. 'Participatory designs and epistemic authority in knowledge platforms for sustainability', *Global Environmental Politics* 21:1, pp. 130–151. https://www.doi.org/10.1162/glep_a_00573

- Hacking, Ian 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge: Cambridge University Press.
- Intepe, Demet, Robert Šakić Trogrlić, Maria Evangelina Filippi, Thirze Hermans, Hannah Bailon & Anuzska Maton 2023. 'The dangers of romanticising local knowledge in the context of disaster studies and practice', in Rohit Jigyasu & Ksenia Chmutina eds, *Routledge Handbook on Cultural Heritage and Disaster Risk Management*, London: Routledge, pp. 147–164.
- IPBES/Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services n.d. 'Decision IPBES-2/4: Conceptual framework for the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services', <https://www.ipbes.net/document-library-catalogue/decision-ipbes-24>, accessed 2 March 2026.
- IPBES/Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. 'Glossary', <https://www.ipbes.net/glossary>, accessed 2 March 2026.
- IPCC/Intergovernmental Panel on Climate Change 2023. 'Technical Summary', in IPCC, *Climate Change 2022—Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press, pp. 37–118.
- Norgaard, Richard B. 1984. 'Traditional agricultural knowledge: Past performance, future prospects, and institutional implications', *American Journal of Agricultural Economics* 66:5, pp. 874–878. <https://www.doi.org/10.2307/1241018>
- Persson, Johannes, Emma L. Johansson & Lennart Olsson 2018. 'Harnessing local knowledge for scientific knowledge production: Challenges and pitfalls within evidence-based sustainability studies', *Ecology and Society* 23:4, 38. <https://www.doi.org/10.5751/ES-10608-230438>.
- Persson, Johannes, Niklas Vareman, Annika Wallin, Lena Wahlberg & Nils-Eric Sahlin 2019. 'Science and proven experience: A Swedish variety of evidence-based medicine and a way to better risk analysis?', *Journal of Risk Research* 22:7, pp. 833–843. <https://www.doi.org/10.1080/13669877.2017.1409251>
- Persson, Johannes, Annika Wallin, Barry Dewitt & Lena Wahlberg 2022. 'Certainty and systematicity of practice-derived evidence matter for its relative importance in professional decision-making: Survey results on the role of proven experience in Swedish medicine, nursing, OT, dentistry, and dental hygiene', *International Journal of Nursing Studies Advances* 4. <https://www.doi.org/10.1016/j.ijnsa.2022.100074>
- Pontin, David Magnus 1891. *Författningar m.m. angående medicinalväsendet i Sverige, omfattande år 1890*, Stockholm: Norstedt & Söner.
- Regeringsformen 1720 (1891), <https://runeberg.org/regform/0115.html>, accessed 2 March 2026.
- Thrupp, Lori Ann 1989. 'Legitimizing local knowledge: Scientized packages or empowerment for third-world people', conference paper presented at the Joint Meeting of the Society for Social Studies of Science/The European Association for the Study of Science and Technology, Amsterdam.

- Toulmin, Stephen & June Goodfield 1965 [1962]. *The Architecture of Matter*, Harmondsworth: Penguin Books.
- Turnhout, Esther, Bob Bloomfield, Mike Hulme, Johannes Vogel & Brian Wynne 2012. 'Listen to the voices of experience', *Nature* 488:7412, pp. 454–455. <https://www.doi.org/10.1038/488454a>
- United Nations 2015. *Paris Agreement*, <https://unfccc.int/documents/37107>, accessed 2 March 2026.
- Vucetich, John A., Sarah R. Hoy & Rolf O. Peterson 2025. 'More reason for humility in our relationships with ecological communities', *BioScience* 75:2, pp. 163–171. <https://www.doi.org/10.1093/biosci/biae129>
- Warren, D. Michael 1991. *Using Indigenous Knowledge in Agricultural Development* (World Bank Discussion Papers 127), Washington, DC: The World Bank.

What's action got to do with it?

MONIKA GRUBER

Introduction

My goal in this paper is twofold. First, I aim to show that the epistemological “standard” view, saying that knowledge is justified and true belief, possibly with additional conditions, is by no means traditional, obvious or indeed standard. I will support my claims by quoting several prominent philosophers dating from antiquity up to today. I will argue that the infamous entailment thesis does not deserve the default status it enjoys, and that it was by sheer inattentiveness that its alleged universal acceptance came to be overstated. I will further argue that knowledge and belief are two disparate and even mutually exclusive states.

Second, I will discuss the role of action in human decision-making. We all have knowledge about certain facts and we also have subjective beliefs about states we hope to be the case. We act on both our knowledge and our beliefs. I will argue that our actions differ depending on whether we act on our knowledge or on our beliefs. This should support my first claim that an agent cannot have a belief that such and such is the case and know it to be the fact at the same time.

A quick look at the history of philosophy

The standard analysis of knowledge in contemporary epistemology is the view that a subject *S* knows a proposition *p* if 1) *p* is true; 2) *S* believes that *p*; 3) *S* is justified in believing that *p*. Hence, knowledge is justified true belief (JTB).¹ The condition that (propositional) knowledge entails belief is called the entailment thesis and has been considered by the majority as uncontroversial. Numerous accounts of knowledge and belief have been constructed involving the entailment thesis, thereby granting

it default status. However, since Edmund Gettier showed in 1963 that true justified belief is not enough to be knowledge, philosophers have been trying to upgrade belief for it to become knowledge. The generally accepted view in today's epistemology is that knowledge is something more than "merely" belief, that it requires further conditions. Barely any arguments for the entailment thesis have been presented, and few philosophers have argued against it.

The JTB view of knowledge is very generally applied, however by no means standard or traditional. As Maria Antognazza points out, ancient philosophers knew better than to confound these two distinctly separate concepts. Plato, Socrates and Aristotle as well as Thomas Aquinas held that knowledge and belief are two separate and completely different notions:

According to these traditional views, knowing and believing are distinct in kind, in the strong sense that they are mutually exclusive mental states: the same cognitive subject cannot, simultaneously and in the same respect, be in the state of both knowing and believing the same thing. Knowing is not 'the best kind of believing'; nor is believing to be understood derivatively from knowledge.²

The supposedly "standard and traditional JTB view of knowledge" is apparently not traditional at all. Let's take a closer look at the writings of the three fathers of philosophy.

First, Plato in *Meno* argues that knowledge and right opinion are very different from each other:

And indeed I too speak as one who does not know but only conjectures: yet that there is a difference between right opinion and knowledge is not at all a conjecture with me but something I would particularly assert that I knew: there are not many things of which I would say that, but this one, at any rate, I will include among those that I know.³

Later Plato argues that belief is of no lesser value than knowledge, i.e., that it is just as useful as knowledge in guiding our actions:

And so long, I presume, as he has right opinion about that which the other man really knows, he will be just as good a guide—if he thinks the truth instead of knowing it—as the man who has the knowledge. [...] Hence true

opinion is as good a guide to rightness of action as knowledge; and this is a point we omitted just now in our consideration of the nature of virtue, when we stated that knowledge is the only guide of right action; whereas we find there is also true opinion. [...] Then right opinion is just as useful as knowledge.⁴

As long as we make them firm with causal reasoning, i.e., when an opinion is fastened up it is worth as much as knowledge. It is just as useful in guiding our actions.

So that right opinion will be no whit inferior to knowledge in worth or usefulness as regards our actions, nor will the man who has right opinion be inferior to him who has knowledge.⁵

Hence, true/right opinion or belief are just as good to have as knowledge since they guide our actions with no lesser success. We should not let our guard down, however, and conclude that knowledge and true belief are the same because they are equally useful.

This is not the case, as Plato emphasizes in one of his longest dialogues, *Theaetetus*. Here, he considers what knowledge is and how we arrive at it:

This is how it is: he who sees has acquired knowledge, we say, of that which he has seen; for it is agreed that sight and perception and knowledge are all the same. But he who has seen and has acquired knowledge of what he saw, if he shuts his eyes, remembers it, but does not see it.⁶

Later, Plato concludes that we acquire knowledge through thinking:

Then knowledge is not in the sensations, but in the process of reasoning about them; for it is possible, apparently, to apprehend being and truth by reasoning, but not by sensation. [...] and indeed now at last it has been made perfectly clear that knowledge is something different from perception.⁷

What sort of opinion or belief can qualify to be knowledge?

THEAETETUS: To say that all opinion is knowledge is impossible, Socrates, for there is also false opinion; but true opinion probably is knowledge. Let that be my answer. For if it is proved to be wrong as we proceed, I will try to

give another, just as I have given this. Assuming that there are two kinds of opinion, one true and the other false, do you define knowledge as the true opinion?

SOCRATES: Yes. That now seems to me to be correct.⁸

The part which interests us the most is where Plato considers whether true belief can be said to be knowledge:

SOCRATES: If, my boy, the command to add reason or explanation means learning to know and not merely getting an opinion about the difference, our splendid definition of knowledge would be a fine affair! For learning to know is acquiring knowledge, is it not?

THEAETETUS: Yes.

SOCRATES: Then, it seems, if asked, "What is knowledge?" our leader will reply that it is right opinion with the addition of a knowledge of difference; for that would, according to him, be the addition of reason or explanation.

THEAETETUS: So it seems.

SOCRATES: And it is utterly silly, when we are looking for a definition of knowledge, to say that it is right opinion with knowledge, whether of difference or of anything else whatsoever. So neither perception, Theaetetus, nor true opinion, nor reason or explanation combined with true opinion could be knowledge.

THEAETETUS: Apparently not.⁹

Plato's dialogues are very didactical. They usually start with a question and suggest possible answers only to reject them in the end and come to an unexpected but well-considered conclusion. The reason why Gettier ascribes the view that justified true belief is knowledge to Plato could be simply that Gettier did not finish reading the very long dialogue *Theaetetus* and took one of the considered hypotheses to be Plato's final opinion. To be fair, Gettier only said in a footnote that JTB seems to be the view Plato is considering in *Theaetetus* and perhaps accepting in *Meno* 98. It is, however, very probable that since Gettier's short paper gained such resonance as it did, even this comment, made in a footnote, got carried along. Further damage was done by other inattentive readers, of both Plato and Gettier.

Aristotle agreed with his teacher's view that knowledge and opinion are two separate and very different states. In *Posterior Analytics* he says:

Knowledge and its object differ from opinion and its object in that knowledge is of the universal and proceeds by necessary propositions; and that which is necessary cannot be otherwise; but there are some propositions which, though true and real, are also capable of being otherwise. Obviously it is not knowledge that is concerned with these; if it were, that which is capable of being otherwise would be incapable of being otherwise.¹⁰

For Aristotle intuition is the starting point of knowledge. Between knowledge, intuition and opinion, only the latter is uncertain and capable of being false and of being otherwise. Furthermore:

Besides, no one thinks that he is “opining” when he thinks that a thing cannot be otherwise; he thinks that he has knowledge.¹¹

Hence, when I believe that such and such is the case I have an opinion, not knowledge. However, knowledge and opinion can have the same object:

If you apprehend propositions which cannot be otherwise in the same way as you apprehend the definitions through which demonstrations are effected, you will have not opinion but knowledge; but if you only apprehend that the attributes are true and not that they apply in virtue of the essence and specific nature of their subject, you will have not true knowledge but an opinion, of both the fact and the reason for it,—that is if you have reached your opinion through the immediate premises; otherwise you will have an opinion only of the fact. [...] it is impossible to have an opinion and knowledge at the same time about the same object; otherwise one would apprehend that the same thing both could and could not be otherwise, which is impossible. Knowledge and opinion of the same object may exist separately in different minds in the sense which we have explained; but they cannot so exist in one mind.¹²

It means simply that one person can know something which another one only believes to be the case. Aristotle has been translated as follows: *episteme* means knowledge, *doxa* is an opinion, and *nous* translates as intuition but also as intellect, mind, sense, or reason. These further translations emphasize how dependable and steadfast intuition was for Aristotle. From these three states, only *doxa* can be false. Knowledge and intuition are always true. Additionally, knowledge is of necessary kind, while opinion is contingent.

Yet another great thinker who holds the view that knowledge and belief are disparate notions and that we can either have belief or knowledge of something, but never both at the same time, is Thomas Aquinas. In his greatest work, *Summa Theologiae*, he states firmly:

Faith implies assent of the intellect to that which is believed. Now the intellect assents to a thing in two ways. First, through being moved to assent by its very object, which is known either by itself (as in the case of first principles, which are held by the habit of understanding), or through something else already known (as in the case of conclusions which are held by the habit of science). Second the intellect assents to something, not through being sufficiently moved to this assent by its proper object, but through an act of choice, whereby it turns voluntarily to one side rather than to the other: and if this be accompanied by doubt or fear of the opposite side, there will be opinion, while, if there be certainty and no fear of the other side, there will be faith.

Now those things are said to be seen which, of themselves, move the intellect or the senses to knowledge of them. Wherefore it is evident that neither faith nor opinion can be of things seen either by the senses or by the intellect.¹³

Thomas sums up the, for us, relevant essence of his deliberations about knowledge and belief:

[I]t is impossible that one and the same thing should be believed and seen by the same person. Hence it is equally impossible for one and the same thing to be an object of science and of belief for the same person. It may happen, however, that a thing which is an object of vision or science for one, is believed by another.¹⁴

What to think then? Some of the earlier voices in contemporary epistemology arguing against the entailment thesis and for the superiority of belief include Colin Radford and Bernard Williams.

Kevin Mulligan presents his insights into the writings of early German phenomenologists who also understood knowledge and belief to be separate states. Mulligan discusses the ideas of Friedrich Bassenge, Dietrich von Hildebrand, Leonard Nelson, Adolf Reinach and Max Scheler among others, and claims that:

Von Hildebrand's main reason for thinking that knowledge, of whatever variety, is not any sort of belief or conviction, is a claim made first of all by Reinach, perhaps the greatest of all the phenomenologists. Belief and conviction, like emotions, are, Reinach argues, *attitudes* (*Stellungnahmen*). These may vary in degrees and usually come in one of two polarly opposed kinds—belief and disbelief, positive and negative conviction, joy and sadness. Attitudes, on this view, are not always propositional: admiration, for example, is a non-thatish attitude. But knowledge is not any sort of attitude—it does not admit of degrees and has no polar opposite. Attitudes are reactions to what is known. Knowledge is no reaction. So knowledge wears the trousers. Belief and convictions are reactions to what we know.¹⁵

The understanding of these notions by phenomenologists is, however, very complex and very different from what concerns me. Nevertheless, my argument for understanding belief and knowledge as disparate notions gains more supporters.

Another philosopher who holds that knowledge and belief are two independent and completely different mental states is John Cook Wilson. He claims that if we have one of these mental states, about a certain proposition, we cannot have the other:

Belief is not knowledge, and the man who knows does not believe at all what he knows, he knows it.¹⁶

What is important for Cook Wilson is that we arrive at different mental states through distinct processes. For example, what distinguishes judgement from knowledge is the agent's decision. There is no decision involved in knowledge. We do not decide to have knowledge, we gain it, we arrive at it. We, however, decide to judge or believe something.

A judgement is a decision. To judge is to decide. It implies previous indecision; a previous thinking process, in which we are doubting. Those verbal statements, therefore, which result from a state of mind not preceded by such doubt, statements which are not decisions, are not judgements, though they may have the same verbal form as judgements.¹⁷

The way he uses the concept of “judgement” here, it could also be a belief. Hence, to believe is to decide. At the same time, Bernard Williams rejects the idea that one could decide what to believe. This has to do with one of

the characteristics he ascribes to beliefs, namely: aiming at truth. It is worth mentioning Williams, for in his 'Deciding to believe' he holds that the entailment thesis expresses "a deep prejudice in philosophy" and that in fact knowledge is an impoverished belief.¹⁸ Perhaps, the most important feature of belief for Williams is that it aims at truth. "To believe that so and so is one and the same as to believe that that thing is true".¹⁹ Furthermore, "although the most straightforward, simple, and elementary expression of a belief by a language-using creature is an assertion, the assertion of p is neither a necessary nor, and this is the point I want to emphasise, a sufficient condition of having the belief that p ".²⁰ It is not a necessary condition, simply because I may have beliefs which I do not express. It is not sufficient, because I may claim that I believe p without it truly being the case, i.e., my assertion that p can be insincere. This feature, according to Williams, is what makes belief superior to knowledge.

Williams presents an idea of a machine producing what he calls *B-states*. Keeping in mind that belief aims at truth, is expressed in an assertion, and that factual beliefs can be based on evidence, the machine would produce true assertions, or rather messages. It would arrive at these messages making an inference based on evidence. In this way it would produce the so-called *B-states*, which are much impoverished versions of our beliefs. The essential reason for this is that the machine would be incapable of making insincere assertions. This, we humans have much less problems doing. It is our decision; we act from our will to withhold our true belief or to lie. The machine cannot do it. It takes the direct route from the gathered information through the preprogrammed aim at truth and produces a true assertion, a *B-state*. If the machine were to print out a false message it would be due to a malfunction, not due to its intention to do it. The key component is the free will. Humans can choose whether to make a true assertion about our belief or whether to withhold the truth. Concluding: "It is, however, a notable feature of this machine that it could produce true *B-states* which were non-accidentally arrived at, that is which were not randomly turned out but were the product of the environment, the programming and so on; and these might be called 'knowledge'."²¹ In this sense we could say that this machine knows the train and bus schedules as well as the city's street plan and even where it is itself. In this usage of "knows" it would mean "has the information". Therefore, knowledge consists solely in having the information, while

belief involves our will to express or withhold it. As much as I agree with Williams' interpretation of knowledge as simply having the information, I believe there is much more to belief than our will to express or withhold it. Humans are not machines without free will, without the capability of deciding what to believe. The argument of not being able to decide what to believe or reveal/withhold was crucial for Williams' machine to always produce true messages. It does not apply to humans. We do decide what we believe. Sometimes our beliefs turn out to be true and sometimes not. To believe is to think, to know does not actively involve thinking or deciding at the precise moment. It is stating a fact: it does not involve a current process of intellectual activity, it is only stating that such and such is the case.

A possible solution?

There certainly is a connection between knowledge and belief, even if they are different, or even mutually exclusive states. We both know and believe things. There are no feelings or emotions when it comes to knowledge, it is factive. States of affairs obtain whether we like it or not, whether we know it or not. We all act on our knowledge as well as on our beliefs, hence it may seem that action is the one thing knowledge and belief have in common. I believe, however, that our actions differ depending on whether we act on beliefs or on knowledge. Frank Ramsey, once again, provides an excellent example:

To say a man has such and such knowledge, beliefs and opinions means then generally something hypothetical, something about what he would think, say, or do in suitable circumstances. It is, in my view, important to realise that it is not only a question of what he would think or say but also of what he would do, for many of our dispositional beliefs are manifested far more in our actions than in our thoughts. For instance, I have a dispositional belief (or perhaps I should rather say knowledge) that the Cambridge Union is in Bridge Street; but this belief is very rarely manifested in an act of thought: I do not often have occasion to judge that that is where the Union is: I only do this when I have to inform a stranger, or just now when I thought of it for an example. On the other hand, this belief of mine is frequently manifested by my turning my steps that way when I want a book from the Union Library, which I do without any process of thought which could properly be called thinking that the Union is in Bridge Street.²²

This, I think, is the major difference between acting on a current belief and on knowledge. When I know something for a fact I do not have to actively think about it, I can just act habitually. When, on the other hand, I'm now forming a belief when I look out the window, see the dark clouds and think that it will rain soon, then based on this belief, I decide to take an umbrella before I leave the house. We act more consciously on our current beliefs than we do on our knowledge. A present belief requires more involvement on the agent's side. Whether we call it assent or a subjective feeling of conviction, it means that we are fully aware of our belief as we act on it and we act because of it. That's why belief is the more complex concept requiring further conditions: mental assent, feeling of conviction, full awareness or consciousness. That's why to believe something requires more current awareness or action than simply to know it.

Action seems to be more crucial when it comes to beliefs than to knowledge. We know a lot of things we never act on. However, for Ramsey, actions define beliefs.

[W]e shall however find it impossible to give any satisfactory account of belief or even of thought without making any reference to possible resulting actions.²³

Indeed, an agent's actions seem central in all discussions about belief and knowledge. Ramsey's example differentiating between acting on belief and on knowledge illustrates clearly how they differ. For him belief is strictly connected to action and only when we act on a state can we say it's our belief. Nothing happens if we just know something without using this knowledge in order to achieve something. If we don't act on our beliefs to fulfil our wishes or desires, are they even our beliefs at all? Or are they just our thoughts? If we don't apply our knowledge and beliefs to our actions, then they are just our thoughts and are of no practical use. So the answer to the title question seems to be "everything!".

Notes

1. See, e.g., Schwitzgebel 2019.
2. Antognazza 2020, p. 3.
3. Plato 1924, *Meno* 98b.
4. Plato 1924, *Meno* 97b–d.
5. Plato 1924, *Meno* 93c.
6. Plato 1921, *Theaetetus* 164a.
7. Plato 1921, *Theaetetus* 186d–e.
8. Plato 1921, *Theaetetus* 187b–c.
9. Plato 1921, *Theaetetus* 209e–210b.
10. Aristotle 1960, *Posterior Analytics*, Book I, XXXIII.
11. Aristotle 1960, *Posterior Analytics*, Book I, XXXIII.
12. Aristotle 1960, *Posterior Analytics*, Book I, XXXIII.
13. Thomas Aquinas 1920, *Summa Theologiae* II-II, Q.1, A.4.
14. Thomas Aquinas 1920, *Summa Theologiae* II-II, Q.1, A.5.
15. Mulligan 2014, pp. 386–387.
16. Cook Wilson 1926, p. 100.
17. Cook Wilson 1926, pp. 92–93.
18. Williams 1973, p. 146.
19. Williams 1973, p. 137.
20. Williams 1973, p. 140.
21. Williams 1973, p. 146.
22. Ramsey 1991, p. 44.
23. Ramsey 1991, p. 45.

References

- Antognazza, Maria Rosa 2020. ‘The distinction in kind between knowledge and belief’, *Proceedings of the Aristotelian Society* 120:3, pp. 277–308. <https://doi.org/10.1093/arisoc/aoaa013>
- Aristotle 1960. *Posterior Analytics*. *Topica* (Loeb Classical Library 391), trans. Hugh Tredennick & E. S. Forster, Cambridge, MA: Harvard University Press. https://doi.org/10.4159/DLCL.aristotle-posterior_analytics.1960
- Cook Wilson, John 1926. *Statement and Inference with Other Philosophical Papers*, vol. 1, ed. Arthur Spencer Loat Farquharson, Oxford: Clarendon Press. <https://archive.org/details/statementinference000owils>, accessed 5 March 2026.
- Gettier, Edmund L. 1963. ‘Is justified true belief knowledge?’, *Analysis* 23:6, pp. 121–123.
- Mulligan, Kevin 2014. ‘Knowledge first—A German folly?’, in Julien Dutant, Davide Fassio & Anne Meylan eds, *Liber Amicorum Pascal Engel*, Geneva: University of Geneva, pp. 380–400. <http://www.unige.ch/lettres/philo/publications/engel/liberamicorum>, accessed 5 March 2026.

- Plato 1921. *Theaetetus. Sophist* (Loeb Classical Library 123), trans. Harold North Fowler, Cambridge, MA: Harvard University Press. https://doi.org/10.4159/DLCL.plato_philosopher-theaetetus.1921
- Plato 1924. *Laches. Protagoras. Meno. Euthydemus* (Loeb Classical Library 165), trans. W.R.M. Lamb, Cambridge, MA: Harvard University Press. https://doi.org/10.4159/DLCL.plato_philosopher-meno.1924
- Ramsey, Frank P. 1991. *On Truth*, eds Nicholas Rescher & Ulrich Majer, Dordrecht: Kluwer Academic Publishers.
- Schwitzgebel, Eric 2019. 'Belief', in Edward N. Zalta ed., *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition). <https://plato.stanford.edu/archives/fall2019/entries/belief>, accessed 5 March 2026.
- Thomas Aquinas 1920. *The Summa Theologica of St. Thomas Aquinas*, Part II-II, 2nd rev. edn, trans. Fathers of the English Dominican province, London: Burns Oates & Washbourne. <https://www.gutenberg.org/ebooks/18755>, accessed 5 March 2026.
- Williams, Bernard 1973. 'Deciding to believe', in Bernard Williams, *Problems of the Self: Philosophical Papers 1956–1972*, Cambridge: Cambridge University Press, pp. 136–151. <https://doi.org/10.1017/CBO9780511621253.011>

On objections to Ramsey on belief

KEVIN MULLIGAN

1. Bourgeois, earth-bound, essence-blind, anti-scholastic pragmat(ic)ism

How did Frank Ramsey think of philosophy? How did he do philosophy? In what follows, I outline part of Ludwig Wittgenstein's answers to these two questions and some related views (1). I then consider some objections to Ramsey's views about belief (2–4), objections which, I suggest, illustrate Wittgenstein's view of Ramsey. I also sketch two views of belief which resemble Ramsey's view in some respects and avoid some of the objections mentioned.

In 1931, Wittgenstein gave a rather definite answer to the two questions:

Ramsey was a bourgeois [*bürgerlicher*] thinker. I.e. his thoughts had the aim of clearing up [*ordnen*] the affairs of some particular community. He did not reflect on the essence of the state—or at least he did not like doing so—but on how *this* state might reasonably be organized. The idea that this state might not be the only possible one partly disquieted him and partly bored him. He wanted to get down as quickly as possible to reflecting on the foundations—of *this* state. This was what he was good at & what really interested him; whereas real philosophical deliberation disquieted him until he put its result (if it had one) on one side as trivial.¹

A bourgeois thinker, Wittgenstein opines, is interested in particular, actual cases and not in the variety of similar cases, actual and merely possible, and their essence or nature. Ramsey's interest in a particular case is a practical interest and so a limited non-theoretical interest. (When Wittgenstein contrasts Beethoven's irony with Wagner's bourgeois irony he calls the former cosmic and the latter earthly [*irdisch*].)

Wittgenstein's verdict clearly recommends thinking about the essence of actual and merely possible cases in "real philosophical deliberation". This should not surprise us. The *Tractatus* is steeped in essentialism. He there inquires into the essence of logic, the essence of language, the essence of thought, the essence of the world and the essences or natures of much, much else.² When he later comes to appreciate that the variety of, for example, understanding consists of numerous relations of merely partial similarity,³ he sometimes opposes such a view to essentialism. But in his later writings he clearly continues to think that real philosophical deliberation is inseparable from an interest in essence. The fact that his understanding of essence differs radically from traditional understandings thereof—essence is now specified by grammar—does not by itself make him an anti-essentialist.

Aspects of Wittgenstein's verdict are confirmed by Ramsey himself for whom "philosophy must be of some use".⁴ What is directly and immediately of use is useful for our relations to contingent, actual particulars. Similarly, Ramsey's condemnation of scholasticism is perhaps not merely a condemnation of mistaking what is vague for something precise but also of the origin of this sort of mistake in the pursuit of exact essences, such as logical categories:

The chief danger to our philosophy, apart from laziness and woolliness, is *scholasticism*, the essence of which is treating what is vague as if it were precise and trying to fit it into an exact logical category. A typical piece of scholasticism is Wittgenstein's view that all our everyday propositions are completely in order and that it is impossible to think illogically.⁵

Ramsey thinks that scholasticism once led him to worry about a question of essence: he mentions that he used to worry about the nature of philosophy "through excessive scholasticism".⁶ Wittgenstein, on the other hand, thought highly of the "strong scholastic feeling" which had "occasioned" his "best discoveries",⁷ which presumably include his discoveries of various essential differences. On the other hand, it might be thought that a philosopher who, like Ramsey in the last quoted passage, refers to the essence of scholasticism cannot really be an enemy of the essentialist point of view. But Ramsey's use of "essence" here and elsewhere is, I suggest, a loose or emphatic way of talking, also to be found in Bertrand Russell, who was no friend of essentialism.

John Maynard Keynes' description of Ramsey's view of and way of doing philosophy also confirms some aspects of Wittgenstein's verdict, provided certain not implausible assumptions are made. Ramsey, he says, is "attempting to distinguish a 'human' logic from formal logic on the one hand and descriptive psychology on the other".⁸ *Descriptive psychology* was the term used by Brentano and his heirs to refer to their favoured essentialist, philosophy of mind, an enterprise they distinguished from causal, genetic, empirical psychology. Edmund Husserl's increasingly baroque essentialism about the mind required, he thought, a new term: *phenomenology*.⁹ Did Keynes have in mind something like this understanding of descriptive psychology? Perhaps. Keynes' attribution to Ramsey of a distinction between formal logic and a "human" logic is not controversial. The difference between the two may reflect a distinction between essentialism and anti-essentialism about logic. In Husserl's famous criticisms of misunderstandings of formal logic in 1900, he not only rejects *psychologism*, the view that formal logic is a matter of psychology, but also *anthropologism*, the view that formal logic is a human logic. Many versions of these two mistakes rely on forms of anti-essentialism. But, Husserl argues, the truths of formal logic are necessary truths grounded in the essences, the exact essences of logical categories. The objections to psychologistic and anthropologistic views of the nature of formal logic do not, of course, by themselves entail that there cannot be a psychology or an anthropology of formal logic.¹⁰

William James' comments on the relations between pragmatism and early Cambridge philosophy point to an undeniable dimension of the scholasticism which Ramsey thought he had grown out of.

If pragmatism has any merit, it is to have rid the world of such diseased-wucherungen [excrescences] as the Moore–Russell–Meinong epistemology.¹¹

James, who had read even more descriptive psychology than Russell, G. E. Moore and George Stout, refers to "the unspeakable [Alexius] Meinong and his English pals".¹² His point, I think, is that the philosophies of Meinong/Russell/Moore are not only useless, they are unhealthy, enemies of life, to use the expression beloved of the philosophy of life. Pragmatism and *Lebensphilosophie* share a complete lack of interest in essentialism, where essences are taken to be the unalterable grounds of

metaphysical necessity.¹³ This is a point made in early, realist phenomenology:

The basic error of the way of thinking [*Denkart*] of all positivism and pragmatism is to overlook the entire realm of [...] essential truths.¹⁴

The pragmatist is someone who is lost [*ein Verirrter*] in the contingent reality of the world. He does not see the essences [*Wesenheiten*], the ideas [*Ideen*] of things.¹⁵

[P]ragmaticist philosophy signifies for us an attempt [...] to reduce all knowledge to useful knowledge [*Arbeitswissen*, work knowledge]—to display the possible modification of the world in the direction of the goals of our will as the only sense and value of knowledge.¹⁶

Max Scheler's rejection of pragmatism as essence-blind, like Wittgenstein's rejection of Ramsey's way of doing philosophy as essence-blind, is by no means an *en bloc* rejection of pragmatism. Pragmatism is not, Scheler thinks, entirely wrong. Its account of the relation between desire, perception and action is almost correct. It is only pragmatism's misunderstandings of meaning, truth and essence which have led to the "pragmaticist swamping [*Versumpfung*] of all philosophy".¹⁷ Pragmatism thus made some headway within early phenomenology but less than it was to make in early Cambridge philosophy.¹⁸

2. Belief and descriptive psychology:

Degrees vs degreelessness

A philosophy of belief which begins to do justice to its variety would be a philosophy of believing that something, believing something to be something, belief in something or someone and believing someone. It would be a philosophy of belief as something which can endure, perhaps because it is a disposition, and of something which occurs ("occurrent belief"). It would adjudicate between the view that belief may be both positive and negative (disbelief, as opposed to belief that not-*p*), between the view that it is a reaction or response to something else and the view that it is no such thing, and between the view that it or some types of belief are affective or intellectual in nature. It would display the relations

between belief, on the one hand, and acceptance, conviction, certainty and judging. It would tell us what belief's objects are. Do believing that p and believing x to be F range over past, present and future, over the contingent and non-contingent? Do these forms of belief also have so-called formal objects? Do they "aim" at, for example, truth or probability? Some of these questions are, of course, such that they cannot be answered without also determining the different roles of belief. (Is belief-that a component of knowledge-that? Are beliefs or believers more or less rational or reasonable? What makes them enjoy these attributes?) Some essentialist philosophers think that their answers to these questions flow wholly or in part from the essence of belief.¹⁹

Belief, Ramsey tells us, is a feeling and comes in degrees. Belief here and in what follows is belief-that. What exactly does Ramsey mean and is he right? A prior question is: Does it matter? This question is justified or suggested by Ramsey's repeated display of indifference about what belief really is. The question about the essence of belief, it may seem, is of less interest to him than the task of understanding the relations between belief, its causal properties, betting and subjective probability.²⁰

One claim he does make about belief which arguably *is* a claim about its nature is a claim which follows from a more general claim he makes about thoughts and the two kinds of character they display. In 1921, he says:

[T]here are *two kinds of characters* which they can have which are of great importance. A thought may be a belief, a doubt, or an assumption, etc. These I call *pistic characters*. I think that *they consist in the presence or absence of feelings*, such as Mr. Russell calls belief feelings, in connection with the thought; but their analysis is not very relevant to the matter in hand and I mention them to distinguish them clearly from the other fundamental kind of characters of thoughts, which I call *referential characters* or references.²¹

This is a version of the distinction famously made by Russell in 1919 between attitudes and propositions, where "proposition" refers to a concrete linguistic or mental entity rather than a so-called singular proposition, objective or state of affairs. In 1927, Ramsey calls the two characters of belief mental factors:

The mental factors of such a belief I take to be words, spoken aloud or to oneself or merely imagined, connected together and accompanied by a feel-

ing or feelings of belief or disbelief, *related to them in a way I do not propose to discuss*.²²

To say that feeling belief towards a sentence expresses such an attitude, is to say that it has certain causal properties which vary with the attitude, i.e. with which possibilities are knocked out and which, so to speak, are still left in.²³

An ancestor of Russell's 1919 distinction is the distinction between modes or attitudes and contents (understood as concrete, mental tokens) in descriptive psychology. Modes are there also called "act-characters" and "act-qualities" and belief and judgment, like assuming or supposing, it is argued, are intellectual and not any sort of affective modes, qualities or characters. Ramsey is familiar with this sort of view:

I speak throughout as if the differences between belief, disbelief, and mere consideration lay in the presence or absence of 'feelings'; but any other word may be substituted for 'feeling' which the reader prefers, e.g. 'specific quality' or 'act of assertion' and 'act of denial'.²⁴

Ramsey's distinction between two types of character resembles the earlier distinctions made by Husserl, Meinong and Russell in one more respect. Ramsey's two characters are wholly disjoint, as are Russell's attitude and proposition and the act-qualities and token contents of the descriptive psychologists.

In the light of this framework, according to which those mental events which are thoughts display two distinct characters, what should we think of Ramsey's claims that belief comes in degrees and is a feeling? On many views of feelings, they (essentially) display degrees. On these views, then, belief would come in degrees because it is a feeling. How, then, does Ramsey understand his two claims and is he right?

The claim that belief comes in degrees might be taken to amount simply to the claim that it is possible for someone to believe that p more than that not- p or for one person to believe that p more than another person. Then belief would resemble regret and sadness. But although we may often observe how someone becomes more and more angry, less and less sad, we do not observe a growth in someone's belief that p even while increasingly good reasons for believing that p are being poured into his ear. The view that belief has degrees has often been rejected. Believing

and disbelieving “do not display degrees”, writes Scheler in 1915–1916,²⁵ as John Newman in the previous century had asserted that assent is degreeless: “We might as well talk of degrees of truth as of degrees of assent”.²⁶ Alan White is equally emphatic:

There are various reasons for denying that belief has degrees. First, belief cannot be increased or decreased in amount nor become more or greater, less or smaller, in the way that, for example, either heat, weight, length and pressure or anger, joy and pride can. I do not believe that *p*, e.g. that London is the capital of the U.K., more than I believe that *q*, e.g. that Lima is the capital of Peru.²⁷

In Ramsey’s claims that belief comes in degrees, that there is “partial belief”, the emphasis lies rather on the existence of beliefs that are certain or less certain, firm or less firm, strong or weak.²⁸ But why should the variability of the semantic value of the adverbial qualification of a verb entail that the semantic value of the verb comes in degrees? If Maria observes Sam attentively or with interest or with difficulty, the fact that her attentiveness, interest or difficulty may grow or diminish while she continues to observe Sam does not entail that she may observe him more or less, although she may of course observe more or less of him. The distinction between the semantic values of the verb “believe” and of its adverbs (“with certainty”, “firmly”, “confidently”) corresponds to a real difference. As White points out,

What has mislead [*sic*] philosophers into the view that belief has degrees is the fact that a belief can vary in various respects which do admit of degrees. A belief can be more or less faint, firm, obstinate or unwavering, obsessive or tentative, confident, hesitant or unquestioning; it can be half-hearted or whole-hearted. But these are not degrees of belief; they are degrees of conviction, confidence, or tenacity with which the belief is held.²⁹

Ramsey seems to be not at all interested in the distinction between mental or psychological states or acts which do and those which do not come in degrees (although his reference to the view that belief is intellectual perhaps shows that he was aware of the distinction). Nor does he pay any attention to the possibility that the certainty of a belief differs from its firmness and both again from its strength.

By way of illustration, consider the following way of understanding the strength or firmness of beliefs:

[O]ne may strongly or firmly believe that p ..., but this does not indicate a degree of belief. It signifies the strength or firmness with which one cleaves to the belief one has. It is the ease or difficulty of shaking the belief in question, and not the belief itself, that has degrees.³⁰

The strength or firmness of a belief, like what Scheler calls the depth of a belief (*Glaubenstiefe*), plays a central role in an account of non-critical but fundamental beliefs famously set out by Wittgenstein towards the end of his life and even earlier by José Ortega y Gasset, who referred to such beliefs as “basic beliefs”. In such an account the firmest beliefs underlie but do not stand in any rational relations to the less firm and more or less critical beliefs on which they are based. They are precisely the beliefs some people find most difficult to shake (“I have two hands”, “My nation-state is legitimate”, “Men know best”, “God is in his heaven”). The firmness of the belief that the ground I stand on is firm is not only a typical example of doxastic firmness but also one of the central metaphors employed to describe doxastic firmness in general.

As a second illustration, consider the familiar sort of view to which Ramsey aims to provide an alternative, according to which doxastic certainty or doxastic strength is typically a reaction or response to (the apparent) goodness or badness or solidity of reasons to believe or disbelieve or to (apparent) relations of probability between states of affairs. On such a view, neither doxastic strength nor doxastic certainty is identical to doxastic firmness, for doxastic firmness is not any sort of rational predicate of beliefs.

A third illustration is the view that doxastic certainty differs essentially from both doxastic strength and firmness. Certainty *is* a feeling but doxastic strength and firmness are not. This view is best introduced in the course of considering Ramsey’s claim that belief is a feeling.

3. Is belief a feeling?

Is belief a feeling? What is a feeling? To be a feeling is to enjoy positive or negative valence. An important counter-example to this claim is sur-

prise, the neutral emotion. But let us simply ignore this. Feelings then seem to come in at least the following main kinds. There are feelings which are unrelated to what Ramsey calls referential characters or to what others call intentionality, for example, localized bodily pain or pleasure and moods (bliss, despair, *Angst*). There are feelings which are (perhaps essentially) related to referential characters, which “colour” (Wittgenstein) these or depend unilaterally on them: Wittgenstein’s admiration of Ramsey’s soul and dislike of his *Geist*, for example. Such feelings depend not only on referential characters but also on pistic characters, in particular beliefs, in a suitably wide sense of the term. Thus Maria’s admiration of Sam is based on her memory-beliefs about him. Hans’ admiration for Sam on his perceptual beliefs about him. Maria’s admiration of the British Empire on beliefs derived from reading history books. And so on.

Suppose now that belief is a feeling. If a belief is a feeling, it has to be—in Ramsey’s peculiar phrase—“directed to” referential characters. It is thus not any sort of bodily feeling or mood. Such a directed feeling, like all feelings other than bodily feelings or moods, must be based on a belief. But a belief, Ramsey says, *is* a feeling. Ramsey’s view, then, on one plausible account of feelings leads to a regress, a vicious mental or psychological regress if the feelings and beliefs are occurrents.

The view that feelings other than moods and bodily feelings must be based on beliefs is a version of what is sometimes called a cognitive view of emotions. On a slightly more sophisticated version of the view, feelings must be based either on beliefs, in a narrow sense of the word, or on memories, perceptions or expectations where these are not forms of belief. Cognitive accounts of emotions were familiar in Cambridge. Moore had endorsed such an account in 1910 and J. M. E. McTaggart followed his lead.

Richard Braithwaite rejected Ramsey’s view that beliefs are feelings: “I cannot accept the feeling as part of the essence of belief. [...] I am convinced that belief and feeling are different, and that the latter is not part of the former”.³¹

But there are at least two views which are close relatives of Ramsey’s view of belief as a feeling. First, belief is inseparable from certainty, indeed from some particular degree of certainty. Certainty, like uncertainty, is something we may feel, and has a characteristic expression. Certainty-

that and uncertainty-whether are indeed feelings. Belief, then, on this view, is inseparable from a feeling which comes in degrees, although it itself does not come in degrees. Scheler puts the view as follows:

Believing something and disbelieving [*nichtglauben*] something have nothing to do with the so-called feeling of certainty and the feelings which accompany them. For certainty stands opposed only to uncertainty (and both certainty and uncertainty are opposed to the oscillation between these and its accompanying feeling, the feeling of doubt) and believing to the equally positive disbelieving, which in turn can be certain or uncertain. Furthermore, *these feelings exhibit degrees, quite unlike believing and not-believing*.³²

White, too, was to note the affective nature of certainty:

Certainty, but not knowledge, is something one can feel, something one can induce in oneself or another, something one can discover in oneself introspectively or in another behaviouristically. We can ask *how* somebody knows that *p*, but not how he is certain that *p*; and, conversely, *why* he is certain, but not why he knows.³³

The strength and firmness of beliefs are, in contrast, not affective states. White also claims that certainty “is not the highest degree of belief, but of conviction, confidence or lack of doubt”.³⁴ A somewhat different view is that conviction is belief which is certain to the highest or to a high degree. (One may feel convinced that *p*). On both views, the relevant type of certainty is subjective, affective certainty. Certainty may also be objective, as when we claim that it is (un)certain (whether) that *p*.

Second, a very close relative of belief is suspecting. (Suspecting, surmising and conjecturing are the three English translations of German *vermuten*, the attitude at the heart of Meinong’s 1915 account of [what he does not call] subjective and objective probability.) In a pioneering but all too brief account of suspecting something, White points out that it comes in degrees—I may suspect that *p* more than that *q*—and argues that to suspect that *p* is to be inclined to believe that *p*.³⁵ Since it is an inclination, it may, he also points out, be felt, unlike belief. It is therefore a phenomenon which involves belief and has an affective dimension. (Why has the philosophy of suspecting-that been so neglected, even by essentialists? Perhaps because of the negative [White says “untoward”] con-

notations of the noun suspicion and of the verb when combined with a nominal complement, a traitor is suspected.)

4. Belief about the contingent vs belief about the non-contingent

Suppose that Ramsey and many others are right to say that knowledge is a suitably qualified type of belief. Suppose, too, that there is knowledge of contingent facts and of non-contingent facts. Then there are beliefs about what is contingent and about what is not contingent. Some philosophers think there are no beliefs about what is not contingent, that belief ranges only over the contingent. Ramsey argues that tautologies and contradictions are not and cannot be believed. Unfortunately, he says very little about this intriguing view.

One might think that his view that belief is intimately bound up with possible action entails that belief must bear only on contingent matters of fact. Whether or not this is the case, Ramsey briefly argues from his already quoted view that

To say that feeling belief towards a sentence expresses such an attitude, is to say that it has certain causal properties which vary with the attitude, i.e. with which possibilities are knocked out and which, so to speak, are still left in.³⁶

to the claim that contradictions and tautologies are never or cannot be doxastically qualified:

If we apply the logical operations to atomic sentences in an indiscriminate manner, we shall sometimes obtain composite sentences which express no attitude of belief. Thus “p or not-p” excludes no possibility and so expresses no attitude of belief at all [...]. Similarly, “p and not-p” excludes every possibility and expresses no possible attitude: it is called a contradiction.³⁷

Is Ramsey thereby committed to the view that his account of belief applies to all types of belief? In order to answer this question, it would be necessary to know whether he allows for beliefs about what is non-contingent but non-formal. For example, beliefs such as that orange lies

between red and yellow or philosophical beliefs such as the belief that to be a person is to be free or that events contain their times. Ramsey's thesis that certain types of sentence express no possible attitude is itself a philosophical but non-formal claim, since an attitude such as belief is not any sort of formal or logical relation. (So, too, is the thesis that Ramsey's formulation apparently disallows that the attitude or pistic character of assuming or supposing can qualify contradictions, can be "directed" to these.) More generally, we do not know what late Ramsey thought about Wittgenstein's view in the *Tractatus* about the relations between what have been called formal and material necessities or necessary connections.

Braithwaite did attempt to answer some of these questions. He develops an account of belief which in several but not all respects resembles that given by Ramsey:

I come to my proposed analysis of belief. And first I wish to consider belief in contingent propositions that are not known directly. I am excluding, that is, beliefs in logically necessary propositions (including the propositions of formal logic and mathematics and instances of them) and beliefs in propositions which are directly known in experience. My theory of belief, therefore, applies to most of the propositions we use in ordinary life but not to some of those which philosophers discuss.³⁸

Braithwaite's analysis of belief, then, is not supposed to apply to all types of belief and so, one might think, cannot capture the essence of belief. Braithwaite also rejects implicitly the view that all knowledge is a type of belief:

Besides contingent propositions directly known, to which I have alluded, I have explicitly omitted from my analysis beliefs in logically necessary propositions. Direct knowledge of these I am pretty clear is something into which belief does not enter: my seeing that $2 + 2 = 4$ is not a believing. My attitude towards such a proposition is more like that towards a proposition about a sense-datum: I have only to understand the words to know whether the proposition is true or false.³⁹

To see or know that $2 + 2 = 4$ does not involve any belief. About knowledge of this type, Braithwaite adopts the Oxonian knowledge-first view.

Braithwaite's view that knowledge of the non-contingent is not any sort of belief is also part of Scheler's view about the essence of belief and its formal object or correlate:

[In] believing and disbelieving, essentially what is given is the being real or irreal of a state of affairs (also of course its becoming, its having been real or timeless being real).⁴⁰

To be real, on this view, is to be causally efficacious and belief must be "about" reality just as (immediate) memory must be "about" the past and knowledge-that "about" facts or factuality. Is knowledge of the non-contingent then merely intuitive knowledge? Is there no conceptually articulated knowledge of the non-contingent? The descriptive psychologists and phenomenologists think that there is such knowledge. Judging, says Scheler, unlike believing, takes both what is real and what is not real as its province:

Believing and disbelieving may not be confused with judging, positive and negative, critical and non-critical [*schlicht*]. Positive and negative judging is also to be found in judgments which have *nothing* to do with reality.⁴¹

The contingent and the non-contingent are essentially *judgeable* (Frege's term). The former but not the latter is essentially *believable*.

5. Conclusions

This brief sketch of what a philosophy of belief might be, of objections to Ramsey on belief and of two possible modifications of his view illustrates, I hope, Wittgenstein's description of Ramsey's approach to philosophy. For many of the questions touched on are just the questions an essentialist would take seriously. It also, I hope, throws some light on the relations between Ramsey, Wittgenstein and pragmatism, one of the red threads running through Sahlin's rich and rewarding guide to Ramsey⁴² and, more generally, on the relations between pragmatism, early phenomenology and early Cambridge analytic philosophy.

In phenomenology and then in analytic philosophy, couples such as conjecture or belief & *probability*, judgment or belief & *truth*, belief or

resistance & *reality*, acceptance & *existence*, memory & the *past*, emotions & *value* received a great deal of attention. Three of the more important options were naïve realism (attitudes or modes owe their correctness or normative status to the exemplification of their correlates or formal objects), buck-passing (the exemplification of such correlates reduces to or is grounded in the normative features of attitudes) and nihilism. One aspect of Ramsey's originality is that he thought that truth was redundant, and he was a nihilist about probability relations, value and pistic characters understood in mentalist terms.

I have not attempted to consider the relation between the views about belief mentioned and Ramsey's account of the relations between beliefs, bets and desires. Nor have I speculated about Ramsey's likely reaction to Wittgenstein's description of his approach to philosophy. I suspect that it would refer to the relation between essences, essentialism and his view that there is no absolute difference corresponding to the distinctions we make between universals and particulars. Possibly also to whistling.

Acknowledgements

Thank you, Monika Gruber, Fraser MacBride, Anna-Sofia Maurin and Peter Simons, for discussion, suggestions and objections.

Notes

1. Wittgenstein 1984, p. 473; my translation.
2. Cf. Mulligan 2024.
3. A view which would not have surprised Ramsey, who refers to "the vagueness of the whole idea of understanding, the reference it requires to a multitude of performances" (Ramsey 1990, p. 2).
4. Ramsey 1990, p. 1.
5. Ramsey 1990, p. 7, 'Philosophy'.
6. Ramsey 1990, p. 1.
7. Wittgenstein 1979, 10.11.14.
8. Keynes 1963.
9. In 1912 Husserl notes that the friends of descriptive psychology and phenomenology "are regularly attacked as scholastics" (Husserl 1981, p. 20).
10. On Ramsey on human psychology and human logic, cf. Suppes 2006.
11. Skrupskelis & Berkeley 2003, p. 5.
12. Quoted in Perry 1935, p. 485.

13. On Charles Sanders Peirce's essentialism about natural necessity, cf. Howat 2024.
14. Scheler 1982, p. 409; my translation.
15. Scheler 1980, p. 465; my translation.
16. Scheler 1980, p. 211; my translation.
17. Scheler 1966, pp. 394–395.
18. In his 1890 review of Ernst Schröder, Husserl is astonished to learn from Schröder that, according to Peirce, “judgments are explained as special cases of habits of thought [*Denkgewohnheiten*]”. This is, he says, an “unbelievable [*aben-teuerliche*] theory”. Scheler claims, perhaps correctly, that many of the views put forward in Husserl's *Logical Investigations* (1900–1901) are criticisms of pragmatist views.
19. The first major 20th-century treatments of belief are phenomenological (Reiner 1934; Scheler 1957), in the traditions influenced by Wittgenstein, Ramsey or Austin (White 1982; White 1993; Hacker 2004) and otherwise analytic (Braithwaite 1933; Price 1969; de Sousa 1971; Cohen 1992).
20. On Ramsey on belief, cf. Sahlin 1990; Gruber 2022.
21. Ramsey 1991, p. 91, ‘The nature of propositions’; my italics.
22. Ramsey 1990, p. 40, ‘Facts and propositions’; my italics.
23. Ramsey 1990, p. 46, ‘Facts and propositions’.
24. Ramsey 1990, p. 40, ‘Facts and propositions’.
25. Scheler 1957, p. 241.
26. Newman 1887, p. 174. Newman rejects John Locke's claim that there are degrees of assent. He identifies degrees of assent with conditional assent and denies their existence. Like later philosophers of belief influenced by Wittgenstein or Austin, he attaches great weight to what we say, do not say or would not say: “we never say that we give [a proposition] a degree of assent” (p. 174). Friends of the degreelessness of belief include Scheler 1957; White 1993; Hacker 2004; Mulligan 2013; Moon 2017; Jackson 2022.
27. White 1993, pp. 81–82.
28. Ramsey also refers to taking for granted, confidence and conviction.
29. White 1993, p. 82.
30. Hacker 2004, p. 192.
31. Braithwaite 1933, p. 142.
32. Scheler 1957, p. 241; my translation and italics.
33. White 1982, p. 76.
34. White 1993, p. 82.
35. White 1993.
36. Ramsey 1990, p. 46, ‘Facts and propositions’.
37. Ramsey 1990, p. 47, ‘Facts and propositions’.
38. Braithwaite 1933, p. 132.
39. Braithwaite 1933, p. 142.

40. Scheler 1957, p. 241; my translation.

41. Scheler 1957, p. 242; my translation.

42. Sahlin 1990.

References

- Braithwaite, Richard Bevan 1933. 'The nature of believing', *Proceedings of the Aristotelian Society* 33:1, pp. 129–146.
- Cohen, L. Jonathan 1992. *An Essay on Belief and Acceptance*, Oxford: Clarendon Press.
- de Sousa, Ronald 1971. 'How to give a piece of your mind: The logic of belief and assent', *Review of Metaphysics* 25, pp. 52–79.
- Gruber, Monika 2022. 'Ramsey's theory of belief', *European Journal of Pragmatism and American Philosophy* 14:2.
- Hacker, Peter 2004. 'On the ontology of belief', in Mark Siebel & Mark Textor eds, *Semantik und Ontologie: Beiträge zur philosophischen Forschung*, Frankfurt: Ontos, pp. 185–222.
- Howat, Andrew 2024. 'Pragmatism', in Kathrin Koslicki & Michael J. Raven eds, *The Routledge Handbook of Essence in Philosophy*, London & New York: Routledge, pp. 53–66.
- Husserl, Edmund 1981 [1912]. *Philosophie als strenge Wissenschaft*, Klostermann, Frankfurt.
- Jackson, Elizabeth 2022. 'Why credences are not beliefs', *Australasian Journal of Philosophy* 100:4, pp. 360–370.
- Keynes, John Maynard 1963 [1931]. 'F. P. Ramsey 1903–1930', in John Maynard Keynes, *Essays in Biography*, New York: Norton, pp. 239–254.
- Moon, Andrew 2017. 'Belief does not come in degrees', *Canadian Journal of Philosophy* 47:6, pp. 760–778.
- Mulligan, Kevin 2013. 'Acceptance, acknowledgment, affirmation, agreement, assertion, belief, certainty, conviction, denial, judgment, refusal & rejection', in Mark Textor ed., *Judgement and Truth in Early Analytic Philosophy and Phenomenology*, London: Palgrave Macmillan, pp. 97–137.
- Mulligan, Kevin 2024. 'Essence and modality: From Husserl to Wittgenstein', in Jimmy Plourde & Mathieu Marion eds, *Wittgenstein's Pre-Tractatus Writings: Interpretation and Reappraisal*, London: Palgrave Macmillan, pp. 145–163.
- Newman, John H. 1887. *An Essay in Aid of a Grammar of Assent*, London: Longmans, Green and Co.
- Perry, Ralph B. 1935. *The Thought and Character of William James*, vol. 2, *Philosophy and Psychology*, Boston: Little, Brown and Company.
- Price, Henry H. 1969. *Belief*, London: George Allen & Unwin.
- Ramsey, Frank P. 1990. *F. P. Ramsey: Philosophical Papers*, ed. David Hugh Mellor, Cambridge: Cambridge University Press.
- Ramsey, Frank P. 1991. 'The nature of propositions', in Frank P. Ramsey, *Notes on*

- Philosophy, Probability and Mathematics*, ed. Maria Carla Galavotti, Naples: Bibliopolis, pp. 86–103.
- Reiner, Hans 1934. *Das Phänomen des Glaubens: Dargestellt im Hinblick auf das Problem seines metaphysischen Gehaltes*, Halle: Niemeyer.
- Sahlin, Nils-Eric 1990. *The Philosophy of F. P. Ramsey*, Cambridge: Cambridge University Press.
- Scheler, Max 1957 [1933]. ‘Glauben als Aktart’, in *Schriften aus dem Nachlass*, Bd. 1, Gesammelte Werke 10, Bern: Francke, pp. 241–250.
- Scheler, Max 1966 [1913–1916]. *Der Formalismus in der Ethik und die materiale Wertethik: Neuer Versuch der Grundlegung eines Ethischen Personalismus*, Gesammelte Werke 2, Bern: Francke.
- Scheler, Max 1980 [1926]. *Die Wissensformen und die Gesellschaft*, Gesammelte Werke 8 [1960], Bern: Francke.
- Scheler, Max 1982. *Politisch-pädagogische Schriften*, Gesammelte Werke 4, Bern: Francke.
- Skrupskelis, Ignas K. & Elizabeth M. Berkeley 2003. *The Correspondence of William James*, vol. 11, Charlottesville & London: University of Virginia Press.
- Suppes, Patrick 2006. ‘Ramsey’s psychological theory of belief’, in Maria Carla Galavotti ed., *Cambridge and Vienna: Frank P. Ramsey and the Vienna Circle* (Vienna Circle Institute Yearbook 12), Dordrecht: Springer, pp. 35–53.
- White, Alan R. 1982. *The Nature of Knowledge*, New Jersey: Rowman & Littlefield.
- White, Alan R. 1993. ‘Suspicion’, in John V. Canfield & Stuart G. Shanker eds., *Wittgenstein’s Intentions*, New York: Garland, pp. 81–86.
- Wittgenstein, Ludwig 1974. *Tractatus Logico-Philosophicus*, London: Routledge & Paul Kegan.
- Wittgenstein, Ludwig 1979. *Notebooks 1914–1916*, eds G. E. M Anscombe & G. H. von Wright, 2nd edn, Oxford: Blackwell.
- Wittgenstein, Ludwig 1984. *Vermischte Bemerkungen* (Werkausgabe Band 8), Frankfurt: Suhrkamp.

Can there be probabilistic knowledge?

PASCAL ENGEL

1. Two dogmas of contemporary epistemology

There are two dogmas of contemporary epistemology about beliefs and knowledge:

- A. Knowledge is a kind of higher-grade belief.
- B. Propositional attitudes are relations to propositions.

The first dogma is associated to the so-called justified true belief analysis of knowledge (JTB) which has prevailed in contemporary epistemology: knowledge is true belief, plus justification, often understood as a condition preventing the risk of having a true belief by chance (an “anti-luck” condition).¹ Knowledge is a kind of belief, but with warrant. This definition also involves a truth condition: knowledge entails truth. This means that beliefs which fall short of being true and which are only probable and even those which are held for good reasons can fail to be knowledge.

The second dogma is the *relational* conception of intentional attitudes: attitudes, epistemic or otherwise, such as beliefs, desires, hopes, etc., are relations between subjects and propositions, which are their contents. A single attitude, such as belief, can be related to different contents or propositions, and a single content or proposition can be the object of different attitudes. Thus, to believe that p is to stand in a three-place relation between an individual s , the relation of believing B , and its object, a proposition p , hence has the form sBp .

Both dogmas have been criticized. Some philosophers have renewed the traditional thesis that belief and knowledge are as different as chalk

and cheese.² Beliefs cannot be knowledge, and they are not a kind of cognition of things. Knowledge and belief are different kinds of states. Knowledge cannot be analysed in terms of belief, and it is primitive. It only can get us to the nature of things. On this “knowledge-first” view,³ knowledge takes precedence over belief, and epistemological notions such as evidence, justification or reliability have to be analysed in terms of knowledge rather than the contrary. Knowledge is by definition factive; it entails truth.

Some philosophers have dissented from the widespread view that attitudes are relations to propositions.⁴ Against the relational view it has been argued that attitudes are not relations between subjects and contents, but that they involve only subjects who are in certain complex contentful states. There are no attitudes as ordinarily construed: there are only kinds of contents, which determine in which kind of attitude a subject is at a given moment. Thus, for a subject *a* to perceive that *p* is for *a* to be in relation to a content, which has some phenomenal properties specific to perception. There is no need to specify a perceptual relation such as “perceiving”. For beliefs, the non-relational view would first identify the content, then treat it as instantiated by a subject. On the view that beliefs-contents are made of concepts the idea would be, for instance, for someone *a* who believes that snow is falling, *a -snow falling instantiation*. For desires, the content would be, say: *snow falling is good* is instantiated in *a*.⁵ Indeed the analysis has to be more complex, and it is unclear that the idea of instantiation completely gets rid of the notion of relation, but the point of the non-relational view is simple: it is the content in which the subject finds himself which specifies the attitude, not the attitude which orients towards the content.

Let us call *revisionist* the “knowledge-first” view and the non-relational view of the attitudes, since they entail strong revisions of our contemporary basic thinking about knowledge and belief. In fact, they seem to conflict: if, according to the first thesis, knowledge is a factive attitude involving a relation to a proposition, and if according to the second thesis knowledge is not a relation to *propositions*, knowledge cannot be an *attitude*. Can one be a revisionist about both? This is not obvious. I want here to consider a view which, on the face of it, seems to be doubly revisionist, in accepting both a “knowledge-first” conception and a non-relational conception of belief. Bayesian conceptions of belief are likely

candidates for non-relational view: they say that beliefs are not relations to propositions, but to degrees of belief or credences.⁶

The main contemporary inspiration for the Bayesian view of belief comes from Frank Ramsey.⁷ After recalling it, I shall examine a version of the Bayesian view which licenses the idea that knowledge can be probable. I would like to suggest that this is not a coherent combination.

2. Ramsey on knowledge and belief

Prima facie Ramsey seems to have accepted the two dogmas—that knowledge is but a higher-grade form of belief and that beliefs are attitudes to propositions. Let us start with his view of belief and knowledge.

Ramsey seems to subscribe to the relational conception of belief, which he takes from Russell: belief is a multiple relation between the mind and the constituents of a judgment.⁸ Thus the ascription “Othello believes that Desdemona loves Cassio” has the form:

Bel (Othello, Desdemona, love, Cassio).

Ramsey agrees, at least *prima facie*, with Bertrand Russell’s view that the contents of beliefs are true when their constituents correspond to the facts. He says that a belief ascription

consists in the holding of some relation or relations between the believer and the referents of the belief, that is, a relation or relations between one’s ‘mind, or [one’s] present mental state, or words or images in [one’s] mind’ and ‘the objective factor or factors’.⁹

But this *facie* is only *prima*, since Ramsey doubted, like Wittgenstein, the correctness of the multiple relation theory and considered that the correspondence theory of truth is wrong, opting for a version of deflationism about truth.¹⁰ He took even more distance from the relational view when he adopted the view that belief has degrees, which are degrees of subjective probability. Partial beliefs are very different from relational beliefs to propositions. If A believes that *p* to a certain degree, the degree is neither a modifier of the attitude of belief nor a part of the content of a proposition. The form is not

A believes that it is true to degree n that p
 but

A believes to degree n that p .

This means that the degree is *essentially* part of the attitude. In other words, partial belief is not a kind of full belief about the probability of an event. As Richard Holton comments, discussing the Bayesian conception of belief:

Let us be clear on what the credence picture involves. It doesn't simply involve the idea that we can entertain the thought that a certain outcome has a certain chance: that there is a 0.5 chance that a fair coin will come down heads. For that thought can be the content of an all-out belief. If that were all that was involved, we could simply add a few all-out beliefs with a probabilistic content to our stock of beliefs, and there can be no doubt that that would be a very useful thing to do. The credence picture is very different. It is that our beliefs are *essentially* probabilistic. This is because probability is not in the content of a few beliefs, but in the attitude of belief itself. It is not that we believe that something is, say, 0.5, or 0.7 probable. It is that we have a 0.5 or 0.7 credence in it happening. Every belief is thus a probabilistic belief (even if the credence happens to be 1 or 0).¹¹

There are good reasons to attribute to Ramsey such a view: there are not two kinds of beliefs, partial on the one hand, and full or categorical belief on the other hand, but only one, for all beliefs are partial, although some can be of degree 1 or 0, in which case these are said to be true or false. The result is radical: if beliefs are identified with degrees of belief, they are not relations to propositions which could be their contents, and they are not propositional attitudes in the Russellian sense. They would be relations to propositions only when they are "full" or "categorical".¹²

If this Bayesian view of belief can be attributed to Ramsey, a psychological consequence and an epistemological consequence follow. The psychological consequence is that beliefs are not propositional attitudes. Although this view is different from the neo-Brentanian non-relational view of the content of belief, it agrees with it in rejecting the relational conception. Nevertheless, Ramsey still speaks of propositions and truth.

But his revisionist leanings are quite clear, especially when it comes to general propositions, which according to him cannot be true at all.¹³

The epistemological consequence concerns the nature of knowledge. It seems that Ramsey is not a revisionist relative to the nature of knowledge. In a note of 1929, 'Knowledge', he says:

I have always said that a belief was knowledge if it was (i) true, (ii) certain, (iii) obtained by a reliable process.¹⁴

As Nils-Eric Sahlin says, commenting on this definition, although it looks superficially like the JTB definition, it is very different.¹⁵ First, the truth of a belief or judgment is to be interpreted as *full* belief. Second, most of our beliefs are not certain, but only probable. But when they are such, they are not knowledge. They are knowledge only when they are reliable. It is tempting here to recognise a condition similar to that which figures in the reliabilist views of knowledge proposed later by philosophers like Alvin Goldman or Fred Dretske.¹⁶ But as Sahlin shows, the full definition has to be that "a belief is knowledge if it is obtained by a reliable process and if it always leads to success".¹⁷ This definition combines the feature of reliability with the idea that the truth of a belief is associated to the success of the actions to which it leads.¹⁸

However, the problem of beliefs which could meet these conditions, but which would nevertheless be true just by luck—the Gettier problem—resurfaces: how can we be sure that we are certain, as Ramsey says, "when we reflect on our reliability"? At this point Sahlin proposes his own *evidentiary value model*:¹⁹ it is not enough that we have a certain amount of evidence for our beliefs, for all kinds of evidence are not equally good, but that there be *mechanisms* which tell us something about the connection between the evidentiary fact and the thesis. These are the reliable processes. Making assessments of the evidentiary value in the light of

evidentiary mechanisms thus prompts us to avoid making decisions based on wrong connections between evidentiary facts and theses. So a person X's belief, that *p*, is knowledge if it is obtained by a satisfactorily reliable evidentiary mechanism and if this belief is true (i.e., if it always leads to success).²⁰

It can hardly be said that on such a model knowledge is first. On the contrary, as in the Russellian and traditional picture, knowledge is true

belief, warranted by reliable mechanisms which are Gettier stoppers, so to say. Ramsey was aware of John Cook Wilson's conception of knowledge, according to which it is an unanalysable intuition of the nature of things, but he could not accept it.²¹ I shall come back to this below. He wanted, like the pragmatists, to define scientific knowledge as apt and reliable true belief.

3. Probabilistic beliefs

I have made this quick detour through Ramsey's views in order to assess a recent proposal by Sarah Moss in her book *Probabilistic Knowledge*, that there can be probable knowledge.²² The Bayesian probabilist denies that the notion of knowledge makes sense: there is only probable belief.²³ Indeed if "knowledge" is understood as justified and certain belief, but failing to be true, the notion of probable knowledge poses no problem. But Moss's project is to combine a Bayesian picture of belief as credence with the claims about knowledge which are distinctive of the "knowledge-first" view. In that, she is revisionist both about belief, which she takes not only to be probabilistic but also not a feature of attitudes but of contents, and about knowledge, which she takes to be probable. The proposal looks bold. Is it coherent?

Moss takes a radical Bayesian view: belief is by nature partial belief, or as she says, probabilistic belief. She grants the existence of full beliefs but takes these to be certainties of degree 1. So she seems to defend a disjunctive view: belief is either partial or full. Actually, she defends a wholesale reduction of all beliefs to degrees of belief or credences. But she does not draw the conclusion that many Bayesians draw, that we can never reach full knowledge, and that we can only have beliefs with a higher probability. In that respect she is in agreement with Ramsey's conception of knowledge as full belief.²⁴ She claims that credences enjoy all the properties that knowledge is supposed to have according to "knowledge-first" theorists: they can be factive, they can be the contents of assertions, and they are norms for actions. So she parts company with Bayesians who typically talk of justification rather than knowledge. For her the concept of knowledge has just as much applicability to states of partial credence as it does to full beliefs.

Bayesianism is already pretty radical in the claim that there are no

beliefs, but only degrees of belief: full beliefs do not play any role in deliberation.²⁵ But Moss's view is even more revisionist, for she also rejects the relational view of attitudes. She takes propositions to be, in a classical sense, the set of possible worlds in which it is true. A proposition is something that is true or false, depending on how the world is; at a first pass, we can identify it with a "truth condition", the set of possible worlds in which it is true. On the standard picture of the mind, attitudes like belief are relations to such propositions, their contents. And, assuming that people have credences that are not reducible to beliefs in propositions (e.g., about objective chances or evidential probabilities), these credences are different relations to these same objects. In other words, for Moss probabilistic belief that it will snow this week is not:

- belief that the objective probability that it will snow this week is high, or belief that the frequency of snow in some relevant reference class is high, or
- belief that the evidential probability that it will snow this week, conditional on some specific body of evidence, is high.

In contrast, probabilistic beliefs, in Moss's sense, are not the subjective probabilities of the Bayesian. They are sets of "probability spaces". Each space in the set has its own set of worlds, and a probability function defined over those worlds. They are not supposed to be things that the believers need to grasp, but they are the information and possibilities that speakers take into account in their beliefs.

According to Moss, a *probability space* is a pair consisting of a set of worlds *O* (the possibilities "opened" by the space) and a function from propositions to real numbers obeying the axioms of the probability calculus and assigning non-zero probability only to propositions overlapping *O*. A *probabilistic content* is then a set of probability spaces. Thus if I am unsettled as to whether to take a possibility seriously at all—if I am not sure whether something might be true—then it will be included in some of the probability spaces in my belief set, but not others. If I am unsettled as to how seriously to take some possibility—if I am not sure how likely it is—then not all of the measures in my belief set will assign it the same probability. Probability spaces are not possible worlds, they are states of information. Thus one might think that it is more likely to rain

than not. What you think expresses a constraint on probabilities, viz., that the probability of rain is greater than the probability of no rain: $P(\text{Rain}) > P(\text{No Rain})$. We model your opinion by saying that you believe the probabilistic content $\{P \mid P(\text{Rain}) > P(\text{No Rain})\}$, or, say, if you believe that the probability of ice melting in the North Pole is between 0.6 and 0.8, you believe the probabilistic content $\{P \mid 0.6 P(\text{North Pole ice melts}) < 0.8\}$.²⁶ In other words, probability is not an external operator on propositions, it is *directly* embedded in the propositions.

On this view there is no need to show that degrees or beliefs which are incoherent lead to Dutch books, for it is enough that the probability spaces are inconsistent. It is quite important in this picture that the probabilistic contents are not, as in the classical Bayesian picture, part of the attitude of belief, in the manner indicated by Holton in the quote above.²⁷

Moss contrasts two views of the attitude of belief:

- (i) *Complex attitude*: belief to degree 0.6 that p (p being a proposition)
- (ii) *Simple attitude*: belief with complex content (a probability space).

On the complex content account, you have a belief with a probabilistic content, namely the set of probability spaces that assign higher probability to Jones smoking than Brown, conditional on Smith smoking. On the complex attitude account, your attitude is much more complicated.²⁸

Moss sides with the simple attitude (*believing*) with complex content account. On this view, belief is not a relation of an individual to a content. It is a content, exemplified in the individual. This looks like a version of the non-relational view of belief, but not quite, since Moss continues to talk of relations, but the move is clear: having a 0.75 degree of belief that p , or being 75% confident that p is not understood as involving a new relation (the *being-75%-confident-that* relation), but a belief relation to the set of probability spaces that assign 75% probability to p . Complexity in the contents allows for simplicity in the relations.

According to Moss, probabilistic contents play all the roles of the propositions in the classical picture. She entertains a kind of functionalist picture, according to which probabilistic contents exist because they fulfil these roles.

The first role played by contents of belief concerns explanations of action: same actions entail same contents: if you and I both open our

umbrellas, this is a strong reason to attribute to us the belief that it rains. This applies to probabilistic content: the fact that agents believe the same probabilistic content can help explain why those agents act the same way: you and I each grab an umbrella simply because each of us have some fairly high credence that it will rain. The second role for belief contents is to explain agreement: we agree because we believe the same thing. The third role is that contents explain how we can change our minds. The fourth role is explanation of the rationality of beliefs: your beliefs are consistent just in case they have consistent contents, and inconsistent just in case they have inconsistent contents.²⁹ These roles do not entail that we reject full beliefs in explanation, but in conditions of uncertainty, probabilistic beliefs explain better. In explaining actions, one may sometimes charitably interpret agents as acting on full beliefs in accordance with principles of instrumental reasoning. But one may sometimes charitably interpret agents as acting on probabilistic beliefs in accordance with principles of standard decision theory. Insofar as your actions are explained by appealing to the contents of your beliefs, decision-theoretic explanations of your actions must appeal to probabilistic contents of belief.

4. Probable knowledge

Having reinterpreted beliefs in probabilistic terms, Moss proceeds to defend her main thesis: there can be probabilistic knowledge, which is not knowledge of objective chances, but which has all the properties of ordinary knowledge. It is factive, argues Moss, because there is a sense in which we can trivially attribute probabilistic knowledge, if we understand truth in a deflationary manner: thus “Jones probably smokes” says just “It is true that Jones probably smokes”. Moss adduces certain linguistic intuitions. She gives arguments like this:

We do, in fact, use ‘know’ (and other factive attitude verbs such as ‘realize’ or ‘remember’) to embed sentences featuring epistemic vocabulary: “She realized she probably liked him as more of a friend”; “He remembered that they might not be home yet”; “She gave him the money despite knowing full well that he’d probably never repay her”. Moreover, the relevant epistemic vocabulary seems exactly the kind that resists propositional analysis: following up on “she realized she probably liked him as more of

a friend” with “though in fact she was romantically attracted to him” feels odd, requiring us to revisit and reinterpret the initial claim. If that sentence were merely attributing knowledge of propositions about probabilities, this would be puzzling. So it must attribute knowledge of a non-propositional, probabilistic content.³⁰

Moss also claims that probabilistic knowledge can constitute genuine knowledge because it is liable to produce Gettier cases of the following sort:

Alice has high credence that she finds Bert attractive on the basis of her fluttering nerves. Her fluttering nerves are actually not caused by Bert, but by a recent injection of adrenaline that she does not know about. Alice has plenty of evidence that she finds Bert attractive. Her evidence is misleading, but it would nevertheless be irrational for her to ignore it. Alice’s total evidence may well demand that she have high credence that she finds Bert attractive. According to the third proposal, it follows that she knows that she probably finds Bert attractive. But intuitively, Alice does not know that she probably finds Bert attractive, as her high credence is the result of intervening epistemic luck.³¹

Moss also argues that the knowledge norm of assertion—that one asserts that *p* if and only if one knows that *p* requires probabilistic knowledge, and that probabilistic knowledge can be a norm of action:

Because credences stand apart from full beliefs in propositions, the latter opinions cannot always be substituted for the former in explanations of action. To state a more serious argument in a similar spirit: probabilistic beliefs figure in intuitive explanations of action under conditions of uncertainty. In explaining actions, one may sometimes charitably interpret agents as acting on full beliefs in accordance with principles of instrumental reasoning. But one may sometimes charitably interpret agents as acting on probabilistic beliefs in accordance with principles of standard decision theory. Insofar as your actions are explained by appealing to the contents of your beliefs, decision theoretic explanations of your actions must appeal to probabilistic contents of belief.³²

But this hardly shows that probable knowledge can be a norm of action. It only shows that credences have to be backed by requirements of safety, or some other requirements, such as the requirement that the

evidence is obtained through some form of reliable cognitive ability.³³ These requirements look pretty much like the theory of evidentiary value proposed by Sahlin. But Sahlin never said that high evidentiary value would “constitute” knowledge. Do these arguments show that there can be probabilistic knowledge?

First the claim that some contents containing adverbs like “probably” and other epistemic modals can be included within probability spaces is trivial, for the notion of “probability space” is so large that it includes any content whatsoever, including those which are straightforward truth-conditional. The fact adduced by Moss that many assertions express the contents of beliefs, including high probability beliefs, does not show at all that probabilistic contents can be the subject of assertions, if the norm of assertion is knowledge. This is question-begging, because high probability is not knowledge, as the familiar Lottery cases show: if you have a ticket in a lottery of 1,000 tickets and say “My ticket will not win” this will elicit the answer: “But you don’t know that”.³⁴ Moss argues that what matters for the factivity of knowledge is that *S knows that it’s likely that p* entails *it’s likely that p* ; what matters for sensitivity or safety is that *S* wouldn’t have believed *p* was likely if it hadn’t been, or that it couldn’t easily have happened that *S* believes *p* to be likely even though it’s not. But that is not what the notion of factivity means. It does not mean that we can predicate some contents as true, but that this predication entails that they are genuinely true. At one point Moss claims:

There is a clear affinity between my claims about truth and knowledge and claims defended by some ethical expressivists. At a first pass, expressivists say that ethical sentences express different sorts of attitudes from the beliefs expressed by ordinary simple sentences. As developed by quasi-realists such as Blackburn and Gibbard, ethical expressivism is nevertheless consistent with the existence of moral truths and moral knowledge. Ascriptions of moral knowledge can be used to express the same sort of attitude as the sentences containing ethical vocabulary that they embed.³⁵

But this is surprising, since moral expressivists precisely do not say that we have genuine moral knowledge, but just that our assertions mimic realist truth and knowledge ascriptions, while in fact being neither true nor false. If Moss wants to claim that ordinary knowledge ascriptions are in fact probabilistic, the expressivist view is not the right move to make.

All of Moss's arguments that probabilistic beliefs can "constitute" knowledge presuppose that beliefs with a high probability *can*, in certain contexts, be understood as claims to knowledge. But this in turn presupposes that some form of contextualism about knowledge is correct. Contextualism about knowledge is the thesis that knowledge is relative to knowledge ascriptions.³⁶ But contextualism is a thesis about linguistic ascriptions, not about the nature of knowledge as a property independent of such ascriptions. The "knowledge-first" view does not make sense unless knowledge is a real property of individuals. The fact that these ascriptions manifest certain linguistic features does not show that the former entails the latter.³⁷

5. Knowledge first and what it takes

The main claim of the "knowledge-first" program can be understood in two ways:

Weak: Knowledge is first because some high credence beliefs, or beliefs passing some epistemic condition of warrant, anti-luck, or safety are met and can "constitute" knowledge. But this phrase is ambiguous, because "constitute" may mean that some beliefs have the characteristics of knowledge attributions, without these beliefs being true, as they should be if they were indeed knowledge.

Strong: Knowledge is first because it is different *in kind* from belief.

Moss defends a version of the weak view. The strong thesis, as it is defended by Timothy Williamson, says that knowledge is not a kind of belief and that knowledge is primitive and cannot be analysed in terms of belief plus some conjunctive factor, although knowing entails believing. This is not the version of "knowledge first" that Moss defends in her book. Nevertheless, in many cases, Williamson and other proponents of the "knowledge-first" program accept that credences, when backed by our knowledge of probability, can be considered as knowledge.

Thus John Hawthorne and Jason Stanley suggest that cases where we appear to be acting appropriately on the basis of our credence or partial belief in *p* are in fact cases where we are acting on the basis of our know-

ledge, namely, our knowledge *of the probability of p*. Stanley illustrates this idea as follows:

[T]here are lotteries in which it is rational for me to buy a ticket, even though I do not know that I will win; when pressed to defend my purchase, I will respond that there is a chance I will win. But this is just to say that there are certain types of actions that I perform on the basis of beliefs about chances. In order for these actions to be acceptable, such beliefs must still constitute knowledge.³⁸

But what kind of probability is this? The relevant kind of probability is epistemic probability, where the latter is understood as probability relative to one's knowledge or evidence.³⁹ Stanley tells us that we don't need to know the exact numerical probability—to justify taking the umbrella, we don't need to know, for example, that the epistemic probability of rain is precisely 0.4. Rather, what we need to know is “that it is reasonably likely that it will rain, that is, that it is likely enough to warrant the disutility of bringing an umbrella on my walk.” But in the first place, this is not the kind of probable knowledge that Moss has in mind, because she tells us that probabilistic beliefs, when they are knowledge, and not full beliefs about probabilities. In the second place, we can doubt that even reasonable beliefs with a high degree of credibility can be knowledge. As Ramsey said, other conditions have to be met: reliability and truth.

Now there is an even stronger sense in which knowledge can be first and different in kind from belief. This is the sense to which Nils-Eric Sahlin alludes to when he quotes Cook Wilson,⁴⁰ who at the time when Ramsey was discussing these issues in Cambridge, was the Wykeham Professor of Logic at Oxford. Cook Wilson argued, in *Statement and Inference*, that knowledge cannot be associated to any belief, because belief in order to be knowledge needs to be supported by evidence, and the evidence itself has to be known, so that the view that knowledge is a form of belief guaranteed by evidence is circular.⁴¹ It follows, according to Cook Wilson, that knowing has nothing to do with the “strength” of evidence, since the evidence may fail to be conclusive. So knowledge is not a kind of belief. It is a primitive, undefinable, notion.⁴² Williamson agrees, and there is a trace of this Cook Wilsonian conception in his thesis that knowledge is evidence, and that one's evidence is one's knowledge, without the evidence *justifying* a form of belief.

Does it follow that for Williamson knowledge has nothing to do with belief? Actually, he claims that knowledge entails belief, even if belief is “botched knowledge”.⁴³ Moreover, the fact that he identifies knowledge and evidence shows that he does not take evidence to justify knowledge, hence avoids the circularity that Cook Wilson discerned in this view. And Williamson does not reject the idea that certain characteristics of knowledge can be applied to it, such as safety and reliability, although these do not define knowledge. Cook Wilson takes a harder line, which was present in the philosophical tradition long before. It is the view that knowledge is a form of intuition of essences. In order to know that *p*, one must in some sense *see*, by some form of intellectual intuition, or *grasp* that *p* and such seeing is analysable, close to what Aristotle called *noesis*. This, Maria Rosa Antognazza has argued, was the traditional, pre-Gettier, pre-Russellian and indeed pre-Ramseyan conception of knowledge.⁴⁴ In this tradition, not only is belief not part of knowledge, but belief does not even *aim* at knowledge. As Antognazza argues, Hume is largely responsible for classifying all cognitions under the species “belief”, but in the Anglo-American tradition there is a remnant of the notion of knowledge as intellectual grasp of the nature of things in Russell’s notion of acquaintance: acquaintance is a kind of primitive contact with things, which cannot be explained through belief and evidence.⁴⁵

Another heir to this essentialist tradition is Edmund Husserl.⁴⁶ He too took knowledge to be a form of *noesis*, implying the presence of the thing to the mind, and the grasp of its essence. And here we have the basis of a link between the rejection of the two dogmas from which we started: if belief is not a relation to propositions, and if knowledge is first in the sense of an intuition of essences, we come back full circle to a view which one might call pre-Ramseyan and pre-Russellian.

This is not the place here to discuss this conception of knowledge. But one may raise some doubts. If knowledge is a form of undefinable intuition, does it follow that we cannot explain it at all? And if it has nothing whatsoever to do with belief and evidence, how can it be acquired and learnt? Can’t believing with some form of justification and warrant be at least in part a means to reach knowledge? Is knowledge acquired by magic or by some sort of illumination? I gather that this is one of the reasons why Ramsey did not accept Cook Wilson’s conception of know-

ledge as primitive and took it as “an Oxford folly”. Knowledge may be first without being an unexplainable and ineffable quality.

To conclude, the notion of probable knowledge makes good sense when it refers to knowledge of probabilities in the objectivist sense. It also makes sense when “knowledge” is meant as a kind of higher-grade belief, which can be rationally justified but which may fail to be true. But if knowledge is supposed to be a factive and primitive attitude in the sense of the friends of the “knowledge-first” program in epistemology, it can hardly be understood in this sense. Sarah Moss’s functionalist conception of “probabilistic” knowledge makes room for a form of contextualization of this notion, but it turns out to be a sophisticated version of Bayesianism. Neither does it reestablish the relational view of attitudes. I have not examined the other options, but the prospect for a full-blown “knowledge-first” conception of probable knowledge looks bleak.

Acknowledgement

I have for many years been inspired and learnt from Nils-Eric Sahlin’s work. This article is a small testimony of my gratitude.

Notes

1. This is the view that Gettier (1963) has attacked, showing that there could be cases of justified true beliefs which fail to be knowledge. Most counterexamples rest on the idea that one could know by chance, hence the view that one needs to show how knowledge has stood immune from chance. For objections to the idea that this view is the traditional one in philosophy, see Dutant 2015 and Antognazza 2024.

2. Ayers 2019; Williamson 2000; Antognazza 2024. The latter argues, against many commentators who took Plato’s *Theaetetus* as the forerunner of the JTB view, that Plato had in fact the view that knowledge is not a kind of belief.

3. Williamson 2000.

4. Brentano 1889; Montague 2022; see also Kriegel 2018.

5. Montague 2022.

6. Gärdenfors & Sahlin 1988.

7. Jeffrey 1968; Sahlin 1990.

8. Russell 1910.

9. ‘Facts and propositions’ in Ramsey 1990, pp. 35, 38.

10. Ramsey did not hold the so-called “redundancy” view often attributed to him, but clearly defended a form of minimalism about truth; see Dokic & Engel 2002.

11. Holton 2014, pp. 13–14.

12. Even in this latter case, it is not clear that Ramsey wanted to countenance the existence of propositions as special entities which would be the objects of beliefs. For an analysis of difficulties of this dual view (that there are both degrees of belief and “flat-out beliefs”) see Frankish 2009.

13. ‘General propositions and causality’, in Ramsey 1990.

14. Ramsey 1990, pp. 110–111.

15. Sahlin 1990, p. 81.

16. Dretske 1970; Goldman 1979.

17. Sahlin 1990, p. 93.

18. For an analysis of this “success semantics”, cf. Dokic & Engel 2002.

19. Adapted from an earlier proposal by Halldén and others, see Sahlin 1986.

20. Sahlin 1990, p. 100.

21. Cook Wilson 2002.

22. Moss 2018; see also Moss 2013.

23. Jeffrey 1968.

24. But Moss does not quote Ramsey on knowledge. Her conception may have something in common with Isaac Levi’s notion of knowledge as “serious possibility” (Levi 1980) but she does not quote him in this respect.

25. See Maher 1986.

26. Moss 2018, pp. 6–7.

27. Holton 2014, pp. 13–14.

28. Moss 2018, p. 4.

29. Moss 2018, p. 7.

30. Moss 2018, p. 10.

31. Moss 2018, p. 119.

32. Moss 2018, p. 8.

33. Konek 2016.

34. Williamson 2000, p. 246.

35. Moss 2018, p. 121.

36. See, e.g., DeRose 2009.

37. See Williamson 2005; *pace* Ichigawa 2017.

38. Stanley 2005, p. 10.

39. Hawthorne & Stanley 2008, p. 582.

40. Sahlin 1990, pp. 85–86.

41. Cook Wilson 2002, § 49.

42. See Prichard 1950; Marion 2000.

43. Williamson 2000, p. 47.

44. Antognazza 2024.

45. Russell 1910.

46. See Mulligan 2014.

References

- Antognazza, Maria Rosa 2024. *Thinking with Assent*, Oxford: Oxford University Press.
- Ayers, Michael 2019. *Knowing and Seeing*, Oxford: Oxford University Press.
- Brentano, Franz 1889. *Vom Ursprung sittlicher Erkenntnis*, Hamburg: Meiner, trans. Roderick M. Chisholm & Elizabeth H. Schneewind 1969. *The Origin of Our Knowledge of Right and Wrong*, London: Routledge & Kegan Paul.
- Cook Wilson, John 2002 [1926]. *Statement and Inference with other Philosophical Papers*, vol. 2, Oxford: Clarendon Press; reprint: Bristol: Thoemmes Continuum.
- DeRose, Keith 2009. *The Case for Contextualism*, Oxford: Oxford University Press.
- Dokic, Jérôme & Pascal Engel 2002. *Ramsey, Truth and Success*, London: Routledge.
- Dretske, Fred 1970. 'Epistemic operators', *The Journal of Philosophy* 67:24, pp. 1007–1023. <https://doi.org/10.2307/2024710>
- Dutant, Julien 2015. 'The legend of the justified true belief analysis', *Philosophical Perspectives* 29:1, pp. 95–145. <https://doi.org/10.1111/phpe.12061>
- Frankish, Keith 2009. 'Partial belief and flat-out belief', in Franz Huber & Christoph Schmidt-Petri eds, *Degrees of Belief: An Anthology*, Berlin: Springer, pp. 75–93.
- Gärdenfors, Peter & Nils-Eric Sahlin eds 1988. *Decision, Probability and Utility*, Cambridge: Cambridge University Press.
- Gettier, Edmund L. 1963. 'Is justified true belief knowledge?', *Analysis* 23:6, pp. 121–123. <https://doi.org/10.1093/analys/23.6.121>
- Goldman, Alvin 1979. 'What is justified belief?', in George S. Pappas ed., *Justification and Knowledge: New Studies in Epistemology* (Philosophical Series Studies in Philosophy 17), Boston: D. Reidel, pp. 1–25.
- Hawthorne, John & Jason Stanley 2008. 'Knowledge and action', *Journal of Philosophy* 105:10, pp. 571–590. <https://doi.org/10.5840/jphil20081051022>
- Holton, Richard 2014. 'Intention as a model for belief', in Manuel Vargas & Gideon Yaffe eds, *Rational and Social Agency: The Philosophy of Michael Bratman*, New York: Oxford University Press, pp. 12–37.
- Ichigawa, Jonathan J. 2017. *Contextualizing Knowledge*, Oxford: Oxford University Press.
- Jeffrey, Richard C. 1968. 'Probable knowledge', in Imre Lakatos ed., *The Problem of Inductive Logic*, Amsterdam: North Holland Publishing Company. Reprinted in Peter Gärdenfors & Nils-Eric Sahlin eds, *Decision, Probability and Utility: Selected Readings*, Cambridge: Cambridge University Press 1988, and in Richard C. Jeffrey, *Probability and the Art of Judgment*, Cambridge: Cambridge University Press 1992.
- Konek, Jason 2016. 'Probabilistic knowledge and cognitive ability', *Philosophical Review* 125:4, pp. 509–587. <https://doi.org/10.1215/00318108-3624754>
- Kriegel, Uriah 2018. *Brentano's Philosophical System*, Oxford: Oxford University Press.
- Levi, Isaac 1980. *The Enterprise of Knowledge*, Cambridge, MA: MIT Press.
- Maher, Patrick 1986. 'The irrelevance of belief to rational action', *Erkenntnis* 24:3, pp. 363–384.

- Marion, Mathieu 2000. 'Oxford realism: Knowledge and perception', *British Journal for the History of Philosophy* 8:2, pp. 299–338, 485–519. <https://doi.org/10.1080/09608780050043235>
- Montague, Michelle 2022. 'Rethinking the attitudes', in Uriah Kriegel ed., *Oxford Studies in Philosophy of Mind*, vol. 2, pp. 269–290.
- Moss, Sarah 2013. 'Epistemology formalised', *The Philosophical Review* 122:1, pp. 1–43.
- Moss, Sarah 2018. *Probabilistic Knowledge*, Oxford: Oxford University Press. <https://doi.org/10.1215/00318108-1728705>
- Mulligan, Kevin 2014. 'Knowledge first—A German folly?', in Julien Dutant, Davide Fassio & Anne Meylan eds, *Liber Amicorum Pascal Engel*, Geneva: University of Geneva, pp. 380–400. <http://www.unige.ch/lettres/philo/publications/engel/liberamicorum>, accessed 5 March 2026.
- Prichard, Harold A. 1950. *Knowledge and Perception. Essays and Lectures*, Oxford: Clarendon Press.
- Ramsey, Frank P. 1990. *F. P. Ramsey: Philosophical Papers*, ed. David Hugh Mellor, Cambridge: Cambridge University Press.
- Russell, Bertrand A. W. 1910. *The Problems of Philosophy*, London: Williams and Norgate.
- Sahlin, Nils-Eric 1986. 'How to be 100% certain 99.5% of the time', *The Journal of Philosophy* 83:2, pp. 91–111.
- Sahlin, Nils-Eric 1990. *The Philosophy of F. P. Ramsey*, Cambridge: Cambridge University Press.
- Stanley, Jason 2005. *Knowledge and Practical Interests*, Oxford: Oxford University Press.
- Williamson, Timothy 2000. *Knowledge and its Limits*, Oxford: Oxford University Press.
- Williamson, Timothy 2005. 'Contextualism, subject-sensitive invariantism, and knowledge of knowledge', *Philosophical Quarterly* 55:219, pp. 213–235. <https://doi.org/10.1111/j.0031-8094.2005.00396.x>

A hundred years of rational decision-making — and then what?

NILS-ERIC SAHLIN

Truth and probability

2026 marks the centenary of Frank Ramsey's 'Truth and probability', the essay in which he formulated what can be seen as the first complete theory of rational decision-making.¹ Aristotle formulated a logic of preference intended to help us choose in situations where there is no uncertainty, only a difficulty in ranking the options. In the centuries that followed, there were scattered attempts to formulate some kind of theory of informed decision-making. The need for a normative theory, or at least a well-thought-out guide, is apparent. "Anything goes" is not an alternative.

The importance of Ramsey's theory is demonstrated by some of the Nobel Prizes in economics. The core of his theory has been important for a number of Nobel laureates, including Kenneth Arrow, Gérard Debreu, Maurice Allais, Harry Markowitz, John Harsanyi, Amartya Sen and Daniel Kahneman. In their work they have reconstructed and strengthened the original theory. Many other economists, philosophers, psychologists, mathematicians and statisticians have also taken part in developing some aspect of Ramsey's theory. This wide-ranging interest indicates the impact of Ramsey's ideas on a number of scientific disciplines, not only theoretical ones but also empirical sciences such as psychology and decision sciences.²

So what does Ramsey teach us?

In his classic paper 'Psychology as the behaviorist views it' from 1913, John B. Watson writes:

Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no essential part of its methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness. The behaviorist, in his efforts to get a unitary scheme of animal response, recognizes no dividing line between man and brute. The behavior of man, with all of its refinement and complexity, forms only a part of the behaviorist's total scheme of investigation.³

It would be a mistake to say that Ramsey subscribes wholeheartedly to Watson's manifesto. Ramsey's ideas about beliefs, a belief being a map, do not fully harmonize with the conceptual structure of orthodox behaviourism.⁴ Nevertheless, and indirectly, 'Truth and probability' advocates a method that can be seen as an "objective experimental branch of natural science": a theory whose goal is to describe, predict and perhaps ultimately control behaviour, but not to explain why an agent has the beliefs or preferences he/she has, and definitely not in terms of introspection. "We shall find", Ramsey says, "that it is very appropriate to the measurement of belief as a basis of action, but in no way related to the measurement of an introspected feeling". It is fair to say that Ramsey gives us "a unitary scheme of [ideal] animal and [human]" decision-making.

Unlike Watson, Baruch Fischhoff and Paul Slovic, Ramsey does not study human behaviour. His research subject is the perfectly rational decision-maker, an ideal agent. He shows how, under ideal conditions, our degrees of belief (credences) and desires can be measured by means of applying a betting method cleverly designed by him. Ramsey demonstrates how his method can be used to measure the strength of our preferences. Assuming the operation of some intuitive principles of rationality, the strength of our preferences can be measured and represented as a utility measure determined up to a positive linear transformation (on an interval scale). Given this utility measure and the betting method, he shows that our "degrees of belief" can be measured and that the measure obtained will satisfy the laws of probability.

In other words, our degrees of belief can be represented by a probability measure. In addition, Ramsey was the first to state the Dutch book theorem. His paper does not supply formal proof, but in passing he mentions that if the ideal agent violates the axioms of probability theory, a book can be made against him and the agent will lose money whatever happens. Laying the groundwork of modern utility theory, Ramsey—also in passing—gives us the decision theory that has led to so many Nobel Prizes. He had proof of the value of collecting evidence years before it became established through the independent works of Leonard J. Savage and I. J. Good; he took higher-order probabilities seriously; and, in a derivation of the “rule of succession”, he introduced the notion of “exchangeability” (but without giving it that name). Ramsey’s theory is just about as complete as any such theory can be⁵—complete in the sense that it gives us answers to almost all the fundamental questions we have about rational decision-making. And most important of all, the theory answers the question: How should we make decisions? Even so, its assumptions may be called into question, and so may the validity of its applicability.

It is important to remember that Ramsey’s theory is a descriptive theory. A single choice that the agent makes does not tell us much. But if we systematically study his/her choices, we can detect a pattern. From the perspective of the observer—from our perspective—it looks as if the agent chooses the action that maximizes subjective expected utility; the agent appears to combine personal probabilities and utilities. Indirectly, however, it is a normative theory. Who does not want to make decisions the way the ideal, the perfectly rational, agent does?

Ramsey’s heirs

Post-Ramsey, a rich variety of theories on rational decision-making have been developed. Many of them are discussed in Peter C. Fishburn’s classic paper ‘Subjective expected utility: A review of normative theories’.⁶ The theories discussed by Fishburn differ from one another in several respects. They make different “ontological” assumptions (e.g., assuming the existence of worlds, propositions, states, outcomes, actions and so on), different epistemological assumptions (what it is supposed that an agent knows and what cognitive capacities he/she has), different operational assumptions, and so on. The theories therefore provide us with

partly diverging representation theorems. The main message, however, is the same: an ideal agent maximizes (subjective) expected utility.

Know the limits of your theory

Ramsey seems to be very well aware of the shortcomings of his theory.

I have not worked out the mathematical logic of this in detail, because this would, I think, be rather like working out to seven places of decimals a result only valid to two. My logic cannot be regarded as giving more than the sort of way it might work.”⁷ The theory, despite being so complete and well thought out in every detail, comes with a couple of problems in terms of applicability and assumptions about the precision with which we can measure values and degrees of beliefs. This is perhaps what Ramsey is alluding to when he says “working out to seven places of decimals a result only valid to two.”⁸

Since the 1950s, empirical researchers have shown that for many reasons, we are unable to make decisions in the way that the ideal agent does. This research has provided us with an amazing and very useful body of knowledge about human decision-making. These results seem to show that as decision-makers we are one-eyed, short-sighted and have a seriously refractive error.⁹ Fischhoff and Slovic in particular, but also many other decision scientists have shown that this research can be used in a constructive way. Knowledge of his/her weaknesses may help a decision-maker to make better-informed and deliberate decisions, and we can also help him/her to do so, if desired.

Moral philosophers tell us that “ought implies can”. If we cannot do something, it cannot be the case that we ought to do it. This dictum, if applicable, calls into question the normativity of the normative theories of subjective expected utility. However, there are other, more theoretical problems with these theories.

Savage’s insights and position

In *The Foundations of Statistics*, Savage writes:

[T]here seem to be some probability relations about which we feel relatively “sure” as compared with others. When our opinions, as reflected in

real or envisaged action, are inconsistent, we sacrifice the unsure opinions to the sure ones. The notion of “sure” and “unsure” introduced here is vague, and my complaint is precisely that neither the theory of personal probability, as it is developed in this book, nor any other device known to me renders the notion less vague. There is some temptation to introduce probabilities of a second order so that the person would find himself saying such things as “the probability that B is more probable than C is greater than the probability that F is more probable than G.” But such a program seems to meet insurmountable difficulties.¹⁰

Savage notes that our state of knowledge may be such that we are unsure of the accuracy or precision of the uncertainty assessments we make. The insurmountable obstacles he is talking about are two. First, if we are uncertain about the uncertainties, we can simply combine the uncertainties at the different levels to obtain a single new uncertainty (probability) assessment, and nothing would be gained. This composite probability, as Savage calls it, can then be used as a basis for action. I have previously argued that Savage’s proposed solution is not entirely satisfactory, as it may lead to the use of a probability assessment (the composite probability) that—given our state of knowledge—is not a justifiable (reliable) assessment.¹¹ Second, Savage also sees a danger in allowing policymakers to have uncertain uncertainties. That kind of higher-order uncertainty, second-order probabilities, he says, can lead to an endless hierarchy of uncertainties (probabilities) and a less realistic theory. Here, too, one may object to Savage’s argument. It seems that the problem arises only if the probabilities at each level of the hierarchy are interpreted in the same way.¹²

The nub of the problem is that maximization requires precision. The information, knowledge and values of the ideal agent must be represented with absolute precision by means of probability and utility functions. In a footnote later added to his book, Savage writes, “One tempting representation of the unsure is to replace the person’s single probability measure [...] by a set of such measures, especially a convex set”.¹³ But then, a point which Savage does not mention in his note, we can no longer maximize expected utility, and new decision rules need to be invented.

Proposed solutions and new problems

Philosophers in particular, but also economists and mathematically inclined psychologists, have tried to develop theories of decision-making that allow the decision-maker a greater degree of indeterminacy, both in terms of his/her knowledge and with regard to his/her values.¹⁴ The best-known prescriptive theory is definitely that of Isaac Levi.¹⁵ Among the theories derived from empirical research on human decision-making, Daniel Kahneman and Amos Tversky's famous paper 'Prospect theory' is by far the best-known example.¹⁶ One theory I am particularly enamoured with is the one developed by Peter Gärdenfors and myself.¹⁷ Levi's theory advocates a lexicographic decision rule, Gärdenfors and Sahlin's theory prescribes that you maximize the minimal expected utility; and Kahneman and Tversky have another rule. These are just three instances; there are many more examples of "generalized" decision rules.

It is well known that Levi's theory violates one of the fundamental ordering principles of rational decision-making, the principle of the independence of irrelevant alternatives (if a choice alternative is sub-optimal in one situation, it should not be possible to make it optimal by adding new alternatives). The independence axiom, Savage's sure-thing principle, is another type of rationality axiom (it tells us that if an alternative A is judged to be as good as another B in all possible states, and better than B in at least one, then a rational decision-maker prefers A to B). As Teddy Seidenfeld¹⁸ points out, there are basically only two types of rationality axioms that are of interest: ordering assumptions and independence assumptions.

In his paper, Seidenfeld gives an example which shows that theories like prospect theory violate the independence axiom, and so does our theory. These types of theories can make counterintuitive recommendations in consecutive choice situations. Giving up the independence axiom to obtain explanatory power or rational resilience comes at a cost—a cost that Levi's theory, with its lexicographically ordered set of decision rules, does not have to pay. But as was observed above, there are other problems with his theory.

Adam Elga has argued that rationality demands sharpness.¹⁹ Unsharp probabilities, uncertain uncertainties, unreliable or indeterminate probabilities and new types of decision rules lead to problems. In order to illustrate this difficulty, Elga designs a clever choice problem. He demon-

strates that Levi's theory does not solve the problem as desired, but he does not show that all theories of the type mentioned above fall into the same trap. Some theories that allow for unsharp probabilities and have alternative decision rules prescribe the "right" action alternative.²⁰

Nonetheless, Elga is right—*rational decision-making demands precision*. I have no mathematical proof of this conviction. But it seems obvious that if we are to allow uncertain uncertainties, we must also invent new decision rules, which means that the theories we formulate will inevitably violate one or other of the fundamental axioms of rationality (ordering or independence). There is no way out of this dilemma. The best thing would have been if there had been a meta-theory that could be used in situations of deep uncertainty—a theory helping us to choose which theory to apply in a particular situation, and which new type of decision rule could be used there and then. Such a meta-theory does not exist, and it seems from what has been said above that it is not possible to formulate such a theory.

Solar radiation modification (SRM)

One piece of advice we all received is to begin with an example. I have chosen to end with one. My example is a real policy problem, a problem that presents those who have to deal with the issues with hard choices, difficult decisions and challenging recommendations. Policymakers have to do what they do without the guard-rails of the theory of rational decision-making. Not only that; the available alternatives seem to be anything but impeccable.

In late 2024, the European Group on Ethics in Science and New Technologies (EGE), the Group of Chief Scientific Advisors to the European Commission (GCSA) and Science Advice for Policy by European Academies (SAPEA) published three independent reports on solar radiation management. SAPEA forms part of the Scientific Advice Mechanism to the European Commission. Their systematic evidence report on solar radiation modification (SRM) provides a systematic overview of the present state of knowledge. The GCSA is part of the so-called Scientific Advice Mechanism, an expert group within the Directorate-General for Research and Innovation of the European Commission. One of their tasks is to make policy recommendations on the basis of publicly avail-

able scientific evidence. The EGE is an independent advisory body of the President of the European Commission. The group provides the European Commission with advice on all aspects of EU legislation and policies, ethical, societal and fundamental rights issues intersecting with the development of science and new technologies. I bring this up because what I say below about SRM relies entirely on these groups' analyses of the research on SRM and the applicability and risks of technologies—what uncertainties exist and their respective degree of uncertainty, the nature, number and severity of the risks that can be identified, and what knowledge gaps there are.

Solar radiation modification has been defined as “a deliberate and potentially large-scale intervention in the Earth’s climatic system”.²¹ The aim is to temporarily or permanently change this system by reducing the adverse effects of greenhouse-gas concentrations. Measures discussed include stratospheric aerosol injection (SAI), cloud brightening (CB), cirrus-cloud thinning, surface brightening and space mirrors. The aim is to reflect sunlight away from the Earth or increase the escape of infrared radiation.²²

According to SAPEA, SAI is the most widely studied SRM technique. The technique mimics the aftermath of a volcanic eruption, its effect being achieved by, for example, injecting precursor gas sulphur dioxide into the stratosphere. Of course, no large-scale trials have been undertaken with this technique; but existing climate models suggest that extensive global cooling is achievable using SAI. As SAPEA points out, SAI is surrounded by uncertainties, uncertain uncertainties and more or less extensive knowledge gaps. It is unclear which aerosol types might result in the smallest side-effects. It is also not known what amounts of aerosols are required to achieve the desired favourable effects on the climate. SAPEA also emphasizes that “[a]brupt phase-in or phase-out of SAI can lead to large and rapid global climate impacts, while unequal injections between hemispheres can lead to uneven distribution of these impacts”.²³ And these are just a handful of the (profound) uncertainties and knowledge gaps that surround SAI.

Cloud brightening means increasing the reflectivity of shallow low-level liquid clouds, thereby mitigating warming. SAPEA emphasizes that there are significant uncertainties with this technique. The model-based evidence available is permeated by uncertainties owing to “empirical

parameterisation of crucial processes and coarse model resolutions”.²⁴ There are, says SAPEA, “significant technological challenges to address before CB could be operationally deployed”.²⁵ I take that to mean that in the case of cloud brightening, the uncertainties are also both uncertain and numerous and the knowledge gaps, if not legion, at least so many as to call for vigilance.

As regards cirrus-cloud thinning, the scientific basis is less robust than for cloud brightening. It is, says the evidence report, unclear whether it actually possesses the desired potential for global cooling. The aim of surface brightening is to increase the fraction of sunlight that is reflected. The available evidence is based on modelling studies and seems to suggest that this type of technique will form a comprehensive obstacle to global warming. Finally, climate models support the potential, and positive, cooling effects of space mirrors. But the step from what can be done in a model to what happens when a technology is introduced on a large scale in a natural setting is fraught with uncertainties and knowledge gaps.

EGE and GCSA agree that this type of intervention comes with severe risks. For example, there is a risk that interventions of this kind might have undesirable impacts on the climate in unpredictable parts of the world, with undesirable and far-reaching consequences for the people living there as well as for the environment and wildlife. A risk that is also mentioned is that certain techniques may be used for military purposes.

Because of the significant knowledge gaps and high risks surrounding solar radiation modification, both EGE and GCSA recommend the European Commission to issue a moratorium. EGE writes: “Issue a moratorium on any large-scale use of SRM (on all experiments and applications of SRM that could have effects beyond the specific locus of intervention) within the territory of the EU, due to the extensive knowledge gaps around SRM technologies and the potential severity of risks they imply”.²⁶ GCSA writes: “Agree on an EU-wide moratorium on the use of SRM as a measure for offsetting climate warming”.²⁷ The two groups also issue other important recommendations, but these are not relevant to the purposes of this paper.

SRM is a very good example of a situation where Ramsey’s theory cannot be applied, nor can any of the classic theories that Fishburn discusses, for example Savage’s theory. At least, they cannot be utilized unless we pretend that the vast knowledge gaps do not exist, or that the uncertainties

are not as uncertain as they seem to be. Nor can they be applied unless we pretend that risks and uncertainties can be represented with a mathematical precision that our actual knowledge really does not permit. Only if we lie about the state of knowledge can we achieve the precision we need. And that would not be morally acceptable, would it?

We could, of course, try to employ some alternative to Ramsey's theory, a theory allowing for a set of probability measures. Such a theory would be one of those that try to deal with knowledge gaps and deep uncertainty. But as I have already pointed out, this is a solution beset with problems, and it comes at the price of bewilderment. The theories would probably, depending on which decision rules they favour, recommend different actions. The question then becomes which of the theories we should trust, and what reasons we have for choosing one theory over one of its competitors in this particular situation. We know that they all violate one or other of the classic axioms of rationality, and we know that we lack a meta-theory to help us choose one theory over another.

The EGE's and GCSA's recommendations are based on the premise that if knowledge gaps are too large and uncertainties too uncertain, there is every reason to be cautious. Prudence is advisable. That said, it is important to recognize that the so-called precautionary principle is not a decision rule we can rely on. Martin Peterson has shown that this is a principle that also violates some of the most basic principles of rational behaviour.²⁸

... and then what?

How can we make sound decisions if we do not have a theory to guide us, if there are no guard-rails protecting us from falling into the deep well of confusion and arbitrariness? What we should avoid are three things. First, we should avoid pretending that the knowledge gaps do not exist and that risk assessments and probability statements cannot be uncertain, and we should not act as if we know more than we do, just in order to maximize expected utility. That is definitely not something you should do if you are a policymaker and want to gain public trust. Second, we should not rely on ill-conceived decision principles, however politically attractive they may be. Third, as was pointed out by way of introduction, "anything goes" is not an option.

Notes

1. Ramsey 1931, pp. 156–198; Ramsey 1990, pp. 52–109.
2. Sahlin & Schwaag Serger 2022.
3. Watson 1913, p. 158.
4. Sahlin 1991.
5. Sahlin 2012.
6. Fishburn 1981.
7. Ramsey 1931, p. 180; Ramsey 1990, p. 76.
8. Ramsey 1931, p. 180; Ramsey 1990, p. 76.
9. Sahlin 2012.
10. Savage 1972, pp. 57–58.
11. Sahlin 1993.
12. Sahlin 1993.
13. Savage 1972, p. 58.
14. Gärdenfors & Sahlin 1982; Gärdenfors & Sahlin 1988.
15. Levi 1974; Levi 1980.
16. Kahneman & Tversky 1979.
17. Gärdenfors & Sahlin 1982; Gärdenfors & Sahlin 1988.
18. Seidenfeld 1988.
19. Elga 2010.
20. Sahlin & Weirich 2014.
21. SAPEA 2024a, p. 10.
22. SAPEA 2024a.
23. SAPEA 2024a, p. 11. Cf. SAPEA 2024b.
24. SAPEA 2024a, p. 11.
25. SAPEA 2024a, p. 11.
26. EGE 2024, recommendation 1, p. 47.
27. GCSA 2024, recommendation 2, p. 7.
28. Peterson 2006; Peterson 2007; Peterson 2026 (in this volume).

References

- EGE/European Group on Ethics in Science and New Technologies 2024. *EGE Opinion on Solar Radiation Modification*, Luxembourg: Publications Office of the European Union.
- Elga, Adam 2010. ‘Subjective probabilities should be sharp’, *Philosophers’ Imprint* 10:5, pp. 1–11.
- Fishburn, Peter C. 1981. ‘Subjective expected utility: A review of normative theories’, *Theory and Decision* 13, pp. 139–199.
- Gärdenfors, Peter & Nils-Eric Sahlin 1982. ‘Unreliable probabilities, risk taking, and decision making’, *Synthese* 53:3, pp. 361–386. <http://link.springer.com/article/10.1007/BF00486156>

- Gärdenfors, Peter & Nils-Eric Sahlin eds 1988. *Decision, Probability, and Utility: Selected Readings*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511609220>
- GCSA/Group of Chief Scientific Advisors to the European Commission 2024. *Solar Radiation Modification*, Luxembourg: Publications Office of the European Union (European Commission, Directorate-General for Research and Innovation).
- Kahneman, Daniel & Amos Tversky 1979. 'Prospect theory', *Econometrica* 47, pp. 263–292.
- Levi, Isaac 1974. 'On indeterminate probabilities', *Journal of Philosophy* 71:13, pp. 391–418.
- Levi, Isaac 1980. *The Enterprise of Knowledge: Essay on Knowledge, Credal Probability and Chance*, Cambridge, MA: MIT Press.
- Peterson, Martin 2006. 'The precautionary principle is incoherent', *Risk Analysis* 26:3, pp. 595–601.
- Peterson, Martin 2007. 'Should the precautionary principle guide our actions or our beliefs?', *Journal of Medical Ethics* 33:1, pp. 5–10.
- Peterson, Martin 2026. 'The precautionary principle has no exceptions', in Göran Hermerén, Johannes Persson, Eva Sjöstrand & Robin Stenwall eds, *Ethics, Risks and Difficult Decision-making: Examples and Epistemic Foundations*, Stockholm: The Royal Swedish Academy of Letters, History and Antiquities (this volume).
- Ramsey, Frank P. 1931. *The Foundations of Mathematics and Other Logical Essays*, ed. Richard Bevan Braithwaite, London: Kegan Paul, Trench, Trubner & Co.
- Ramsey, Frank P. 1990. *F. P. Ramsey: Philosophical Papers*, ed. David Hugh Mellor, Cambridge: Cambridge University Press.
- Sahlin, Nils-Eric 1991. 'Baconian inductivism in research on human decision making', *Theory & Psychology* 4:1, pp. 431–450. <https://doi.org/10.1177/0959354391014003>
- Sahlin, Nils-Eric 1993. 'On higher order beliefs', in Jacques-Paul Dubucs ed., *Philosophy of Probability*, Dordrecht: Kluwer, pp. 13–34.
- Sahlin, Nils-Eric 2012. 'Unreliable probabilities, paradoxes, and epistemic risk', in Sabine Roeser, Rafaela Hillerbrand, Per Sandin & Martin Peterson eds, *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, Dordrecht: Springer, pp. 477–498.
- Sahlin, Nils-Eric & Sylvia Schwaag Serger 2022. 'Decision-making in a time of spin and unspoken values', in Mats Benner, Göran Marklund & Sylvia Schwaag Serger eds, *Smart Policies for Societies in Transition: The Innovation Challenge of Inclusion, Resilience and Sustainability*, Cheltenham: Edward Elgar Publishing, pp. 228–244. <https://doi.org/10.4337/9781788970815.00015>
- Sahlin, Nils-Eric & Paul Weirich 2014. 'Unsharp sharpness', *Theoria: A Swedish Journal of Philosophy* 80:1, pp. 100–103. <https://doi.org/10.1111/theo.12025>
- SAPEA/Science Advice for Policy by European Academies 2024a. *Evidence Review Report: Solar Radiation Modification*, Berlin: SAPEA.

- SAPEA/Science Advice for Policy by European Academies 2024b. *Ethics of Solar Radiation Modification: A Literature Review*, Berlin: SAPEA.
- Savage, Leonard J. 1972 [1954]. *The Foundations of Statistics*, New York: John Wiley; 2nd rev. edn, New York: Dover.
- Seidenfeld, Teddy 1988. 'Decision theory without "independence" or without "ordering": What is the difference?', *Economy and Philosophy* 4, pp. 267–315.
- Watson, John B. 1913. 'Psychology as the behaviorist views it', *Psychological Review* 20:2, pp. 158–177.

The contributors

WÄNDI BRUINE DE BRUIN

Schaeffer Institute of Public Policy and Government Service, Sol Price School of Public Policy, and Dornsife Department of Psychology, University of Southern California, Los Angeles

Professor Bruine de Bruin utilizes behavioral science to better understand how we respond to information and make decisions about environmental and health risks. She has published more than 175 peer-reviewed publications on these topics and also served on expert panels on science communication, including for the National Academies of Sciences in the United States.

PASCAL ENGEL

L'École des hautes études en sciences sociales, Paris

Professor Engel has written on the notion of truth, philosophy of mind, epistemology and the philosophy of logic. He has advocated a realist theory of truth and a normativist conception of belief. Among his books are *Truth* (Acumen 2002) and *Frank Ramsey: Truth and Success* (Routledge 2002, with Jérôme Dokic). In *What's the Use of Truth?* (Columbia University Press 2007) he engaged in a debate with Richard Rorty.

MONIKA GRUBER

Department of Philosophy, University of Vienna

Dr Gruber's research interest concerns early analytic philosophy and logic, especially Alfred Tarski's formal notion of truth and Frank Ramsey's pragmatism. She is currently working together with Nils-Eric Sahlin, Fredrik Stjernberg and Robin Stenwall in trying to develop a more up-to-date interpretation of Ramsey's distinctively British version of pragmatism.

CONTRIBUTORS

GÖRAN HERMERÉN

Department of Medical Ethics, Lund University

Professor emeritus Hermerén has published books and papers on research ethics and health care ethics, contributed to encyclopedias, been involved in many EU-funded research projects, and served as external examiner in bioethics at the National University of Ireland. He has also served as chair of the European Group on Ethics in Science and New Technologies, Brussels, as chair of the working group for science and ethics of ALLEA (All European Academies), Berlin, and as member of the National Council on Medical Ethics, Stockholm.

KEVIN MULLIGAN

Institute of Philosophical Studies, Università della Svizzera italiana, Lugano

Professor Mulligan has written about ontology and metaphysics (tropes, events, states of affairs), the philosophy of mind (belief, feelings, knowledge, perception) and Austrian thought from Bolzano to Musil and Wittgenstein. In 1984, together with Peter Simons and Barry Smith, he published the paper ‘Truth-makers’ which helped to launch truthmaker theory, a research program at the borders of semantics, logic and ontology that is still very much active today.

JOHANNES PERSSON

Department of Philosophy, Lund University

Professor Persson has worked on topics in metaphysics and philosophy of science, with an interest in what local knowledge and proven experience are and how they relate to scientific evidence in decision-making.

MARTIN PETERSON

Department of Philosophy, Texas A&M University, College Station

Professor Peterson has worked on topics in normative ethics, applied ethics and decision theory. He has been instrumental in the development of engineering ethics and the ethics of technology and has authored several books on these topics, including *The Ethics of Technology: A Geometric Analysis of Five Moral Principles* (Oxford University Press 2017), *The*

CONTRIBUTORS

Dimensions of Consequentialism (Cambridge University Press 2013) and *An Introduction to Decision Theory* (Cambridge University Press 2009).

NILS-ERIC SAHLIN

Department of Philosophy and Department of Medical Ethics, Lund University

Professor emeritus Sahlin is a theoretical philosopher and medical ethicist, focusing on theoretical and applied issues related to decision-making and risk-taking. In particular, he has been interested in the philosophy of Frank Ramsey.

PAUL SLOVIC

Department of Psychology, University of Oregon, Eugene

Professor emeritus Slovic is a leading theorist and researcher studying the psychology of risk and decision-making and has published extensively on topics such as human judgment and risk perception. His work on the phenomenon of *psychic numbing* shows the way to better comprehend the cost of human tragedy and become more compassionate.



Speakers and some of the participants at the symposium on ethics, risks and difficult decision-making, held at The Royal Swedish Academy of Letters, History and Antiquities, 17–18 October 2024. From left to right: Pascal Engel, Martin Peterson, Johan Hirschfeldt, Ben Thierren, Wändi Bruine de Bruin, Fredrik Stjernberg, Lena Wahlberg, Dag Prawitz, Robin Stenwall, Monika Gruber, Anna-Sofia Maurin, Gunn Johansson, Nils-Eric Sahlin, Paul Slovic, Daniel Västfjäll, Johannes Persson, Göran Hermerén, Eva Sjöstrand and Lotta Ericsson. Photo: Staffan Eriand Isa/Kungl. Vitterhetsakademien.

Index

- Agrawal, Arun 69
 Allais, Maurice 127
 Antognazza, Maria 80, 122, 123n
 Aquinas, Thomas 80, 84
 Aristotle 80, 82, 83, 122, 127
 Arrow, Kenneth 30, 37, 39, 40n, 127

 Bassenge, Friedrich 84
 Berkes, Firket 66, 68, 72, 73
 Braithwaite, Richard Bevan 99, 102, 103
 Brentano, Franz 93
 Brook, Ryan 71
 Bruine de Bruin, Wändi 10, 19

 Cook Wilson, John 85, 114, 121, 122

 Debreu, Gerard 127
 Dewitt, Barry 67
 Dokic, Jérôme 123n
 Dretske, Fred 113

 Elga, Adam 132, 133
 Engel, Pascal 11, 109

 Fischhoff, Baruch 128, 130
 Fishburn, Peter C. 129, 135
 Frege, Gottlob 103

 Gärdenfors, Peter 132
 Gettier, Edmund L. 80, 82, 113, 114, 118, 122, 123n

 Goldman, Alvin 113
 Good, Irving John (I. J.) 129
 Goodfield, June 65
 Gruber, Monika 11, 79, 104

 Harsanyi, John 127
 Hawthorne, John 120
 Hellström, Ann 46, 50, 57
 Hermerén, Göran 9, 10, 45
 Holton, Richard 112, 116
 Husserl, Edmund 93, 96, 104n, 105n, 122

 Jackson, Frank 33, 41n
 James, William 93

 Kahneman, Daniel 48–49, 127, 132
 Keynes, John Maynard 93

 Levi, Isaac 124n, 132, 133

 Markowitz, Harry 127
 McLachlan, Stéphane 71
 McTaggart, John (J. M. E.) 99
 Meinong, Alexius 93, 96, 100
 Mies, Richard 14
 Moore, George Edward (G. E.) 93, 99
 Moss, Sarah 114–119, 121, 123
 Mulligan, Kevin 11, 84, 91
 Nelson, Leonard 84

 Newman, John H. 97, 105n
 Nollkaemper, André 31, 32

 Ogden, Pete 23, 24
 Ortega y Gasset, José 98

 Persson, Ingmar 51
 Persson, Johannes 10, 65, 67
 Peterson, Martin 10, 29, 136
 Plato 11, 80–82, 123n

 Radford, Colin 84
 Ramsey, Frank P. 9, 11, 87, 88, 91–99, 101–104, 105n, 111–114, 121, 122, 124n, 127–130, 135, 136
 Reinach, Adolf 84, 85
 Reiner, Hans 105n
 Ross, W. David 37, 38, 41n
 Rumphius, Georg Eberhard 73
 Russell, Bertrand 92, 93, 95, 96, 111–113, 122

 Sahlin, Nils-Eric 7, 9–11, 13, 29, 40, 46, 52, 58, 60, 61n, 67, 103, 113, 119, 121, 123, 127
 Savage, Leonard J. 129–132, 135
 Scheler, Max 84, 94, 97, 98, 100, 103, 105n
 Schwaag Serger, Sylvia 58
 Seidenfeld, Teddy 132
 Sen, Amartya 38, 40n, 41n, 127

INDEX

- | | | |
|--|---|--|
| <p>Slovic, Paul 10, 13, 51, 128,
130</p> <p>Smith, Michael 33</p> <p>Socrates 80–82</p> <p>Stanley, Jason 120, 121</p> <p>Stefánsson, H. Orri 29, 40</p> <p>Stout, George 93</p> <p>Szent-Gyorgyi, Albert 13</p> | <p>Toulmin, Stephen 65</p> <p>Tversky, Amos 48, 49, 132</p> <p>von Hildebrand, Dietrich
84, 85</p> <p>Wahlberg, Lena 67</p> <p>Wallin, Annika 67</p> <p>Watson, John B. 128</p> | <p>White, Alan 97, 100</p> <p>Williams, Bernard 84–87</p> <p>Williamson, Timothy
120–122</p> <p>Wittgenstein, Ludwig
91–94, 98, 99, 102–104,
105n, 111</p> |
|--|---|--|

Should decisions be based on knowledge, defined—as Plato proposed—as true and justified beliefs? In a multicultural society there are frequently different and diverging views about what we know. Our choices are influenced by the alternatives we see, how they are described, and how we suppose that their consequences will affect us and our descendants. If we want to be able to predict preferences, we need to understand the importance of how the problems are presented and how our minds work.

The idea that the rational individual maximizes his or her expected utility when making a decision—derived from Frank Ramsey’s essay ‘Truth and probability’, published in 1931—is an important point of departure in this collection of essays. It comprises case studies of difficult decision-making when faced with challenges, such as nuclear strategy, the climate crisis and care of extremely preterm-born children, as well as perspectives on philosophical issues raised by making rational decisions.



ISBN 978-91-88763-78-5

