
Open
MIND

Thomas Metzinger & Jennifer M. Windt (Eds). *Open MIND.*

Thomas Metzinger & Jennifer M. Windt (Eds).

Open MIND

Frankfurt am Main: MIND Group

To our partners, Stefan Pitz and Anja Krug-Metzinger,

and to all the students and scholars of philosophy and cognitive science

who do not have easy access to scientific literature.

Imprint

Philosophisches Seminar / Gutenberg Research College
Jakob Welder-Weg, 18
Johannes Gutenberg-Universität Mainz
D-55099 Mainz

Production: Satzweiss.com Print Web Software GmbH, Saarbrücken

ISBN: 978-3-95857-102-0

www.open-mind.net

Table of Contents

About this Collection

Introduction to the Open MIND Project

Thomas Metzinger

Author

Thomas Metzinger
metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

1 What is this?

This is an edited collection of 39 original papers and as many commentaries and replies. The target papers and replies were written by senior members of the MIND Group, while all commentaries were written by junior group members. All papers and commentaries have undergone a rigorous process of anonymous peer review, during which the junior members of the MIND Group acted as reviewers. The final versions of all the target articles, commentaries and replies have undergone additional editorial review.

Besides offering a cross-section of ongoing, cutting-edge research in philosophy and cognitive science, this collection is also intended to be a free electronic resource for teaching. It therefore also contains a selection of online supporting materials, pointers to video and audio files and to additional free material supplied by the 92 authors represented in this volume. We will add more multimedia material, a searchable literature database, and tools to work with the online version in the future. All contributions to this collection are strictly open access. They can be downloaded, printed, and reproduced by anyone.

2 What is the MIND Group?

The MIND Group is an independent, international body of early-stage researchers, which I founded in 2003. It is formed of young philosophers and scientists with a strong interest in questions concerning the mind, consciousness, and cognition. They come from various disciplines such as philosophy, psychology, cognitive science, and neuroscience.

Over the past decade, the MIND Group has cooperated with a number of institutions, such as the Frankfurt Institute for Advanced Studies, the *Meditationszentrum Beatenberg*, the *Wissenschaftskolleg zu Berlin*, and the *ICI Kulturlabor Berlin*. I first founded the group at the *Johannes Gutenberg-Universität* in Mainz in 2003, but soon had to relocate it to Frankfurt am Main, where we meet twice a year. Meetings typically involve two or three public lectures at the *Johann Wolfgang Goethe-Universität*, delivered by highly prominent guests, most of whom are now authors of the target papers in this collection and senior members of the group. In addition, our invited speakers offer extended, closed workshops, where advanced students have the opportunity to give short mock-lectures in English.

This format was inspired by a question which kept confronting me in my teaching: namely why are there so many excellent, smart young philosophers in Germany, who nevertheless are—and often remain—almost completely invisible on the international stage? More than half a century after World War II, only three or four German universities rank among the top 100. The established philosophical community is still largely disconnected from many of the latest and most exciting developments in modern philosophy of mind. One result of my thinking about this

Metzinger, T. (2015). About this collection - Introduction to the Open MIND project.
In T. Metzinger & J. M. Windt (Eds). *Open MIND:* 0(T). Frankfurt am Main: MIND Group. doi: 10.15502/9783958571037

problem was that this lack of integration into the global research context was caused, in part, by the language barrier. The biggest psychological obstacles for many young German philosophers seem to be, quite simply, to prepare a talk in English; find the courage to travel to an international conference in another country; and actually present their work there. One of the things we practice at MIND Group meetings is to prepare them for this.

The MIND Group sees itself as part of a larger process of exploring and developing new formats for promoting junior researchers in philosophy of mind and cognitive science. One of the basic ideas behind the formation of the group was to create a platform for people with one systematic focus in philosophy (typically analytic philosophy of mind or ethics) and another in empirical research (typically cognitive science or neuroscience). One of our aims has been to build an evolving network of researchers. By incorporating most recent empirical findings as well as sophisticated conceptual work, we seek to integrate these different approaches in order to foster the development of more advanced theories of the mind. One major purpose of the group is to help bridge the gap between the sciences and the humanities. This not only includes going beyond old-school analytic philosophy or pure armchair phenomenology by cultivating a new, type of interdisciplinarity, which is "dyed-in-the-wool" in a positive sense. It also involves experimenting with new *formats* for doing research, for example, by participating in silent meditation retreats and trying to combine a systematic, formal practice of investigating the structure of our own minds from the first-person perspective with proper scientific meetings, during which we discuss third-person criteria for ascribing mental states to a given type of system.

In addition to bridging geographical and disciplinary gaps, the MIND Group also aims to bridge conventional gaps produced by institutionalized hierarchies in academia. If you will, this is simply the academic variant of the generation gap: Few things are more intimidating to young researchers than being confronted, at a conference, with criticism from a researcher who has long been one of their intellectual heroes, known only from textbooks, university classes, and research articles. For this reason, the MIND Group meetings have provided a protected space for promoting supportive and collegial interactions between senior and junior group members. In particular, the meetings of the MIND Group have helped establish and cement collaborations both among junior members and between junior and senior members. In some cases this has led to research visits, joint research projects, or long-term mentoring relationships. One motivation for founding the group, after all, was to smooth the path from university studies to being a professional academic for advanced students and young researchers.

3 Why did we do this?

We wanted to make a contribution by offering a freely available resource to others. When we first started thinking about what to do for the 20th meeting of the MIND Group, we knew we wanted it to be something special, some way of sharing with the interested academic public some of the expertise and collegial atmosphere we had built up over more than 10 years of working together. Initially we considered inviting everyone to a big four-day conference at an attractive location. But then we decided that we would do something more substantial and innovative - rather than creating a transient event and an enormous CO_2 footprint. We wanted to create a resource of lasting value that will subsist for years to come, and most importantly something that really is accessible for everybody—not only for people in affluent parts of the world, like ourselves. There seemed no better way to do this than by providing a large, open-access collected edition showcasing the work of our senior and junior members.

It quickly became clear that because of the scope of the project, and also because we had specific ideas about how it should be realized, this was going to be an experiment in autonomous open-access publishing. The MIND Group is an independent body, and apart from evening lectures by our invited speakers, its meetings are not open to the public. One goal of the Open MIND project was to first publish our scientific work without the support of a publisher, who would

eventually sell our own intellectual property back to us and our peers and simultaneously make it inaccessible to students in Brazil, India or China by locking it behind a paywall. We wanted to see if we could successfully establish a professional form of quality control via a systematic, journal-independent peer review process—and also if we could make it happen faster than existing and established institutions of academic publishing. We gave authors a deadline of 1st March 2014, and planned to publish the entire collection (including commentaries and replies) on January 15th 2015. We knew that these two pillars—speed and quality control—would be crucial to the success of the project. Academics are sometimes reluctant to publish their work in edited collections that often only appear years after the manuscripts have been submitted. We suspected that we would only succeed in obtaining state-of-the-art research papers if we could guarantee that the research discussed within them would not be out-of-date by the time the collection went online.

This publication format is also novel in another sense. Because a selected subset of junior group members acted as reviewers and commentators, the whole publication project is *itself* an attempt to develop a new format for promoting junior researchers, for developing their academic skills, and for creating a new type of interaction between senior and junior group members. Many of the reviewers and commentators in this edited volume have never actively participated in any scientific review process before, and, for many their commentary is their first ever publication. Throughout the project, all junior members were able to play different roles: they acted as reviewers, trying to improve and constructively criticize the target articles submitted by senior group members and commentaries submitted by their peers. Sometimes, reviewers were asked to go back and revise their reviews—and sometimes their reviews also led to the rejection of target papers altogether. They also acted as authors; and because their commentaries also went through a review process, they got to experience the review process from the other side as well.

This collection, therefore, is the result of a three-layered interaction between junior and senior members: personal (through meetings), editorial (through implementing a common publication project), and philosophical and scientific (through writing commentaries and replies). Throughout this process, we were often surprised and impressed by the results—and we hope that you will be, too.

4 Who did this?

Many people have made this contribution possible and many hours of unpaid work have gone into it. Here are the most important supporters.

4.1 The editors

As founder and director of the MIND group, I consider myself to be neither a junior nor a senior member. Therefore, I have not contributed a target paper or a commentary. If anything, my contribution lies in the choice and selection of authors and in the work, together with my collaborator Jennifer Windt, of bringing this project to completion.

4.2 Financial funding

All in all it has cost about € 241.000, to realize this project. First and foremost, the *Barbara Wengeler-Stiftung* needs to be mentioned: not only has it supported the current project with € 80,000, but over the years it has enabled the MIND Group to stay independent, and to realize a long series of fruitful meetings, during times when it was difficult to get support elsewhere. It has also supported some members by providing PhD and travel grants and by offering the annual € 10,000 *Barbara Wengeler*-Prize, awarded at our meetings in Frankfurt. The Gutenberg Research College and the *Vokswagen-Stiftung* have generously supported the project by providing two editorial staff positions for David Baßler, Daniela Hill, and Dr. Ying-Tung Lin, and by awarding a five-year Research Fellowship, beginning in April 2014, to me, Thomas Metzinger. This work was also partly supported by the European FP7 collaborative project VERE (contract no. 257695).

What Does it Mean to Have an Open Mind?

Thomas Metzinger & Jennifer M. Windt

Authors

Thomas Metzinger
metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt
jennifer.windt@monash.edu
Monash University
Melbourne, Australia

1 Instead of an introduction

In our discussions leading up to the Open MIND collection's going online, we thought long and hard about how exactly to showcase the vast material in this collection and the ideas and motivations behind the project in our editors' introduction. We first thought about using the introduction to briefly summarize the take-home message of every single target article, commentary and reply, as is customary in introductions to edited collections. This struck us, however, as being both unwieldy and redundant: It would have entailed summarizing and commenting on a total of 117 texts. More importantly, due to the online format of the collection (including in-text search functions) and the inclusion of abstracts and keywords in the papers themselves, the authors have already provided concise introductions to their own texts. Retracing their steps in an editorial introduction would not have added to the value and usability of the collection.

We then considered using the introduction to create our own personal best-of-Open-MIND list, discussing what we take to be the very best ideas and most valuable insights in every single article, or perhaps even focusing on the contributions that we personally take to be the most theoretically important. Though our own list of personal favorites seemed to flow naturally out of the editing process, this strategy quickly struck us as being at odds with the motivations driving the collection. Using the editors' introduction to create a personal best-of list would have been highly selective and biased by our own personal research interests and styles in a way that we felt would contradict our own ideal of open mindedness.[1]

These considerations naturally gave rise to a more difficult and more profound question: What exactly do we mean by open mindedness, not just in general, but in the context of interdisciplinary research on the mind? The strategy of using the contributions to the Open MIND collection as a foil for the more general academic variant of open mindedness was tempting. But we quickly realized that this approach would again strike many readers (as well as, perhaps, some of our own authors) as highly idiosyncratic and perhaps even arbitrary or self-important.[2]

1 In fact, for this reason, we intentionally omit any references to the contributions to Open MIND in this introduction.

2 This is not, of course, to deny that we take "Open MINDedness" (as broadly practiced in the context of this collection) to be an example of "open mindedness" as a more general epistemic stance. And we are certainly proud enough of what we like to think of as our little star-collection to allow ourselves at least a few words on why we think this is the case. To begin with, on many levels, Open MIND was an exercise in edit-

Metzinger, T. & Windt, J. M. (2015). What Does it Mean to Have an Open Mind?
In T. Metzinger & J. M. Windt (Eds). *Open MIND:* 0(C). Frankfurt am Main: MIND Group. doi: 10.15502/9783958571044

So we decided to use our editors' introduction to do something more humble and, in a way, more classically philosophical. Here, we want to briefly address a deeper and more difficult problem: that of what genuine open mindedness really is and how it can contribute to the Mind Sciences. The material in the collection speaks for itself. Here, and in contrast to the vast collection that is Open MIND, we want to be concise. We want to point to a broader context for a particular way of thinking about the mind. And we want to introduce an idea, namely that of what open mindedness could mean in the context of the contemporary, interdisciplinary Mind Sciences. This variant of open mindedness is characterized by epistemic humility, intellectual honesty and a new culture of charity. It also has a pragmatic dimension: open mindedness of this kind is research generating and fosters an environment of sincere and constructive interdisciplinary collaboration. It is also profoundly inspired by the classical ideals of philosophy as a pursuit of genuine insight and rational inquiry, the importance of a critical and in a certain sense non-judgmental attitude, and the deep relationship between wisdom and skepticism as an epistemic practice. And finally, and again very classically, open mindedness has an ethical dimension as well: It implies sensitivity to normative issues, including ones of an anthropological, sociocultural, and political kind. By bringing these different strands of ideas together and creating a big (and admittedly still sketchy) picture of what open mindedness might mean in the interdisciplinary Mind Sciences, we hope to begin a conversation on how an open minded attitude and a charitable culture of collaboration can be cultivated in the future. This is very much intended as an invitation to further think about and develop this topic. We hope our readers join us in this endeavor.

orial open mindedness. The authors and commentators asked to contribute to this collection were explicitly encouraged to discuss any topic they themselves thought to be relevant. The only restriction, in terms of content, was that the target articles fall within the scope of the Mind Sciences. We also tried to foster a particular type of intellectual atmosphere by encouraging authors, commentators and reviewers to be consistently constructive and charitable. Our hope was that this approach would bring out the best in our contributors in the different stages of the project. In many cases, we explicitly encouraged our authors to write in a way that would be accessible to readers from different academic backgrounds and to take different disciplinary perspectives into account. Generally, the publication of academic articles always involves a process of give and take between authors, editors, and reviewers. And we strongly felt that it would be a good indicator of the success of our collection if at the end of the day, our authors were themselves happy and proud of their contributions. This entailed carefully calibrating our own roles as editors and in many cases leaving the final decision to our authors. Finally, and perhaps most importantly, our choice of the title Open MIND also reflects the that by introducing a two-way interaction between senior target authors and junior commentators through the commentaries and replies, we wanted to help our junior group members and commentators enter into a discussion with more senior and prominent representatives of the field. All of these points – as well as the free availability of the Open MIND collection to students and researchers from anywhere in the world, free of charge – exemplify theoretical and practical dimensions of what we consider to be academic open mindedness. In addition, many of the papers published here explore new ways of thinking, in the broadest sense, about the mind and new and innovative ways of driving forward research.

2 Open mindedness as an epistemic stance

Open mindedness is not a theoretical position but an epistemic practice. Clearly, there are many different kinds of open mindedness, and the precise way of characterizing the relevant kind of open mindedness will depend on the subject matter in question, or, more simply, on what it is that one is open minded about. Yet as a first pass at a definition we might say that open mindedness, in its most general sense, is characterized by epistemic humility and adherence to a general ideal of intellectual honesty. This is true for open mindedness in general, but also for the specific variants we are interested in here, namely open mindedness in academic research, including interdisciplinary scientific discourse on the mind.

Whatever else it may be, open mindedness is also an *attitude* that is now shared by a growing number of researchers in philosophy of mind, cognitive science, neuroscience and artificial intelligence (AI). We are all interested in the deep structure of the human mind and of conscious experience, but we also recognize how far we are still away from a unified theoretical model that would satisfy philosophers and scientists alike, a model that is conceptually convincing and able to integrate all existing data and make use of different methods at the same time. We do not want to fool ourselves. Although great progress has been made during the last five decades, it is not at all clear which

combination of methods and which type of theoretical approach will generate the final breakthrough or even facilitate epistemic progress. We, meaning researchers of different stripes and from different disciplines comprising the Mind Sciences, including the authors contributing to this collection, are all in the same boat: We share a common epistemic goal, and we find ourselves in a major historical transition. Progress in the empirical sciences of the human mind is certainly impressive and is continuously gaining momentum, generating large amounts of new and sometimes surprising data. At the same time, exciting new approaches in formal modeling and philosophical meta-theory are increasingly opening up new perspectives. Yet it is not at all clear that we are already asking the right kinds of questions and exactly what combination of conceptual and empirical tools will do the trick. Seeing this fact clearly has already begun to change our attitude. Researchers from different disciplines are listening and talking to each other in new ways. Developing news forms of interdisciplinary collaboration[3] is an integral part of this process. "Having an open mind" also refers to a kind of scientific practice that involves honestly listening to representatives of exactly those approaches and academic disciplines you may not have expected to make a contribution.

At the same time, open mindedness, understood as a fruitful and research generating epistemic practice, should be clearly distinguished from arbitrariness, indecisiveness, lack of specificity, and, especially in the context of philosophy, lack of conceptual precision. Open mindedness is not just any kind of openness, and it is different from simply being non-committal or from hedging. The challenge is to develop an understanding of open mindedness that is guided by theoretical considerations and empirical research findings alike. Ideally, this account of open mindedness should also suggest specific strategies for cultivating forms of sincere interdisciplinary collaboration, sharpening the underlying conceptual issues, and developing precise predictions for future research. Open mindedness of the epistemically fruitful type will often be more about asking better questions than about committing to specific answers. It will involve an attitude of willingness to question or even reject one's own prior commitments. It will be inherently critical (cf. Lambie 2014). And it will, perhaps, have more to do with striving for genuine understanding than with the search for truth and knowledge (Taylor 2014). One core idea of the great philosopher of science Karl Popper, which is now reappearing in the latest mathematical theories of brain functioning, was that we are always in contact with reality at exactly the moment at which we falsify a hypothesis: The moment of failure is exactly the moment at which we touch the world.[4] Similarly, the best scientific theories will be those that most easily lend themselves to falsification. For this reason, open mindedness involves, among other things, endorsing very specific theoretical positions purely for the sake of epistemic progress, rather than for the sake of being right, advancing one's career, publishing in a high-impact journal, and so on. Open mindedness is not so much about the specific content of a belief, be it personal or theoretical, but about the way in which it is held.

Open mindedly searching for the right kind of question is a good first-order approximation to the specific type of attitude we are trying to describe. Another way to characterize the epistemic stance one might call "open mindedness" is to say that it is an interdisciplinary variant of the principle of charity. It is not just that philosophers should be empirically informed or that neuroscientists should listen carefully to constructive attempts at conceptual or methodological clarification. We need to develop a new culture of charity, and this will require new and sustainable forms of interdisciplinary cooperation. In philosophy, the "principle

3 And the same is true, of course, for intradisciplinary collaboration.

4 Here is what he said about the fundamental principle of any ideological form of rationalism turned weltanschauung: "Uncritical or comprehensive rationalism can be described as the attitude of the person who says 'I am not prepared to accept anything that cannot be defended by means of argument or experience' . . . Now it is easy to see that this principle of an uncritical rationalism is inconsistent; for since it cannot, in its turn, be supported by argument or by experience, it implies that it should itself be discarded". (cited from Popper 2013, p. 435; originally in Popper 1945/2003; see Metzinger 2013c for a popular discussion).

of charity" has long been recognized and investigated as reading others' statements according to their best, strongest possible interpretation, that is, to never attribute irrationality, falsehoods or fallacies without necessity. But we also all know how hard this can be. Agreement should be optimized and as each other's interpreters, we should always, when possible, prefer the most coherent reading in order to maximize the truth or rationality of what another researcher (or philosopher) says. We now need an interdisciplinary variant of this principle, and not only in bridging the gulf between the humanities and the so-called hard sciences of the mind, but also in organizing novel and more efficient forms of cooperation. This point applies not only to the relationship between disciplines, but also to that between different generations of researchers. An optimization problem has to be solved: What is the best way of pooling intellectual resources and of efficiently structuring research? Therefore, a second step in approximating the undogmatic attitude to which we refer in the title of this collection is to characterize it as an openness to the possibility that, for mind and consciousness, there may be no such thing as a single dominating discipline, no *Leitwissenschaft*, as we say in German. Rather, not only the connectivity between already existing research programs has to be strengthened, but the overall pattern of scientific practice will have to be given a new internal structure as well. What is needed is a new and, as we will argue later, genuinely philosophical way of thinking.

A genuine receptiveness to new or unexpected ideas and disciplinary perspectives also presupposes a certain set of abilities. One of them is tolerance of ambiguity: to not only tolerate transient cognitive, conceptual and theoretical inconsistencies between disciplines or generations, but to view certain kinds of ambiguity as actually desirable, as a source of progress. Again, the challenge will be to distinguish productive types of ambiguity from ones that are overly cautious or vague, thus hampering real progress. The same is true, of course, within academic disciplines themselves. Academic disciplines are not natural kinds. Contrary to what some might think, there might be no single authoritative or right way of doing philosophy, and there may be no clean way to distinguish philosophy from the empirical sciences. Open mindedness of the constructive kind will not waste time by worrying too much about disciplinary demarcation criteria or mere labels, but will also be open towards different methods and approaches within individual disciplines. Put differently, it may turn out to be less important whether a given question or position is philosophical (in the sense, perhaps, of being of a purely conceptual nature) or empirical than whether it genuinely helps advance the overall debate. Open mindedness thus clearly also has an inherently pragmatic dimension. When this kind of tolerance of ambiguity, for instance towards disciplinary borders, but also towards different (and, ideally, complementary research methods) is paired with conceptual clarity and precision, it can be turned into a driving force for research. This balancing act is what academic open mindedness, or so we claim, is all about.

3 Open mindedness and the phenomenology of (un)certainty

Having an open mind involves, among other things, a specific way of being noncommittal with respect to the truth of a proposition. As pointed out earlier, this is not the same as hedging: one can investigate and even defend the truth of a proposition or the adequacy of a given theoretical-conceptual or empirical model while at the same time acknowledging that it might be false. In the history of the philosophy of science, this continued openness to the falsifiability of a scientific hypothesis, often associated with attempts to bring about specific ways of establishing and testing its falsity, is commonly regarded as a marker of good scientific practice. It is the core of intellectual honesty. As Russell tells us, "intellectual integrity [is] the habit of deciding vexed questions in accordance with the evidence, or of leaving them undecided where the evidence is inconclusive" (2009, p. 579). The moment at which we give up this openness, we lapse into dogmatism. The real

danger, says Russell, is never the content of a doctrine, be it religious or political, but always "the way in which the doctrine is held" (Russell 2009, p. 582). Of course, this connection between wisdom and the recognition of not-knowing is much older (Ryan 2014). In the *Gorgias*, Socrates explicitly claims that he is happy to be refuted if he is wrong. In fact, he claims he would rather be refuted than to refute someone else because it is better to be delivered from harm oneself than to deliver someone else from harm. And in the *Apology* (21d), after being accused of blasphemy and of corrupting the youths of Athens, Socrates famously said before the tribunal of 501 Athenians: "I neither know nor think that I know".

This is not the place to enter into a discussion of open mindedness in the context of philosophy of science or of the history of philosophical theorizing about wisdom. Because we are interested in open mindedness in the context of the interdisciplinary Mind Sciences, we do, however, want to draw attention to a related point: open mindedness as an epistemic practice involves a specific kind of mental attitude and is closely related to certain kinds of phenomenal states. Cultivating the relevant kinds of phenomenal mental states and epistemic feelings makes a real difference, or so we suspect, for scientific research by facilitating the development of a research climate that is conducive to constructive and genuinely fruitful discourse and new forms of collaboration. This is an empirical prediction, which might be false. For now, our claim simply is that exactly the kind of open mindedness described here is needed to even begin investigating its truth. If at the end of the day, this strategy should turn out to fail – that is, if there turn out to be good empirical reasons for rejecting the claim that there actually are specific phenomenological profiles and mental attitudes that decisively facilitate progress in interdisciplinary research on the mind – this would be a valuable insight. But this insight about the value of open mindedness in scientific discourse itself depends on an initial willingness to cultivate exactly the kind of epistemic practice in question.

If this is right, there is another reason to be interested in open mindedness in the present context. This is that open mindedness, as an epistemic practice and mental attitude, is itself a potential target for interdisciplinary consciousness research. Philosophy of mind in particular can contribute by laying the theoretical-conceptual groundwork for the further empirical investigation of open mindedness in academic life and proposing points of contact with psychology and cognitive neuroscience. To make this inner connection more clearly visible, let us briefly sketch the outlines of such an account.

Where might one begin investigating open mindedness as a mental state? At the outset, it stands to reason that the relevant form of open mindedness has precursors in the history of philosophy and might also be interestingly related to current debates on philosophical methodology. After all, the principles of epistemic humility, intellectual honesty, charitability, and searching for more accurate questions while cultivating a productive form of tolerance of ambiguity are deeply rooted in the history of philosophy. On a systematic and more general level, one would expect philosophy, as the discipline traditionally most concerned with the status of knowledge and truth and the practice of inquiry itself, to be able contribute to an analysis of what open mindedness really is. Based on these considerations, let us have a short look at four main questions: One, what is the relationship between open mindedness and intuitions? Two, what is the relationship between open mindedness and the tradition of philosophical skepticism? Three, what would answers to the first two questions tell us about the relationship between open mindedness and the allegedly most pressing problem for interdisciplinary consciousness research, the subjectivity of phenomenal mental states? Might we even use the analysis of open mindedness to formulate principles for the investigation of phenomenal states and the status of first-person data? And four, how can open mindedness as an epistemic stance be related to ethical and practical and questions? In particular, how can it contribute to normative issues related

to neurotechnological interventions in the human brain, and on how to cultivate new forms of interdisciplinarity?

3.1 Intuitions and the phenomenology of certainty

The concept of intuition has a long philosophical history and is also firmly rooted in everyday language and folk psychology.[5] Intuition, in everyday language, refers to immediate and direct insight independent of reflection or instinctively grasping or sensing a matter of fact. In the history of philosophy, the concept of intuition often has dual epistemic and experiential readings, and this is true for the traditions of rationalism and empiricism alike. In the *Rules for the Direction of the Mind* (Rule 3), Descartes describes intuitions as an immediate, effortless and indubitable kind of seeing with the mind, which is even more reliable than deduction. In his *Essay Concerning Human Understanding* (IV.II.I), Locke tells us that intuition involves a direct perception of ideas that is, once more, the basis of all forms of knowledge. The close relationship between intuitions and sensory perception, and especially seeing, is already evident in the Latin verb *intueri*, which means to look and observe, but also to examine or consider. The central underlying element is the immediacy and directness of perception, which is imported into the concept of intuition via an implicit analogy between the phenomenology of sensory perception and genuine insight in an epistemic sense.

At the same time, the epistemic status of intuitions as well as different ways of defining the concept of intuition are a matter of controversy in the current debate on philosophical methodology. The debate on intuitions stands at the center of the confrontation between classical (and allegedly intuition-based; see for instance Cappelen 2012 for a critical discussion) conceptual analysis conducted in the proverbial philosophical armchair and recent claims that experimental philosophy, mostly by collecting layperson's responses to questionnaires (often involving vignettes inspired by well-known philosophical thought experiments; for discussion, see Knobe & Nichols 2008; Alexander 2012; for a general introduction to intuitions in philosophy, see Pust 2014), offers a new, empirically based method for collecting intuitions. According to some experimental philosophers (for discussion and further references, see Alexander & Weinberg 2007), these results even cast doubt on the reliability of intuitions as a mark of philosophical expertise more generally.

Here, we would like to propose a definition of intuitions that is compatible with the historical literature on intuitions as well as phenomenologically and empirically plausible. Departing from our brief remarks on the history of intuitions in philosophy, we might say that intuitions are the "phenomenal signature of knowing", a seemingly direct and effortless way of perceiving or seeing with one's mind independently of a prior process of reflection. This analogy between intuiting and perceiving provides an entry-point, or so we claim, for a naturalized concept of intuition. But it also suggests a potentially dangerous equivocation between the phenomenological and the epistemological readings of the concepts of intuition. If the phenomenology of intuiting is indeed similar to that of perceiving in virtue of its effortless and seemingly direct experiential quality, then this immediately invites the problem that the phenomenology of intuiting and perceiving can be deceptive: what seems, subjectively, to be a case of veridical perception can always turn out to be a hallucination or an illusion (for an introduction to the problem of perception, see Crane 2014), or a nocturnal dream (see Windt & Metzinger 2007; Metzinger 2013a; Windt 2015). Similarly, what seems to bear the marks of genuine insight can always turn out to be an epistemic illusion.[6]

If intuitions are indeed mental states characterized by a specific phenomenology, this suggests that the attempt to simultaneously char-

5 This section draws on arguments first presented in Metzinger & Windt (2014).

6 For a striking case study of two patients who experienced strong feelings of subjective certainty, including religious beliefs, during epileptic seizures, see Picard (2013). This is all the more interesting as these beliefs contradicted the patients' longstanding convictions, but still, subjectively, seemed entirely convincing to them during the seizures themselves.

acterize them as involving genuine insight and as the basis of knowledge rests on what elsewhere we call the *"E-error"*: a category mistake in which we ascribe epistemic properties to something that does not possess them (Metzinger & Windt 2014, p. 287). If our account of intuitions is on the right track, then intuitions are dangerous, because in virtue of their phenomenology and of their possessing an occurrent conscious character of "insight", they predispose us to believe certain propositions merely on the basis of seemingly "understanding" them. The phenomenology of intuitions is such that it immediately and effortlessly creates a bias towards accepting the truth of propositions that we simply *know* or feel to be true, while simultaneously preventing us from seeking further justification because these truths also seem unconstructed, indubitable, and self-evident. If so, then one of the factors underlying intuitions and intuitive plausibility is that because of their phenomenal character, they prevent an open minded assessment. Intuitions turn us into inner dogmatists. And this is true not only for individual propositions held to be intuitively true, but also for the continued adherence to theoretical claims about the status of intuitions as a guide to or even basis of knowledge and genuine insight. The phenomenal character of intuitions even predisposes us towards certain meta-theoretical intuitions about the general epistemic status of intuitions, and we can see the marks of this throughout the history of philosophy as well as in the contemporary debate (e.g., Bealer 1998; Chudnoff 2013). The analysis of intuitions clearly should not itself be driven by intuitions. Instead, the analysis of intuitions is a prime example of where an open mind is needed.

Our own account starts out from the assumptions that intuitions are a specific class of phenomenal states. Human beings can direct their introspective attention towards the content of these states and, at least partly and under certain conditions, report about it. Many higher animals very likely also possess intuitions even if they are not able to directly attend to or verbally report their intentional contents. Before the evolution of biological nervous systems and before the emergence of phenomenal consciousness, no intuitions existed on our planet. Patients in coma or human beings in unconscious, dreamless deep sleep have no intuitions in the sense intended here. At the same time, intuitions likely have a long evolutionary history: There must have been a point in time in which the first intuition appeared in the mind of some conscious organism and this specific type of inner state then propagated itself across thousands of generations while its functional profile became ever more differentiated. Plausibly, one could describe the having of intuitions as an *ability* – a mental ability that was adaptive and was acquired gradually.

If one takes the phenomenal character of intuitions seriously, this ability clearly seems to be an epistemic ability: *Prima facie,* to have an intuition means to have the subjective experience of knowing something, directly and immediately, without necessarily being able to express this knowledge linguistically or to provide an epistemic justification. Typically, inner experience seems to present knowledge to the subject of experience, even if one does not know *how* and *why* one possesses this knowledge. Intuitions are the phenomenal signature of knowing, a seemingly direct form of "seeing" the truth with one's own mind. The *E- error* then consists in confusing the phenomenal character a conscious state with its epistemic status. "Epistemicity", the phenomenal quality of "knowingness", or the feeling of being a knowing self, as such is only a phenomenal quality, just as redness, greenness or sweetness are. The well-known philosophical problem is that the phenomenological and epistemological readings can always come apart, because what phenomenologically appears as a kind of perception could really be a hallucination or an illusion. Subjectively indistinguishable mental states do not necessarily have the same epistemic status. Trivially, the difference between veridical perception and hallucination (in the philosophical sense; see Macpherson 2013; Crane 2014;) is not available on the level of subjective experience itself, and therefore the confusion between phenomenal character and epistemic content is naturally grounded in the transparent phenomenology, the seeming directness and immediacy of

sensory perception. The same is true for the phenomenology of intuition. Conflating epistemic status and phenomenal character becomes particularly dangerous if it is imported into theoretical debates, and if the phenomenal quality in question is that of "epistemicity", of direct and non-inferential knowing itself. The important lesson is that *as* phenomenal states, such states are neither necessarily veridical nor necessarily non-veridical, because experience as such is not knowledge. *As* inner experiences they possess no intentional properties and cannot be semantically evaluated by concepts like "truth" or "reference". Phenomenal transparency is not epistemic transparency.

Many, but not all, of our philosophically relevant intuitions are characterized by an additional element of *certainty*, of *just knowing* that one knows. Here, the phenomenal signature of knowledge does not only refer to the content of what is seemingly known in a direct, and non-inferential manner, but to our higher-order, subjectively experienced knowledge itself. This means that the phenomenal character of "epistemicity" that accompanies and tags the respective mental content as an instance of knowing has itself become transparent. Its representational character is not introspectively available anymore: the fact that epistemicity is itself the content of a non-conceptual mental representation, that it is internally constructed and always contains the possibility of misrepresentation, is veiled by an experience of immediacy. Transparency is a special form of darkness. Something constructed is experienced as a *datum*, as something given. Therefore, in stable intuition states we not only experience the first-order content as directly given, but the epistemicity of the state itself. Let us call such states *intuitions of certainty*. Referring to G. E. Moore[7] one might say that the phenomenal signature of knowing has itself become diaphanous or transparent: according to my own subjective experience I simply *know* that I know and the possibility of error and falsehood it is not given on the level of conscious experience itself. From the fact that a conscious perception instantiates the phenomenal quality of "greenness" it does not follow that the underlying process or even the perceptual object are green. Analogously, the same is true for the "phenomenal signature of knowing" that characterizes intuitions.

Intuitiveness is a property of theoretical claims or arguments, relative to a class of representational systems exhibiting a specific functional architecture. Conscious human beings are one example of such a class. The brains of human beings are naturally evolved information-processing systems, and when engaging in explicit, high-level cognition they use specific representational formats and employ characteristic styles of processing. Whenever we try to comprehend a certain theory, an argument or a specific philosophical claim, our brains construct an internal model of this theory, argument, or claim (Johnson-Laird 1983, 2008; Knauff 2009). This mostly automatic process of constructing mental models of theories possesses a phenomenology of its own: some theories just "feel right" because they elicit subtle visceral and emotional responses, some claims "come easily", they are sound and healthy, and some arguments (including implicit assumptions they make in their premises) seem "just plain natural". Some forms of skepticism appear "healthy" to us, others do not – there seems to be a deep connection between sanity and reason.

There may be two overarching reasons for this well-known fact. First, theories that are intuitively plausible exhibit a high degree of "goodness of fit" in regard to our network of explicit prior convictions and, more generally, by optimally satisfying the constraints provided by our conscious and unconscious models of reality as a whole. These microfunctional constraints implicitly represent both the totality of the knowledge we have acquired during our lifetime, as well as certain assumptions about the causal deep structure of the world that proved functionally adequate for our biological ancestors.

7 In *The Refutation of Idealism*, G. E. Moore wrote: "The term 'blue' is easy enough to distinguish, but the other element which I have called 'consciousness' - that which a sensation of blue has in common with a sensation of green – is extremely difficult to fix… And in general, that which makes the sensation of blue a mental fact seems to escape us; it seems, if I may use a metaphor, to be transparent – we look through it and see nothing but the blue; we may be convinced that there is something, but what it is no philosopher, I think, has yet clearly recognized." (1903, p. 446)

Theories that immediately feel good because they are characterized by a high degree of intuitiveness maximize a specific kind of internal harmony. What we introspectively detect is a high degree of consistency, but in a non-linguistic, subsymbolic medium. Therefore we could also replace the term "intuitiveness" by a notion like "intuitive soundness" or "introspectively detected consistency or goodness of fit" (relative to a preexisting model of reality). In principle it should be possible to spell out this point on a mathematical level, by describing the underlying neural computations and their properties in a connectionist framework, or by utilizing the conceptual tools provided by dynamical systems theory or predictive coding.

A second perspective might be to look at intuitions not from a representationalist, but from biophysical point of view. We are embodied beings, and there are different levels of embodiment. Computational, but also thermodynamical imperatives guide the self-organization of representational states in our brains. One major causal factor underlying the conscious experience of "intuitive soundness" might simply be the amount of energy it takes to activate and sustain a mental model of a given theory, plus the amount of energy it would take to permanently *integrate* it into our pre-existing model of reality. Our mental space of intuitive plausibility can in principle be described as an energy landscape: claims that "come easily" do so because they allow us to reach a stable state quickly and easily, theories that "feel good" are theories that can be appropriated without a high demand of energy. Theories that *don't* feel good have the opposite characteristics: they "don't add up", they "just don't compute", because they endanger our internal harmony and functional coherence, and it would take a lot of energy to permanently integrate them into our overall mental model of reality. They are costly. However, in a biophysical system like the human brain there may well be a direct connection between thermodynamic efficiency and reduction of complexity on the level of information processing. If biological self-organization in the human brain takes place in a way that continuously attempts to minimize the prediction error generated by the flow of "hypotheses" originating in its current model of reality, then the process that creates what today we call our deepest "theoretical intuitions" may also be described as the continuous attempt at reducing variational free energy. While, on a more abstract level, this process can be seen as minimizing representational complexity while simultaneously maximizing the evidence for the overall model, it is at the same time a physical process that is not guided by abstract rationality constraints, but simply one that optimizes metabolic and statistical efficiency at the same time (Sengupta et al. 2013; Friston 2010; Hohwy 2013).

We need an open mind, because many of the best future theories about the human mind and conscious experience may just "not compute" for beings like us. However, what does or does not compute is, in part, a contingent fact determined by the functional architecture of our brain, shaped by millions of years of biological evolution on this planet, as well as—to a much lesser degree—by our individual cognitive history and a given cultural/linguistic context. The phenomenology of intuitive soundness—the fact that some arguments seem "just natural"—is a biological phenomenon, with a short cultural history supporting it as well. However, the inner landscape of our space of intuitive plausibility is not only contingent on our evolutionary history and on certain physical and functional properties of our brains—it was also optimized for *functional adequacy* only. It serves to sustain an organism's coherence and physical existence, but this does not mean that the *content* of intuitions is epistemically justified in any way. This is especially true because the evolved functional adequacy of intuitions applies to everyday action in practical contexts—not to abstract reflection in theoretical ones or cognitive environments. This is why searching for a comprehensive theory of the conscious mind presents such a major challenge to our intellectual honesty: It demands that we investigate a claim even if it contradicts our deepest intuitions, something that somehow is "just too radical", way too costly, painful or even self-damaging, and seems to cry out for a more moderate, weaker version