



Report on the **Futures Fellow Program**

Project Title:

Rethinking Suicide Risk

Author

Dr. René Baston
Eltlingstrasse 50
45141 Essen

Mail: rene.baston@tu-dortmund.de

Technische Universität Dortmund
Instituts für Philosophie und Politikwissenschaft
Emil-Figge-Str. 50
44227 Dortmund

Raum 2.249

Date: January 26, 2026

Content

1. Introduction.....	1
2. Brief description of the project	2
3. Description of the research stay.....	3
4. Results	9
5. Conclusion	13

1. Introduction

Hosted by UA Ruhr's North America office, the Futures Fellow program is a competitive fellowship for researchers affiliated with the University Alliance Ruhr's three universities. It reflects UA Ruhr's commitment to global engagement and its goal of strengthening the transatlantic science corridor. Building on the Research Alliance launched in 2021 to advance internationally oriented research, the program supports durable collaborations between UA Ruhr institutions, research centers, and partner universities in the United States across a wide range of disciplines. Futures Fellows are encouraged to pursue projects at a partner university of their choice, exchange knowledge and expertise, and build interdisciplinary networks that improve research quality through collaboration and shared resources. The fellowship particularly aims to support early to mid-career researchers, including junior professors, by enhancing their professional portfolios, expanding international grant and research opportunities, and increasing the visibility of UA Ruhr work through public presentations. Above all, it is designed to create ongoing, reciprocal relationships that lead to sustainable, long-term partnerships.

I had the opportunity to participate in the Futures Fellow program, which enabled me to devote three full months to my project. For this period, I was granted unpaid leave, and I rearranged my teaching obligations into a single block seminar, allowing me to focus entirely on the fellowship. During my stay, my official host and supervisor was Professor Brian McLaughlin.

During my Futures Fellow stay, I focused on extending my Normic account of suicide risk, first developed in "Beyond Prediction: A New Paradigm for Understanding Suicide Risk" (Synthese, 2024). This work responds to a well-documented challenge in clinical psychology: meta-analyses indicate that many standard risk factors have limited predictive value, and factor-based assessment models often produce clinically unhelpful results, including high false-positive rates. Rather than treating risk as something that must be grounded in probabilistic prediction, I argue for an alternative framework that reframes suicide risk in terms of normalcy and intelligibility within a possible-worlds approach. On this view, risk increases when suicidal action becomes easier to render intelligible against norms of practical rationality and the balance of reasons for living and reasons for dying, including cases where life-sustaining reasons are fragile or absent. The project also connects this reconceptualization to broader philosophical debates about non-probabilistic notions of risk, drawing on historical work suggesting that reasoning about risk predates modern probability theory and highlighting how a normic framework can better capture the explanatory and normative structure of risk situations.

2. Brief description of the project

Standard suicide risk assessment in clinical psychology is in trouble: large meta-analyses suggest that the usual risk factors have very limited predictive power, and clinicians cannot reliably distinguish who will attempt suicide from who will not. Even when people are classified as “high risk,” false positives are overwhelming, and many suicides still occur in so-called lower-risk groups. That leaves clinicians with very little that is genuinely action-guiding for an individual case.

My project develops an alternative that does not treat “risk” as a probability estimate. The Normic account reframes risk in a possible-worlds framework, where what matters is how normal an outcome is, given relevant norms. Applied to suicide, the key idea is that risk rises when suicidal action becomes easier to make sense of in light of norms of practical rationality and the person’s reasons for living and reasons for dying. On this view, someone can be at high risk when life-sustaining reasons are weak or absent, even if standard factor-based models fail to flag them.

This approach also connects suicide research to broader philosophical debates about whether risk can be understood in objective, non-probabilistic terms. Historically, people reasoned about risk before modern probability theory, which supports the thought that “risk” is not just about likelihood. A normic, possible-worlds approach aims to capture what is distinctive about risky situations: the explanatory and normative structure that makes an outcome feel close at hand, because it would not take much to see how it could happen.

In the book *Against the Gods*, Peter Bernstein notes that probability theory was only developed in the seventeenth century, even though people had long been forced to make decisions under risk. In maritime trade, for example, investors still had to judge whether a voyage was worth backing, and they did so by drawing on the best information available rather than on statistics. Bernstein describes how reports from sailors familiar with a route, as well as letters and other firsthand accounts, could serve as key sources for assessing how risky a journey might be. These judgments were clearly forms of risk assessment, but they operated without probability calculus or modern statistical tools. Against this background, it is plausible that the concept of risk captures more than its contemporary technical meaning. My project takes this thought seriously by investigating and refining alternative, non-probabilistic notions of risk, with the aim of showing how such concepts might illuminate clinically relevant questions in suicide risk assessment.

3. Description of the research stay

The stay took place from 09.01.2025 to 14.12.2025. Professor Brian McLaughlin gave me access to his office, and I went to the University about 2 times a week. The following table provides an overview of the research trip:

13.10.2025	<ul style="list-style-type: none">• First time in the office, met Juan Comesana, Jonathan Schaffer, Dean Zimmerman• Rutgers Annual lectures in Philosophy, Prof. Susanna Siegel (Oct. 13, 15 & 16) from Harvard 2-4pm at the Alexander Teleconference Hall, room 403
15.10.2025	<ul style="list-style-type: none">• Meeting with Albert Newen und Sofia Predrini for lunch• Visiting the second talk of Prof. Susanna Siegel
16.10.2025	<ul style="list-style-type: none">• Preparing the presentation for the workshop on Friday (17.10)
17.10.2025	<ul style="list-style-type: none">• Participated at the Bochum-Rutgers Mini-Workshop 2025 on "Belief and Imagination", with Jenny Wang (Rutgers), Tania Lombrozo (Princeton), and Christian Scholz (RUB).• Gave the talk "Can Belief Explain Implicit Bias? Not Without a Theory of Reasoning"
21.10.2025	<ul style="list-style-type: none">• Met at Rutgers with Carolina Sartorio and discussed my draft "The Hidden Core"
23.10.2025	<ul style="list-style-type: none">• Met with Louis Sass Professor in the Department of Clinical Psychology at Rutgers
24.10.2025	<ul style="list-style-type: none">• Participated at the Value Reading Group with Carolina Sartorio and Frances Kamm
28.10.2025	<ul style="list-style-type: none">• Met with Priya Nayar at the New York Office to talk about the scholarship program
02.11.2025	<ul style="list-style-type: none">• After reworking my draft, I submitted the paper "The Hidden Core" at the Journal of Moral Philosophy, which I discussed with Carolina Sartorio.
03.11.2025	<ul style="list-style-type: none">• Lunch and coffee with Preston Lennon, Haley Linscott, and Luke Quinton
06.11.2025	<ul style="list-style-type: none">• Meeting and lunch with Brian McLaughlin at Rutgers

11.11.2025	<ul style="list-style-type: none"> Met with Dean Zimmerman at Rutgers to discuss possible world formalism.
21.11.2025	<ul style="list-style-type: none"> Presentation of the Normic Account of Suicide Risk in the Value Reading Group
26.11.2025	<ul style="list-style-type: none"> Met at Rutgers with Brian McLaughlin
03.12.2025	<ul style="list-style-type: none"> Christmas Event at the German Embassy
10.12.2025	<ul style="list-style-type: none"> Met again with Professor Louis Sass at Rutgers Went to the Department's Christmas Party
14.12.2025	<ul style="list-style-type: none"> Flying back to Germany

17.10.2025

- Participated at the **Bochum-Rutgers Mini-Workshop 2025 on "Belief and Imagination"**, with Jenny Wang (Rutgers), Tania Lombrozo (Princeton), and Christian Scholz (RUB).
- Gave the talk "Can Belief Explain Implicit Bias? Not Without a Theory of Reasoning"

The workshop "Belief and Imagination" focused, as the title suggests, on two core themes, and one of them was the mental state of belief. In my talk, I examined whether implicit prejudices can count as beliefs, and argued that the answer depends in part on how a mental state responds to reasons.

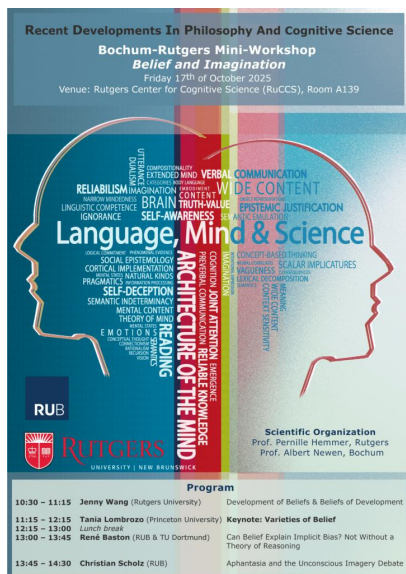


Image 1: Flyer of the Workshop

Beliefs are usually taken to be guided by facts and to change when the evidence changes, which is why philosophers often characterize belief as reason-sensitive. I suggested that implicit prejudices do not fit this profile. Empirical work shows that implicit evaluations can be produced under laboratory conditions using evaluative conditioning, where a neutral stimulus is repeatedly paired with a positive or negative stimulus. A classic example would be pairing the sound of a bell with an electric shock, so that the bell comes to trigger a negative automatic response. My question was: why should this count as something like a rationally grounded attitude toward the bell? It is obvious that the bell is not the

cause of the shock, so there is no good reason for the subject to develop a negative stance toward

bell sounds. That mismatch between the attitude and the reasons that could justify it supports the conclusion that implicit prejudices are not beliefs.

21.10.2025

- Met at Rutgers with Carolina Sartorio and discussed my draft “The Hidden Core”

I drew on feedback from **Carolina Sartorio**, whom I know from my research stay at the University of Arizona in Tucson, where she was a professor at the time. She has since moved to Rutgers University, and we met to discuss one of my drafts because I knew she is an expert on the topic. Carolina offered sharp, critical comments on several metaphysical issues that the text raises but did not sufficiently address, including questions about the nature of values. In particular, she pressed me to clarify whether



Image 2: Carolina Sartorio and me

the relevant values should be understood in a constructivist way as grounded in our attitudes and practices, or in a realist way as independent of our perspective.

In the paper, I argue against a Strawsonian assumption that the value of close relationships depends on anger-like reactive attitudes such as resentment. In intimate contexts, I suggest, resentment is often a secondary emotion that covers up more basic primary emotions like fear

of rejection, hurt, shame, or longing. Because resentment can escalate conflict and distort understanding, it is a poor candidate for what supposedly “keeps” relationships meaningful. Instead, I propose that what really sustains the value of close relationships is vulnerability: being willing to make those primary emotions explicit, which better expresses how much we care.

To support this, I draw on ideas from Emotionally Focused Therapy, where relationship distress is often traced to unmet attachment needs and defensive interaction cycles. On this view, expressing primary emotions can improve meaningfulness, fulfillment, and stability by inviting reassurance, responsiveness, and mutual support, whereas resentment tends to provoke defensiveness. The paper also argues that accountability does not require resentment. Partners can still call each other to account through non-resentful forms of blame or forward-looking moral protest that aim at protection, repair, and moral growth rather than payback. The upshot is that we can preserve what matters in close

relationships while making room for responsibility practices that do not threaten the relationship's value.

6.11.2025

- Meeting and lunch with Brian McLaughlin at Rutgers



Image 3: Me in Brian's office

In my meeting with Brian McLaughlin, I presented my approach to risk, focusing in particular on suicide risk. I explained why I am drawn to non-probabilistic risk concepts that work with possible worlds, such as Duncan Pritchard's modal account of risk, and I suggested that these frameworks are not essentially about prediction. In medicine, I argued, risk matters largely because resources are limited and must

be allocated, so those at high risk are precisely those who should receive immediate support and intervention. Brian challenged a key step in this line of thought. He pointed out that I was moving too quickly from the claim that a risk concept is non-probabilistic to the conclusion that it is non-predictive. On his view, switching to possible-worlds talk does not, by itself, solve the underlying problem that prediction still does not work in a clinically reliable way. If prediction remains out of reach, then it is not clear that the new framework puts us in a better position than traditional assessments, which also fail to predict individual behavior yet can still be used, in practice, to guide how scarce resources are distributed.

This discussion led me to focus more directly on the question of how risk might be understood without treating it primarily as a claim about the future. In response, I began to develop a way of thinking about risk that does not rely on forecasting individual outcomes. The preliminary conclusions of this shift are summarized in the "Results" section (A), where I record the main ideas and open questions that emerged from this reorientation.

11.11.2025

- Met with Dean Zimmerman at Rutgers to discuss possible world formalism.

Duncan Pritchard's idea of *modal risk* is not about how likely something is, but about how easily it could happen. In terms of "nearby" possibilities: an outcome is riskier if you can get from the actual situation to a situation where it happens with only small changes. That is why a bomb that explodes if a lottery ticket wins can count as riskier than a bomb that explodes only if a bizarre chain of events occurs, even if both have the same probability. The lottery outcome is a normal way things could go, while the bizarre chain would require the world to be very different.

A problem with this approach is that it treats what is actually true as maximally risky, simply because the actual world is the closest possible world to itself. If the wall in your house really contains asbestos, then on the modal view the risk is automatically maximal, which is unhelpful because you wanted to assess the risk before you know the truth. The *normic account of risk* tries to fix this by saying that risk depends on what would be normal given your evidence. If, based on what you know, asbestos would be surprising (abnormal), then the risk can be low even if asbestos is in fact there. Me and Dean Zimmerman talked about the fact that even the normic view can run into a similar issue if it lets one especially "normal" scenario decide everything. We discussed a practical fix (see Results C).

21.11.2025

- Presentation of the Normic Account of Suicide Risk in the Value Reading Group

In that session of the reading group, I gave a presentation about my normic account of suicide risk. The account treats risk as a matter of how close a person is to suicidal action in terms of intelligibility, not in terms of statistical likelihood. On this view, someone is at higher risk when, given what we know about their reasons for living and dying, theorist can make a suicidal's persons state of mind intelligible through norms of practical rationality. The key is the person's balance of reasons for living and reasons for dying, and how stable those life-sustaining reasons are under pressure. If reasons for living are fragile or absent, then suicidal action becomes easier to understand as a "normal" next step within a nearby possible scenario, even if standard risk-factor models would not predict it. The goal is to circumvent the problem that suicidal action is, in fact, not predictable, which is why some clinical psychologists are suggesting to drop the notion of suicide risk altogether.

The discussion revealed that one major problem of my approach is how I conceptualized suicidality and its relation to the agent's reasons for living and dying. In my approach, "reasons for living" are connected to identity-constituting projects and unconditional desires, which makes suicidality look like a one-way slide that begins once these projects collapse. That picture struggles with what clinicians and first-person reports emphasize most, namely that people can still report meaning, attachment, and reasons to live while nonetheless feeling pulled toward death. It also leaves too little room for the

affective side of suicidality, especially the way pain, hopelessness, relief-seeking, and emotional exhaustion can become action-guiding even when a person's evaluative outlook is not simply 'life has no meaning' (see Results **B**).

4. Results

During the research stay, I made numerous advances with my research project on normic suicide risk. The key results are listed below.

A) Risk: Non-predictive approaches

I have considered three different ways in which we can think about risk. First, *predictive risk* is the conception most familiar from statistics, economics, and everyday forecasting. Here, risk is understood in terms of likelihood: it concerns how probable it is that some undesirable event will occur in the future. Weather forecasts, medical prognoses, and financial risk models all presuppose this framework. The central question is epistemic and forward-looking: given the evidence available, how likely is the bad outcome? On this conception, risk assessment is primarily a matter of modeling uncertainty, updating beliefs, and improving prediction. Examples for this type of risk:

- “There is a 60% risk of rain tomorrow.”
- “If you undergo this surgery, there is a significant risk of complications.”

Fact-relative risk is not about what we can predict but about what is actually the case. It concerns present conditions that objectively ground reasons for concern, independent of anyone’s degree of uncertainty. When an engineer says a bridge is unsafe, or when a health authority says that asbestos in a building poses a serious risk, the judgment is anchored in features of the current state of affairs. Even with complete information, the situation would remain risky, because the risk is grounded in how the world is, not in ignorance about what will happen. This conception therefore treats risk as a normative-evaluative category tied to norms. Examples for this type of risk:

- “There is a real risk from asbestos in these walls.”
- “This playground equipment is dangerous in its current condition.”

Vulnerability risk shifts the focus from events and conditions to the position of agents. Here, to say that someone is “at risk” is to say that they are insufficiently protected or exposed in ways that make certain harms easier to inflict upon them, even if no particular harmful event is probable. Undocumented workers are at risk of exploitation and marginalized communities are at risk from environmental hazards. These are not primarily predictive claims about what will happen, nor merely descriptions of present hazards, but diagnoses of structural arrangements that generate asymmetries of power and protection. This conception of risk is inherently normative. Examples for this type of risk:

- “Migrant workers are at risk of exploitation because they lack legal protection.”

- “Communities without access to healthcare are at risk during public health crises.”

The vulnerability risk conception looks especially promising for applying it in clinical psychology. On a *vulnerability-based conception adapted to clinical psychology*, suicide risk should be understood as an assessment of how exposed and fragile the patient is given their situation and their reasons. The clinician’s task begins with reconstructing the agent’s lived standpoint: what pressures the person experiences, what losses or conflicts structure their world, what resources, relationships, or commitments still support them. Risk assessment then proceeds normic rather than statistical: from the standpoint of practical rationality, one asks whether, given these perceived reasons and constraints, suicidal action would be intelligible or make sense to the agent. Where suicide appears as a coherent option within the agent’s deliberative landscape, vulnerability is high (high-risk). On this view, risk is not a forecast but a normative evaluation of exposure and fragility, and it grounds professional responsibility: because clinicians occupy a role defined by duties of care, they are obligated to respond to such vulnerability with protective intervention (if justified and necessary) and support.

B) Reasons for living and dying, well-being, and suicidality

I have reconsidered how to deal more appropriately with reasons for living and dying. Those reasons should not be treated as items we can count, subtract, or balance on a mental scale. They are treated as the kinds of considerations that make a person’s stance toward their life intelligible, both to themselves and to others. The central thought is that suicidality typically involves an overall evaluation like “life is not worth living,” which is not just a report of how things are going (like, I am not happy) but a global negative assessment of whether continuing makes sense. From there, I insist on a phenomenologically important distinction: in most suicidal states, the person is not primarily drawn by the thought that death is valuable in itself, but pushed by the sense that living has become unbearable or no longer worth sustaining. So, reasons for dying that matter most are usually reasons that speak against continuing to live, such as chronic pain, hopelessness, entrapment, or a collapse of meaning, rather than reasons that treat death as a goal with its own value. This also clarifies why suicidal ambivalence, being torn between life and death, is so common. A person can have serious, even compelling reasons to go on living while simultaneously experiencing strong pressures that make rejecting life seem understandable, especially when suffering crosses a threshold and the remaining goods of life feel inaccessible.

To make this clinically and theoretically usable, I draw on Alan Goldman’s book *Values of Life*. It treats pleasure, happiness, meaning, and the pursuit of central goals as key sources that normally make a life feel worth continuing. When these dimensions are absent or inverted, a negative life-evaluation

becomes understandable, and suicidal thinking starts to look coherent rather than mysterious. Importantly, the aim is not to compute a total score. It is to interpret patterns and their coherence: some profiles fit together in a way that makes a suicide attempt sadly intelligible, while other profiles raise puzzles that call for further explanation, such as hidden despair, masked symptoms, or disruptive psychopathology. This is why *explanatory effort* is important in this context: The more a potential suicide coheres with the person's evaluative pattern across well-being dimensions, the less explanatory work we need to understand it; the more it clashes with that pattern, the more unexpected it is, and the more we should look for what the initial picture is missing.

C) A pragmatic way of dealing with possible worlds for normic risk

According to Hirvelä and Paterson (2024), the normic account of risk is subject to the *privileged world problem* "according to which a proposition's truth value at a single world can make it maximally risky" (Hirvelä & Paterson, 2024, p. 8). The problem mirrors the problem of the modal account: risk is determined by the status of a single privileged world rather than by the overall modal profile of the proposition. On the normic account, the risk of a proposition is fixed by the normalcy of the most normal world in which that proposition is true. This means that if a proposition holds at a maximally normal world, no proposition can be riskier than it, irrespective of how exceptional the proposition may be elsewhere. Since a proposition need only be true at a single such world, contingent propositions can be assigned maximal risk simply in virtue of holding at a privileged world, namely a most normal world. This result conflicts with the intuitive thought that maximal risk should be reserved for necessary truths or for propositions that obtain across a wide range of possibilities.

I think there is a pragmatic way to solve the privileged world problem for the normic account of risk. Possible worlds, as I see it, are tools of thinking, such as irrational numbers. We can use such tools to solve theoretical problems, and we can use them in a variety of ways. Currently, the following paradigm is at play:

1. Fix an ordering over worlds
2. Identify the best (closest / most normal) world W where p holds
3. The world W determines the risk of p

The privileged world problem is not inevitable, but it is a consequence of a particular design choice. The privileged world problem arises because risk is fixed by a single world, which is maximally ranked as normal or similar. The following approach avoids this, because there is no maximally normal world doing the decisive work. The actual world functions as a reference point, and risk depends on

explanatory deviation (which indicates some distance to Lewisian similarity orderings (Lewis, 1973, 1986)).

Here is a different way to look at the normic account of risk:

1. Fix the actual world as the center
2. Partition modal space into coarse-grained risk regions (such as high/mid/low)
3. For a given proposition p , imagine a world where p holds
4. Assess how much explanatory work is required to accommodate p
5. Place that world in the appropriate risk sphere

Instead of finding the most normal world in which p holds, possible worlds are constructed from the actual world by adding p . The task is, then, to place that constructed world into a risk sphere. Thereby, no world is privileged in itself, and risk is not fixed by truth, but by explanatory costs. Thereby, the privileged world problem does not appear for the normic account of risk.

5. Conclusion

During the fellowship, the three major results of my project came together into a more defensible version of the normic account of suicide risk. I began by addressing a serious theoretical challenge to normic risk, namely the privileged world problem. The worry was that normic risk might be artificially “maxed out” by the truth of a proposition at a single especially normal world, which would make risk assessments unstable and conceptually distorted. Building on discussions that were significantly shaped by Dean Zimmerman’s feedback, I developed a pragmatic solution that changes how possible worlds are used in the account. Instead of letting one privileged world determine the risk of a proposition, risk is assessed by starting from the actual world and asking what explanatory changes would be required, given the evidence, to accommodate the proposition. Risk is then located within coarse-grained regions (high, mid, low) according to explanatory cost. This removes the problematic dependence on a single maximally ranked world and clarifies the methodology: possible worlds function as disciplined tools for exploring intelligibility relative to evidence.

This strengthened framework also helped me respond to an important worry raised by Brian McLaughlin. I had initially suggested that once risk is decoupled from probability, it is no longer about prediction. Brian’s point was that this inference is not justified: operating with possible worlds does not automatically remove the forward-looking dimension, and it does not by itself solve the problem that individual prediction fails in suicide assessment. This pushed me to broaden the conceptual landscape and to articulate non-predictive conceptions of risk that can do genuine clinical work. In particular, I adapted *vulnerability risk* to the context of clinical psychology. On this view, suicide risk is best understood as an assessment of how exposed and fragile a person is with respect to radical self-harm, given their circumstances and their deliberative situation, rather than as a forecast of what they will do. Vulnerability risk is inherently normative: it tracks whether the person lacks protections, resources, or stabilizing commitments. This aligns naturally with normic risk because the assessment relies on norms of practical rationality and ethical norms of care to judge when a person’s situation constitutes a special kind of danger that demands professional response, regardless of any attempt to estimate likelihood.

Finally, the vulnerability-based reconceptualization becomes operational through a revised understanding of reasons for living and dying. Rather than treating reasons as items to be counted or weighed, I treat them as the considerations that make a person’s overall stance toward life intelligible. Suicidality, on this picture, grounds in a global evaluation such as “life is not worth living,” and the crucial clinical task is to understand how that evaluation is supported by the person’s lived experience. To make this interpretation more systematic, I connected reasons to broad dimensions of well-being,

drawing on Alan Goldman's account of well-being. He suggests that what normally makes life good is pleasure, happiness, meaning, and the pursuit of central goals. Suicidal states can then be understood as cases where these dimensions are absent, inverted, or experienced as inaccessible. This yields a clear link back to normic risk: the easier it is to render suicidal action intelligible given the person's perceived reasons and constraints, the higher the normic risk. And it links directly to vulnerability: when this intelligibility comes easily, the person is especially exposed, even if we remain agnostic about what they will do in the future. The overall result is an improved normic account of suicide risk that is theoretically more robust, clearer about its method, and, prospectively, better suited to guiding clinical responsibility under conditions of limited medical resources.