



# The division of cognitive labor and the structure of interdisciplinary problems

Samuli Reijula<sup>1</sup> · Jaakko Kuorikoski<sup>2</sup> · Miles MacLeod<sup>3</sup>

Received: 8 February 2022 / Accepted: 16 May 2023 / Published online: 10 June 2023  
© The Author(s) 2023

## Abstract

Interdisciplinarity is strongly promoted in science policy across the world. It is seen as a necessary condition for providing practical solutions to many pressing complex problems for which no single disciplinary approach is adequate alone. In this article we model multi- and interdisciplinary research as an instance of collective problem solving. Our goal is to provide a basic representation of this type of problem solving and chart the epistemic benefits and costs of researchers engaging in different forms of cognitive coordination. Our findings suggest that typical forms of interdisciplinary collaboration are unlikely to find optimal solutions to complex problems within short time frames and can lead to methodological conservatism. This provides some grounds for both reflecting on current science policy and envisioning more effective scientific practices with respect to interdisciplinary problem solving.

**Keywords** Division of cognitive labor · Interdisciplinarity · Collective problem solving · Modularity

## 1 Introduction

Interdisciplinarity in its many forms is strongly promoted in science policy across the world. It is seen as a necessary condition for providing practical solutions to pressing complex problems for which no disciplinary approach alone holds all the required

---

✉ Samuli Reijula  
samuli.reijula@helsinki.fi

Jaakko Kuorikoski  
jaakko.kuorikoski@helsinki.fi

Miles MacLeod  
m.a.j.macleod@utwente.nl

<sup>1</sup> Theoretical Philosophy, University of Helsinki, Helsinki, Finland

<sup>2</sup> Practical Philosophy, University of Helsinki, Helsinki, Finland

<sup>3</sup> Philosophy Section, University of Twente, Enschede, The Netherlands

answers (Klein, 1990). At present interdisciplinarity is widely incentivized through funding instruments which typically take the form of 3–6 year research projects. These instruments are both provided by funding agencies and increasingly used to structure internal university funding (Lindvig & Hillersdal, 2019; Salmela et al., 2021). In this article we formally model this type of interdisciplinary problem solving as a division-of-labor problem. Our goal is to provide a basic representation of the epistemic dynamics of this type of problem solving, and to explore the relative benefits and costs of employing different cooperative or integrative strategies.

Our motivation for turning to modeling in these cases is to gain insight into how the cognitive structure of interdisciplinary problems—which require more than one type of cognitive expertise to adequately resolve them—might govern the approaches and solutions researchers choose (and ought) to take up, given the potential costs of engaging in collaboration. The costs here include the time and missed opportunities for producing solid disciplinary results. As other authors have noted, the outcomes of interdisciplinary projects often fail to be as integrated as an optimal solution might seem to require (Bruce et al., 2004; MacLeod & Nagatsu, 2016; Mennes, 2018; Salmela et al., 2021) and interdisciplinary research is often perceived as being shallower than allegedly more rigorous disciplinary research (Adams et al., 2007; Larivière & Gingras, 2010; Leahey et al., 2017). While many aspects of interdisciplinarity have been extensively discussed, to our knowledge, no one has tried to investigate the general structure of interdisciplinary problems based on the basic forms and patterns those problems can take irrespective of the subject matter or fields involved. Analyzing this structure (and its variations) promises us potential insight into the costs and benefits of different cognitive problem-solving strategies across a range of cases, and thus the costs and benefits of interdisciplinary problem-solving strategies overall.

Beyond these general observations, theorizing about the potential benefits and costs of interdisciplinarity on a purely conceptual level alone is not an appropriate method for examining the *trade-offs* and the conditions under which the costs of interdisciplinarity may outweigh the benefits (Reijula & Kuorikoski, 2019). Empirical case studies and quantitative analyses of research activities are essential for estimating the effectiveness of solution-oriented, short-to-medium-term interdisciplinary collaborations, but such empirical work ought to be informed by analytical theorizing about the possible mechanisms responsible for the observed results. Yet, purely verbal theorizing is notoriously unreliable when it comes to the possible outcomes of multiple interacting mechanisms and feedback processes. The formal models discussed in this article aim to contribute to such analytical foundations for theorizing about the social production of knowledge.

Many of the existing models in the social epistemology of science portray science as competitive problem solving: all agents compete at solving the same problem (usually defined by single-dimensional epistemic utility). Such models can capture some benefits of epistemic diversity, but are ill-equipped to explore what we call *collaborative division of cognitive labor* (D’Agostino, 2009). Collaborative division of cognitive labor requires *cognitive coordination* that goes beyond the question of how to allocate resources between a set of competing research hypotheses. In the interdisciplinary interactions studied in this article, groups of researchers come together to solve a complex problem beyond their individual competences. This (temporary)

collective must identify and assign its members with appropriate tasks. Such *task partitioning* and differentiation of epistemic roles presupposes that the research problem addressed by the community is at least partly decomposable (von Hippel, 1990). By decomposability we mean that the original problem must be divisible to subtasks so that the epistemic contribution of each subtask is at least partly invariant with respect to changes in the states of the other subtasks (Wimsatt, 2007, Ch. 9).

The general idea that search and discovery proceeds by dividing problems into partially independently solvable subtasks has a long pedigree in philosophy and the cognitive sciences. Most notably, Simon (1962) emphasized the importance of approximations to near-decomposability in understanding complex systems. In philosophy, this general idea has been pursued by Wimsatt (2007), who stresses the value of *reductionist heuristics*—working hypotheses that the whole is merely the sum of its parts—as tractable guides to learning about the organization and interaction of the parts. Drawing on Simon and Wimsatt, Bechtel and Richardson (2010) suggested decomposition and localization as key heuristics in the search for mechanistic explanations in the cognitive neurosciences.<sup>1</sup> Although our approach in this article is largely compatible with these accounts, our notion of decomposition does not refer to the causal-mechanical structure of systems, the relationship between a material system and its parts. Instead, we focus on ways in which an epistemic task can sometimes be decomposed into subtasks like a question can be unpacked into a set of more detailed questions.

We need a model template that allows us to study task partitioning and can be used to weigh the epistemic benefits and costs under different forms of cognitive coordination. Such a template should allow us to capture at least the following key elements: (1) recursive division of a problem into sub-problems or subtasks, (2) allocation of sub-problems to different problem solvers, and (3) different strategies for combining the solutions of sub-problems. At the same time, for the sake of transparency and avoiding unnecessary artifacts, the model should be as simple as possible. The function of the model is to help keep the logic of our argument straight, provide definitions for previously intuitively understood concepts and ideas, and make explicit our previously implicit assumptions—not to create realistic and detailed *in silico* scientific communities with which to experiment in the hope of entirely new empirically interpretable “results” (Kuorikoski & Reijula, 2019).

In the following section we introduce an approach to modeling that meets these needs, originally developed in economics by Luigi Marengo and coauthors (Marengo & Dosi, 2005; Marengo et al., 2000) and by Scott Page and coauthors (Hong & Page, 2001; Page, 1994, 1996). We establish distinctions between different levels of complexity and different forms of cognitive coordination. We illustrate the scenarios with examples from real-world interdisciplinary research. In Sect. 4, we employ simulation methods to derive qualitative results on the effectiveness of different forms of collaboration for problems exhibiting different levels of complexity. In light of our findings, in the final sections, we discuss the epistemic implications of these results. Our main aims are twofold: first, to lay out the model template and demonstrate its

---

<sup>1</sup> For related approaches to scientific discovery in the philosophy of science, see Nickles (1980) and Darden (1997). And for parallel discussions in organization science, see, for example, references in Marengo (2015), Leahey et al. (2017) and in engineering, Sobieszczanski-Sobieski et al. (2015).

potential for modeling collaborative problem solving, and, second, to show how the model can explain the perceived shortcomings of interdisciplinary research as consequences of rational responses to the prevailing institutional incentives for multi- and interdisciplinary collaboration. The ambition of this article is thus to provide a proof-of-concept model and use it to make explicit an empirical worry about transient interdisciplinary projects. Inspired by the model, we suggest new *hypotheses* about possible modifications to these incentives and to problem-solving strategies that might improve interdisciplinary research.

## 2 A model of collaborative division of cognitive labor

To clarify the distinction between cognitive diversity and division of labor (see Sect. 1), we approach the scientific research process as follows. We can think of complex problems as multidimensional tasks consisting of several interdependent components (Newell & Simon, 1972). For instance, the task of designing a guitar amplifier consists of the subtasks of designing the amplifying circuit, the speaker, and the enclosure. Each of these subtasks can further be divided into simpler subtasks, often allocated to experts in different fields. Likewise, a game of chess can be divided up into a sequence of moves, but the usefulness of a move depends on the other moves in the game. The game is difficult because such dependencies are difficult to grasp. A key aspect of the difficulty of a problem is thus the extent to which solving a particular subtask depends on the state of the other subtasks (Marengo & Dosi, 2005).

The analysis of problem solving as search applies to the scientific research process. Scientists generally look for formulations of problems, experimental designs, patterns in data, mechanisms behind the data, and implications of their theories (Simon, 1989). All such activities can be viewed as problem-solving tasks consisting of various interdependent sub-problems. Here the components of the problem include choices concerning the background theories to rely on, which instrumentation and methods to use, answers to particular empirical questions, and so on. When combined, the solutions to the sub-problems constitute an answer to the original problem. Just as in chess and the engineering tasks referred to above, the epistemic value of such an answer is not a simple aggregate of the values of its components, instead, there are often important interdependencies between the solutions offered to sub-problems.

The first aspect of our model is a computational representation of a scientific research problem. As a starting point, note that a problem can consist of several parts, and it can have better and worse solutions (cf. Nickles, 1981). Arguably we can use a *binary string* of length  $N$  as a simple representation of a problem consisting of multiple parts. We define each locus in the string as an atomic component problem, and the possible groupings of individual bits as alternative ways in which the complex problem can be subdivided into a set of sub-problems. Whereas  $N$  could be seen to stand for *problem size*, we define the *complexity of the problem* as the number and strength of interdependencies between the solutions to these component problems, as they contribute to the epistemic value of the overall solution.

These ideas can be made more precise by introducing some notation. The problem comprises of a sequence of *components*,  $S = \{s^1, s^2, \dots, s^N\}$ , where  $s^i \in \{0, 1\}$ . The

index set  $I = \{1, \dots, N\}$  consists of the indices of the individual loci. A *configuration*, or a possible solution to the problem, is a string  $x_i = s^1, s^2, \dots, s^N$ . The set of possible configurations we denote by  $X = \{x_1, \dots, x_{2^N}\}$ .

The configurations can be ordered according to their *epistemic value*. We write  $x_i \succeq x_j$  whenever  $x_i$  is (weakly) epistemically preferred to  $x_j$ . Given this notation, a problem is defined by the pair  $(X, \succeq)$  (Marengo & Dosi, 2005). In essence, the problem is this ordered set of solutions, or at least, is fully determined by it (cf. Nickles, 1981).

Now, as the number of components  $N$  increases, the number of configurations grows exponentially ( $2^N$ ). This means that even for problems consisting of a moderately small number of component parts, an exhaustive search of the problem space is not feasible, and agents equipped with bounded cognitive resources must devise strategies for managing problem complexity.

In order to focus on complexity management strategies, the agents' representation of the problem must be distinguished from the problem itself. For now, let us assume a simple encoding where each configuration in  $X$  has a unique representation in the agents' shared language  $L = \{l_1, \dots, l_{2^N}\}$ , where each string  $l_i$  corresponds to the configuration  $x_i$ ; and the subjective valuation function  $V$  (shared by all agents) respects the objective ordering  $\succeq$ .

A well-known strategy for managing complexity is heuristic search. Heuristic search often functions iteratively by modifying the existing solution  $l_i$  according to a rule in order to produce a new solution candidate  $l_j$ . In other words, a heuristic is a mapping  $\phi : L \rightarrow L$ . The decision rule embodied in the heuristics we study follows the hill-climbing logic: If the epistemic value of the candidate solution is strictly higher than that of the current solution,  $V(l_j) > V(l_i)$ , it is adopted as the new solution. Otherwise the candidate is discarded.

As an example of heuristic search in a multidimensional binary space, consider the model introduced by Hong and Page (2001) that implements diversity in problem-solving heuristics as (group) agents possessing different *flipset heuristics*. A flipset can be thought of as a bit mask, where the bits set to '1' flip the state of the corresponding bit in the target string. For example, the flipset '001' applied to the target string '101' flips the rightmost bit, resulting in the string '100'. In Hong and Page's model, each agent has a small set of such heuristics. A set of heuristics gives rise to a set of states  $N_{l_i} = \{l_j, \dots, l_k\}$  which can be reached from  $l_i$  by a single application of one of the agent's heuristics. We call  $N_{l_i}$  the *neighborhood* of  $l_i$ . Consequently, each distinct set of flipset heuristics results in a characteristic set of possible paths that the agent can follow in the search space. The motivating intuition underlying the well-known diversity-beats-ability theorem (Hong & Page, 2004) is that, under appropriate conditions, the diversity provided by the larger set of heuristics possessed by a random group of problem-solvers is epistemically more beneficial than a less diverse set of individually high-performing expert heuristics.<sup>2</sup>

Notice, however, that such heuristic diversity does not yet amount to division of labor as described in Sect. 1. In the scenario outlined above, the members of a problem-solving group all address the same full problem (of length  $N$ ), albeit with a diverse set of heuristics. In the context of scientific problem solving, this would mean, implausibly,

<sup>2</sup> See Reijula and Kuorikoski (2021, 2022) for a critical examination of Hong and Page's results.

that each scientist addresses the full complex problem on their own, but with different epistemic resources. In order to remedy this shortcoming and to implement genuine division of labor in the model, some more terminology is needed.

Let us call a non-empty subset of the index set,  $b \subseteq I$ , a *block*, and its cardinality  $|b|$  the size of the block. A *partition*  $P$  can be defined as a set of blocks that together cover the all the components of the problem:  $P = \{b_1, \dots, b_k\}$  such that  $\bigcup b_i = I$ .

Division of cognitive labor can now be portrayed as different blocks being allocated to different agents, or more realistically, to different monodisciplinary research teams. In the following, we view a scientific discipline as a group of agents characterized by a set of shared heuristics that manipulate a given block. This corresponds to the common idea that single disciplinary perspectives can only address limited aspects of complex scientific problems. For example, a possible division of labor between two disciplines could involve the first addressing the first  $k$  bits of a problem of size  $N$ , and the second addressing the remaining  $N - k$  bits.<sup>3</sup> To represent this, it is useful to define a *block configuration*  $b_k(x_i)$  as the substring of  $x_i$  of length  $|b_k|$  containing the components of the string that belong to the block  $b_k$ . By  $b_{-k}(x_i)$  we denote the remaining bits in  $x_i$ , and consequently,  $x_i$  can be written as a concatenation of the two substrings:

$$x_i = b_k(x_i)|b_{-k}(x_i)$$

Following Marengo and Dosi (2005), we define the size of a partition  $|P|$  as the size of its largest block.

### 3 Forms of cognitive coordination for different types of problems

The agents addressing a problem are not necessarily—or even usually—aware of its true block structure, i.e., the way and extent to which it can be effectively decomposed into independent subtasks. We call the subjective representations of blocks *covers* (Page, 1994). An exhaustive set of covers (corresponding to the notion of partition) we call a *decomposition*,  $D$ . Manipulating the problem decomposition provides a powerful collective strategy for addressing complex problems.

**Remark 1** For a decomposed problem, the size of the problem space to be covered is not  $2^N$ , but instead, in the order of  $2^{|D|}$

In the limit, a problem decomposition of size one,  $D = \{\{1\}, \{2\}, \dots, \{N\}\}$ , which divides the complex problem into a union of sub-problems of minimal complexity, can be solved in linear time (in  $O(N)$  steps). The maximally decomposed problem is tackled one bit at a time, with the hope that the contribution of the atomic solutions to the overall problem are independent, and that the epistemic value of the overall solution is an aggregate of its component solutions. As the decomposition size increases, the heuristics become more holistic.

<sup>3</sup> An alternative approach would be to constrain the language  $L_i$  of agents belonging to discipline  $i$  so that only those bits of  $X$  that are indexed by the block  $b_i$  would be represented in  $L_i$ .

The decomposition heuristic, however, offers no free lunches, as simplifying the search may reduce the quality of solutions discovered. The following examples illustrate situations where problems of different complexity constrain the adequacy of disciplinary division of cognitive labor.

### 3.1 Very simple and very complex problems

We have thus far only discussed ways of dividing a task into individual subtasks. Whether the problem can be solved in such a piecemeal manner, i.e., whether the problem itself really is effectively decomposable, is a separate issue. Let us consider a problem of length  $N = 4$ . For the sake of simplicity—and without loss of generality—we can assume that problem-solving begins from the epistemically least valuable solution state  $x_{min} = '0000'$ , and the correct, epistemically optimal state is  $x_{max} = '1111'$ .<sup>4</sup>

Now, consider a problem of minimal complexity where the solution of an atomic sub-problem contributes to the total epistemic payoff does not vary with respect to changes in the solutions of the other sub-problems. If the problem is in this way *fully modular*, application of a flipset switching any single bit from '0' to '1' will give an agent an unambiguous signal that the overall solution has been improved. Such a signal can be encoded, for example, in a value function defined simply as the number of '1's in the offered total solution. More generally, we can now see that it is the value function that defines the complexity of the problem and that the *simplest problem type here corresponds to a monotonous and additive value function*.

In contrast, a maximally complex problem is one where everything depends on everything else, and information obtained from trying alternative solutions to a particular sub-problem does not contain any signal as to whether the search as a whole has moved towards the globally optimal solution. Such a property of the problem manifests as ruggedness of the search landscape (Kauffman, 1993). A fully *non-modular* problem can be defined by a value function obtained by randomly shuffling the payoff values of all configurations apart from  $x_{min}$  and  $x_{max}$ . In such extreme cases, only flipset heuristics that address the whole  $N$ -length string at once can be guaranteed to reach the optimal solution. Consequently, if reaching the optimal solution is the aim, no division of labor is possible.

Obviously neither of these extremes is of much interest for exploring the logic of the division of cognitive labor and neither, as such, corresponds to any real cases of collaborative research in science. In the simplest possible case, the task can be divided into its atomic subtasks and solved accordingly. In the most complex case, the only viable research strategy is fully holistic guesswork, blind trial and error with respect to the whole problem, again without any possibility of systematic gains from coordinated division of labor. Arguably, it is thus *partly modular problems*, intermediate levels of problem complexity, which constitute the set of cases appropriate for examining the benefits and costs of division of labor.

---

<sup>4</sup> This assumption is unproblematic, because the agents in the model are blind to us associating state '1' with correctness.



**Table 1** Block-modular value function

String	Block A	"Value <sub>A</sub> "	Block B	"Value <sub>B</sub> "	Total value
0000	00	0	00	0	0
0101	01	0	01	1	1
1010	10	0	10	0	0
1111	11	2	11	2	4
... <sup>a</sup>					

<sup>a</sup>See Supporting Information for the full specification of the value function

### 3.2 Block-modular problems

Consider a simple model problem of length  $N = 4$ . Again '0000' and '1111' stand for the beginning state and the globally optimal solution. One way to implement a partly modular problem is the following. The full string is first divided to two blocks. In our example, *block A* consist of bits 1 and 2, and *block B* of the remaining bits 3 and 4. Then we define a value function which is random inside the blocks, but additive with respect to the contributions of the independent blocks: the "value"<sup>5</sup> of the first two bits and the "value" of the third and fourth bit. Such a problem can now effectively be decomposed into two subtasks, which can be solved independently of one another. See Table 1 for an example of a block-modular value function.

We can see that this problem is not fully modular, because block A cannot be further subdivided into independently solvable atomic subtasks (it cannot be solved simply by flipping one bit at a time). More generally, the formal property characterizing such *block-modular* value functions is that given the block structure, the state of bits not in the block  $k$ , i.e., the block configuration  $x_i(b_{-k})$ , does not influence the order ( $\succeq$ ) of the values of the states that different states of block  $x_i(b_k)$  map onto.

In the context of multi-, inter- and transdisciplinarity, a block-modular research problem is one which *can be solved by a multidisciplinary project*, in which the disciplinary sub-projects proceed independently of each other, and their final solutions are simply combined into the full solution.<sup>6</sup> In this sense, the only kind of cognitive coordination that the problem requires is the division of labor reflecting the true block structure (the 2|2-block structure in our example). Once the correct decomposition is established, no further coordination between disciplines is required.

Such friendly cases of partially modular problems can occur when independent approaches collectively add to the strength of the solution, but neither approach relies on information from the other to develop the best solution for their task. For example, projects combining economic and ecological models for sustainable resource use (such

<sup>5</sup> The scare quotes are there to remind us that the values of the sub-problems do not have any existence independent from the value function defined over the whole string  $x_i$ , and that they are only shorthand for highlighting how the total value function is defined.

<sup>6</sup> Multidisciplinary projects are typically distinguished from interdisciplinary ones whereby for the former problems are divided into distinct disciplinary tasks and the level of interaction or information exchanged between tasks is minimal (Mennes, 2020) Here we give a formal account of how a problem optimally suited for a multidisciplinary approach would be structured.



as forests or fisheries) can sometimes proceed by independent modeling of the resource (its growth rates) and the market for the resource. An optimal harvesting strategy or intensity can be found by sampling the ecological model over different harvesting intensities, and measuring the output until an economic optimum is reached. But finding the best economic and ecological models for the particular resource can be carried out largely independently once a set of economically relevant variables are agreed upon for the ecological model to track (e.g., measures of weight or quality of resource).

However, we should not expect neat modularity to be the default case when dealing with natural phenomena (cf. Kuorikoski & Pöyhönen, 2013). In engineering, modularity is typically an outcome of successful standardization (e.g., standardized electrical components or cargo containers). That is, modularity often requires work (Baldwin & Clark, 2000). Nature, in contrast, does not act like an engineer but instead like a backwoods mechanic making the best use of the components available. The resulting designs are often non-modular.

Moreover, task modularity alone is not enough. It also needs to be correctly perceived and taken into account in the division of cognitive labor. In terms of our model, the decomposition should match the partition. Yet this is often not how decompositions are determined; strong assumptions of modularity are commonly made and built into disciplinary organization and specializations. For example in Strasser et al. (2014) calculating the effect of climate change on ski tourism is broken up into the sub-problems of formulating a relevant climate change model based on climate science, a hydrological snow precipitation model, and two economic models, one of the effect of snow availability on demand and the other of the spatio-economic structure of the industry, to understand the overall effects on ski tourism of precipitation changes. Dividing climate science from snow precipitation tracks different specializations within environmental science, and both models are constructed largely independently (or simply taken off the shelf) and then put in communication with one another. The economics models are assumed to be capable of optimal construction independently of the natural science models. In general much interdisciplinary modeling in the environmental sciences assumes that environmental problems can be divided into disciplinary sub-problems, based on the perceived modularity governing the systems those disciplines track.

However, when tackling many complex real-world problems, a reasonable overlap between problem modularity and the decomposition imposed by disciplinary divisions cannot always be expected, and sticking to them might result in sub-optimal outcomes. For example, a resource management problem might be far more socially and politically entangled between different stakeholders, so decomposing it into economics and ecology components alone may not be sufficient to reach a good solution.

### 3.3 Nearly modular problems

Further complications arise when there are back and forth dependencies between sub-tasks. For example, the solution to sub-problem A can be found based on a solution of sub-problem B, but that solution in turn provides information which could potentially

**Table 2** Nearly modular value function

Line number	String	Block A	Block B	Total value
1	0000	00	00	0
2	1000	10	00	1
3	1100	11	00	0
4	1010	10	10	2
5	1011	10	11	0
6	1110	11	10	3
7	1111	11	11	4

change the best solution to sub-problem B, and so forth. In a collaboration between engineers and business scientists on an optimal design for a particular technology, the business scientists may be charged with investigating the usability and marketability of potential designs in order to assess which one works best. Engineers need the usability information to set constraints on their design space. But, to carry out their analysis, the business scientists need to know the design criteria engineers are working with. A usability analysis might provoke reflection on the basic design assumptions, and require engineers to shift or expand their design space to include different features. This would then require a new usability investigation based on different variables, and so on. As such, finding an optimal design may require an iterative form of problem-solving, in which the moves of each group depend on the criteria the other group is operating with. The problems each group faces can nonetheless be said to be modular to the degree that they can be independently solved given at least temporary criteria being set by their collaborators.

As a simple illustration, let us define the value function as follows (again with lines omitted for the sake of brevity):

Table 2 presents a *nearly modular problem* that is similar to the one in Table 1, except an additional interdependence has been introduced between the two subtasks: team A (Block A) cannot move forward from their local optimum (line 2) until team B (Block B) has taken its first step towards the solution (line 4), and furthermore, team B cannot reach the global optimum before team A has fully completed its search task. This value function thus represents a problem which requires an additional element of cognitive coordination: the task can only be solved by the correct sequencing of individual projects. This could be taken as an example of *the simplest possible case which requires true interdisciplinary interaction, in contrast to mere multidisciplinaryity*. We call different ways of organizing such cognitive coordination *coordination schemes*.

As the examples illustrate, we are still interested in the most minimal setups that allow collective search to reach the global optimal solution. This can be understood as follows: each discipline with its set of block-specific heuristics generates a neighborhood around each solution candidate  $l_i$ . The coordination scheme determines the order in which the heuristics from different disciplines are applied (e.g., a “round robin” in which one discipline searches for its local optimum first, then the second and so on; alternating moves between teams on each turn). Coordination schemes also determine how often the epistemic value of novel candidate solutions is evaluated.

Heuristics, together with a coordination scheme, result in a set of possible paths from group starting positions to the global optimum. Such paths define the reachability of the optimum in these different scenarios. Because epistemic efficiency grows as the size of the decomposition  $D$  decreases (Remark 1), we are interested in the properties of successful combinations of (i) coordination schemes and (ii) minimum-size decompositions for a given problem structure.

In general, for every problem, there exists a decomposition from which the optimum can always be reached, the degenerate decomposition  $D = \{1, 2, 3, \dots, N\}$ . If everything really depends on everything else, no finer decomposition, or any set of smart heuristics for that matter, can guarantee optimal search outcomes. As we pointed out, however, such problems lie beyond the reach of interdisciplinary modes of cognitive collaboration, requiring perhaps a form of transdisciplinarity in the sense of Nicolescu (2002), which is built around the radical reorganization or even dismantling of the disciplinary structure of scientific problem solving.<sup>7</sup>

#### 4 Efficiency of division of cognitive labor

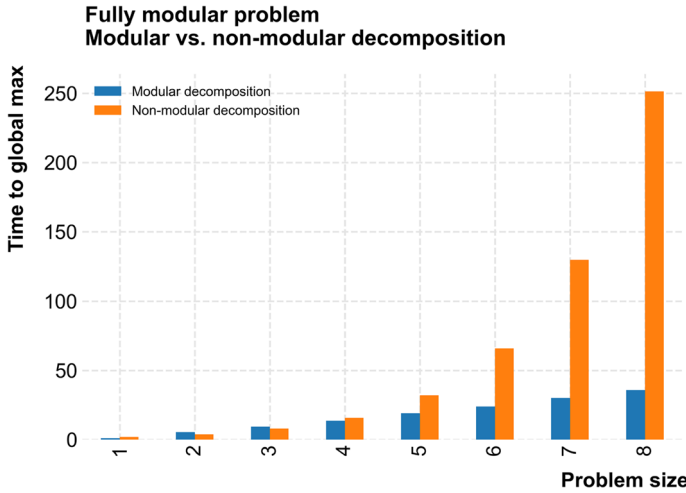
In this section, we continue our inquiry into the connections between levels of problem complexity and forms of cognitive coordination. In particular, whether a set of research heuristics can eventually reach the epistemic maximum is often not the only, or even the most relevant question. It is also important to determine how long it takes to reach the solution. To explore such questions of efficiency, and to shed light on comparative questions of epistemic performance, we now move from the relative informal model-based reasoning employed above to computational simulations.<sup>8</sup>

As the first step, we set some benchmarks relating to the performance of the extremely reductionist and holistic heuristics in solving problems of varying degrees of difficulty and complexity. Figures 1 and 2 illustrate the trade-off between the exhaustiveness of search and the time required. First, Fig. 1 compares the time performance of fully modular (reductionist) decomposition (decomposition size 1), and non-modular (holistic) search (decomposition size = problem size) in fully modular problems of different sizes  $N$ . Although exhaustive search is always guaranteed to reach the global optimum, even in moderately small- $N$  problems it uses significantly more time than the modular strategy.<sup>9</sup> To complement Fig. 1, Fig. 2 illustrates the necessity of engaging in exhaustive search, when the problem manifests no potential for a modular decomposition. Although in small problems  $N \in (1, 5)$  even the modular strategy

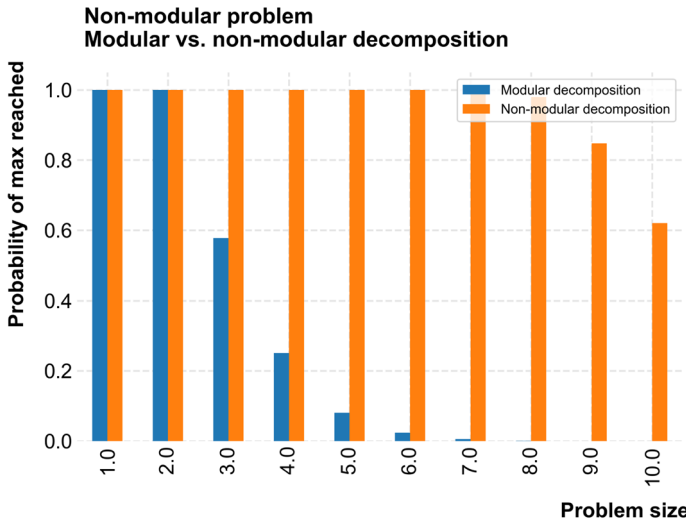
<sup>7</sup> Note however that Nicolescu's use of transdisciplinarity is not typical today. The concept usually refers to research which integrates the values and participation of extra-academic stakeholders. In both cases, however, scholars advocate against traditional disciplinary approaches and disciplinary structures (cf. Bernstein, 2015).

<sup>8</sup> The source code used to generate the findings in this section can be viewed at <https://osf.io/hw7zj/>.

<sup>9</sup> For the sake of simplicity, following Marengo and Dosi (2005) we assume that a discipline possesses the full set of flipset heuristics in its domain. In other words, in each round a discipline applies a random modification to the bits within its domain. Furthermore, we assume a mode of coordination where blocks are modified sequentially, each cover being applied to the position reached by the previous one. See SI for a robustness check of this assumption.



**Fig. 1** Fully modular ( $|D| = 1$ ), and non-modular decomposition ( $|D| = N$ ) applied to a fully modular problem. For all figures, 50 repetitions of the simulation were run over 50 different problem landscapes

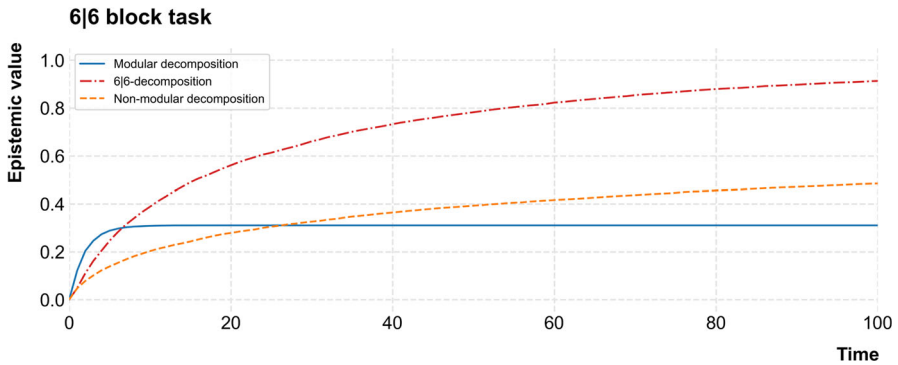


**Fig. 2** Probability of reaching global maximum for modular and non-modular decompositions in a non-modular task

sometimes manages to avoid local maxima, in problem spaces of higher dimensionality it generally fails to reach the optimal solution.<sup>10</sup>

But as already stated, these extreme cases are of little interest in investigating the efficiency of division of cognitive labor and forms of interdisciplinarity. Let us again start with the simplest case of a non-trivial nearly modular problem. Figure 3 presents

<sup>10</sup> The reason why the performance of the non-modular search tapers off on problem sizes  $N > 9$  is that the simulation is terminated after 1000 time steps. This reflects the temporal inefficiency of the non-modular search pictured in Fig. 1.



**Fig. 3** Temporal evolution of epistemic value reached in a block-modular task

the evolution of epistemic value over time in a  $N = 12$  task consisting of two blocks of size 6.<sup>11</sup> As the figure indicates, the fully modular (reductionist) strategy initially outperforms the other two strategies, but its performance quickly plateaus on a local maximum. The exhaustive search (full-size decomposition) does eventually reach the global optimum, but does so significantly more slowly than a search employing the 6|6 decomposition.

The initial success of maximally reductionist heuristics is noteworthy. In their evolutionary simulations, Marengo and Dosi (2005) showed that under similar conditions, highly modular heuristics may outperform even the heuristics employing the optimal decomposition (cover set structure matching the objective block structure). This can happen because highly decomposed search heuristics enable agents to quickly reach sufficiently high, but ultimately sub-optimal and local, gains, leaving the slower agents (including those who are, in a sense, right) in their wake and ultimately exiting the population.

Figure 4 illustrates a further result that complements observations made by Marengo and Dosi (2005). They showed that an optimal trade-off between search accuracy and speed is obtained by a decomposition matching the block structure of the task. In the figure, the horizontal axis represents different two-part decomposition, where the cut between the two covers is placed at a value plotted on the horizontal axis. Value 6, standing for the 6|6 decomposition where the first cover consists of bits 0-5 and the second of 6-11, and therefore maps onto the true block structure of the problem, leads predictably to optimal results.

Perhaps more interestingly, Fig. 4 also conveys the epistemic predicament of *uncertain modularity*. In real-life instances of complex problem solving, an important aspect of the task is to discover the appropriate problem decomposition. Although the 6|6 cover set outperforms the other decomposition strategies, the comparative signal is rather weak: given their tolerable epistemic performance, research projects employing, for example, a 5|7 cover set would hardly realize that their strategy was sub-optimal—at least initially. Furthermore, the extreme heuristics forgoing all attempts at carving the problem at suitable joints are not in fact the worst performing ones, as this honor

<sup>11</sup> In this 6|6 block task, bits 0–5 constitute the first block, and 6–11 the second one.

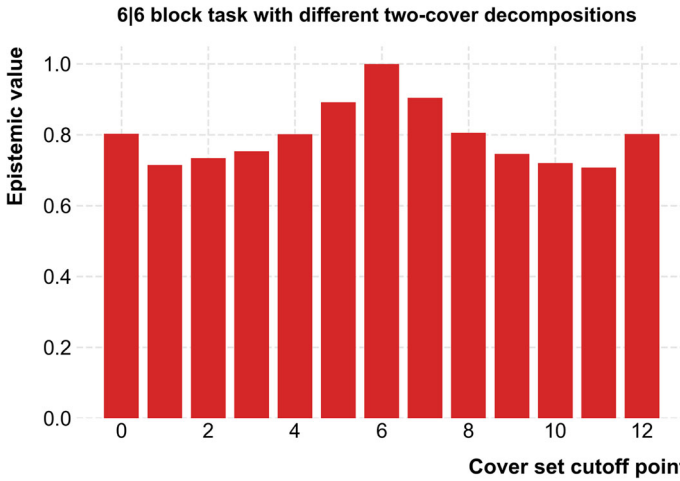


Fig. 4 6|6 block modular task. Expected epistemic value for different two-cover decompositions

Table 3 Summary of problem types, forms of cognitive coordination, and forms of cross-disciplinary organization

Fully modular	Block-modular	Nearly modular	Unknown modular	Non-modular
N independent covers, $ D  = 1$	Two independent covers, $k N - k$	Two covers and turn taking	Decomposition search (see Sect. 4)	Full cover, $ D  = N$
Multidisciplinary	Multidisciplinary	Interdisciplinary	Interdisciplinary	Transdisciplinary

is left for heuristics which do attempt to decompose, but get the decomposition badly wrong (i.e., the 1|11 cover set). This is yet another reminder of the importance of the right conceptualization of the problem.

Compared to, for example, many engineering tasks, assessing the adequacy of decomposition schemes is particularly difficult in many instances of scientific problem solving. Consider two different ways of task partitioning in aircraft design: In the first, one team designs the body and other team, the engine. In the other, team A designs the front half of the aircraft body *and* the engine, and the team B the back half of each (von Hippel, 1990). The latter decomposition strategy is manifestly absurd. However, as we see in the following section, analogous situations are far more difficult to detect in interdisciplinary scientific research.

### 5 Implications for interdisciplinary practice

We can now summarize our model-based reasoning relating different levels of problem complexity to different ways of organizing the collaborative division of cognitive labor. The first line of Table 3 summarizes the levels of problem complexity discussed

above. For simplicity, apart from the fully modular case, only two-block structures are discussed. The second line presents the minimal level of cognitive coordination that guarantees reaching optimal solutions in the problems of that column. The third line associates such forms of cognitive coordination with familiar labels of cross-disciplinary collaboration.

Table 3 does not aim to be exhaustive, but it illustrates some of the dependencies between problem structure and the required form of cognitive coordination made explicit within our template. As pointed out in Remark 1, search time increases exponentially as a function of decomposition size (moving from left to right in the table). This raises the question of whether there are forms of cognitive coordination that can be used to address complex problems without increasing decomposition size. This appears possible with two aspects of coordination discussed within the context of the current model.

A first option is to alter the evaluation of epistemic payoffs. Getting stuck on local maxima can be avoided by evaluating candidate solutions only after several heuristics from different disciplinary perspectives have been applied to the current solution candidate. As can be seen from the model, however, such a strategy does not produce time benefits. For example, if epistemic value is not evaluated after each round but only every second round after the application of covers of size  $n$  and  $m$ , the search space to be traversed is of the size  $2^m * 2^n$ . In terms of expected search time, this is equivalent to searching a solution space of dimensionality  $m + n$ .

That said, this strategy may still have useful practical consequences. Whereas disciplinary perspectives on a problem (i.e., covers) are determined by the disciplinary matrix (e.g., instrumentation, accepted background theories, training) and can be hard to change, the degree and timing of the points at which teams come together and collectively evaluate their progress may be more amenable to change. Consider different engineers collaborating on a technological design for a complex device to perform specific functions. Rather than simply producing their sub-systems and then integrating them at the end, our modeling suggests there can be a benefit to putting prototypical systems together at earlier stage and using the output of that testing to modify individual strategies. This might apply, for instance, when there is anticipated feedback between the systems so that simply producing the optimal sub-systems independently is not likely to produce an optimal overall result. Now what applies to technological design could also apply in more scientific contexts to model or experiment design.

These results, while theoretical, can be used to reflect on aspects of current interdisciplinary practice and science policy. Our simple modeling framework provides potential insight into how the cognitive structure of interdisciplinary problem-solving might influence the strategies researchers pursue and the results they obtain. It is clear from the above results that there is a general trade-off between highly modular searches and more exhaustive holistic searches: highly modular searches are faster but tend to get stuck on local optima whereas fully exhaustive searches are bound to hit the global optimum but may take astronomically long to get there. As such the competitive advantage of highly reductionist over more holistic search is the larger the shorter the time span and the stronger the selection pressure acting on agents (Marengo & Dosi, 2005). These results suggest that part of what might make interdisciplinary results lack depth, when they do, can be attributed to researchers' preference (in the context of



typical problem-driven, fixed-term funding instruments) to treat complex problems as modular and apply familiar disciplinary techniques (heuristics). The resulting search behavior is time effective and may yield promising initial results, but is likely to have sub-optimal long-term outcomes.

On these grounds, the current funding model of bringing together researchers from different disciplines to work on specific problems under considerable time pressure may set up the conditions for interdisciplinary failure. Given the extreme competitiveness of academic systems, short-term project funding heavily incentivizes the search for quicker solutions, thus potentially crowding out of the population of interdisciplinary scholars willing to work slowly towards a better global solution.

Neither our model nor the one presented in Marengo and Dosi (2005) incorporate an explicit decision model: the models do not assume that actors make future-oriented decisions about what kind of heuristics best guarantee their long-term survival in the epistemic game. Real-life researchers invested in interdisciplinary projects, however, undoubtedly do so. If the researchers know that higher epistemic payoffs are more risky and lie in the future—quite likely beyond the project funding period—the sensible thing to do is to play it safe. The relative short-term success of using tried and tested disciplinary heuristics thus also resonates with the widely shared belief (which is notoriously difficult to test empirically), that in contrast to the official gospel of interdisciplinarity, short-term interdisciplinary projects actually favor methodological conservatism over slower and more uncertain methodological innovation and integration of knowledge. In sum, typical funding structures may well put off those more willing to pursue deeper solutions in favor of individuals who aim to collect quick results and build up their own credentials.

The drive to modularize problems at the outset is not purely the result of scientific expedience, but is built into the typical work-package structure characteristic of large-scale projects. Such structures require that researchers divide work into modular problems with individually evaluated milestones. This kind of structure for cognitive coordination restricts the possibility of holistic search at the outset, institutionalizing reductionist local search as the governing paradigm.

The obvious suggestion for achieving deeper results is to be more flexible regarding the time constraints. Six-year projects may not be long enough. Researchers may not be willing to invest enough time and resources into developing new search strategies when the end of the project is in sight, and the expertise they anticipate to need in the post-project future follows established disciplinary traditions. A better situation may be one where a researcher is recruited into a permanent position at a research institution on a specific topic, or generally, when individuals rather than projects are funded (Ioannidis, 2011): If a researcher knows that the integrative work they do can be considered as a rational investment in their future work, introducing changes in search strategies, and ultimately disciplinary decompositions, becomes more feasible.

Our model also tentatively suggests a strategy for improving the chances of even fixed-term interdisciplinary projects providing good, if not optimal, solutions to complex problems. As mentioned above, evaluating the epistemic value of a solution candidate only after the multiple sub-projects have applied their individual heuristics (covers) is formally equivalent to increasing the decomposition size. Thus an interdisciplinary consortium can effectively increase its neighborhood of search by refraining

from proceeding at the outset simply by deferring to individual research, and instead, organizing collaborative activities so that deliverables are produced and evaluated only after collaborative research activities, where insights from several perspectives are combined. Needless to say, this comes at the price of accepting a significantly slower pace of progress—at least initially. The forced introduction of predetermined milestones into multidisciplinary research projects institutionalizes quick epistemic evaluation over more holistic search. In this regard, one thing that funding committees interested in deepening interdisciplinary outcomes in a short-term project might require from applicants is evidence of more nuanced strategies for building joint search spaces collaboratively explored by the different teams in the project.

The final question to consider is whether the epistemic maximum is always even the right goal. Quicker, less optimal solutions may actually be desirable in contexts of urgency. This applies to many areas where interdisciplinary research is often desired, for example sustainability science. A satisficing-based approach, relying on reductionist heuristics without context-sensitivity, may often be what the practical context calls for (cf. Wimsatt, 2007). The same point holds for the most complex class of problems, for which no successful decomposition is possible (what Wimsatt calls causal thickets) or for wicked problems in which the preferences and values of stakeholder are in conflict, context-dependent, and may shift in response to different solutions (Rittel & Webber, 1973). In our template, this would mean that the value function would be undefined and the problem would therefore lack a clear structure for researchers to formulate. The ambition and promise of transdisciplinarity has been to address real-world problems of such a complexity, for which science should strive towards a wholesale rejection and transcendence of traditional disciplinary boundaries (Bernstein, 2015). Yet after all these years, this promise seems as elusive as ever. Even disregarding the institutional and cognitive challenges that an exceedingly holistic approach to science entails, the glacial speed of exhaustive search processes inevitably raises the question whether the rational thing to do, instead, is to utilize existing disciplinary decompositions and make do with the sub-optimal but reachable epistemic goals.

## 6 Conclusions

In this article, we have conceptualized multi- and interdisciplinary research as modes of collective problem solving that rely on different forms of cognitive coordination. We presented a model template for formally reasoning about the relationship between levels of problem complexity and the appropriate form of coordination, or what we called a *coordination scheme*. Coordination schemes, in turn, can be characterized in terms of the heuristics, division of cognitive labor, and scheduling of joint epistemic work and evaluation. By examining the relationships between different coordination schemes and types of problems, we evaluated the adequacy of different organizational arrangements for different problem-solving tasks. Heuristics and decompositions adaptive in a mono- and multidisciplinary contexts no longer lead to optimal solutions in the more complex solution spaces which cross-cut disciplinary perspectives.

Our second aim was to draw tentative conclusions concerning the epistemic effects of current institutional forms of interdisciplinary research. We suggested that under

prevailing institutional conditions, interdisciplinary projects may resort to sub-optimal, overly reductionist strategies. Our modeling framework helps clarify the problem-solving structures and the cognitive incentives with which researchers operate under these institutional conditions, and thus helps us understand why they would resort to seemingly sub-optimal strategies. In contrast to the accepted ideology of interdisciplinarity, which emphasizes innovation, current institutional structures governing interdisciplinary research tend to encourage conservative science. This is because the relatively modular individual heuristics most likely selected (in short-to-medium time frames) usually represent well established ideas and methodologies in the parent disciplines, and the sub-optimal solutions reached by such methods represent, at most, mediocre science.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11229-023-04193-4>.

**Acknowledgements** This article was presented at the *Centre for Philosophy of Social Science* (University of Helsinki) research seminar as well as the DFG research network workshop *Simulations of Scientific Inquiry* (University of Bochum), *Ghent-Brussels Seminars in Logic, History and Philosophy of Science* (Ghent University), and *CamPOS* seminar (University of Cambridge). We'd like to thank the participants at those sessions for their valuable feedback. Further we would like to thank the contributions of the three anonymous reviewers for their questions and comments, Kate Sotejeff-Wilson for editing the manuscript, and Maximilian Noichl for help with the illustrations. This research received funding from the Academy of Finland.

**Funding** Open Access funding provided by University of Helsinki including Helsinki University Central Hospital.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adams, J., Jackson, L. & Marshall, S. (2007). Bibliometric analysis of interdisciplinary research. *Report to the Higher Education Funding Council for England*.
- Baldwin, C.Y. & Clark, K.B. (2000). *Design rules: The power of modularity* (Vol. 1). MIT press.
- Bechtel, W., & Richardson, R. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. MIT Press.
- Bernstein, J.H. (2015). Transdisciplinarity: A review of its origins, development, and current issues. *Journal of Research Practice* 11(1).
- Bruce, A., Lyall, C., Tait, J., & Williams, R. (2004). Interdisciplinary integration in Europe: The case of the Fifth Framework programme. *Futures*, 36(4), 457–470.
- D'Agostino, F. (2009). From the organization to the division of cognitive labor. *Politics, Philosophy & Economics*, 8(1), 101–129.
- Darden, L. (1997). Recent work in computational scientific discovery. In *Proceedings of the nineteenth annual conference of the cognitive science society* (pp. 161–166). Lawrence Erlbaum.
- Hong, L., & Page, S. E. (2001). Problem solving by heterogeneous agents. *Journal of Economic Theory*, 97(1), 123–163.

- Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), 16385–16389.
- Ioannidis, J. P. (2011). Fund people not projects. *Nature*, 477, 529–531.
- Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Press.
- Klein, J. T. (1990). *Interdisciplinarity: History, theory, and practice*. Wayne state University Press.
- Kuorikoski, J. & Pöyhönen, S. (2013). Understanding non-modular functionality. Lessons from genetic algorithms. *Philosophy of Science*, 80(5), 637–649.
- Kuorikoski, J. & Reijula, S. (2019). Making it count. An inferentialist account of computer simulation. SocArXiv. <https://osf.io/preprints/socarxiv/v9bmr/>.
- Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, 61(1), 126–131.
- Leahey, E., Beckman, C. M., & Stanko, T. L. (2017). Prominent but less productive: The impact of interdisciplinarity on scientists' research. *Administrative Science Quarterly*, 62(1), 105–139.
- Lindvig, K., & Hillersdal, L. (2019). Strategically unclear? Organising interdisciplinarity in an excellence programme of interdisciplinary research in Denmark. *Minerva*, 57(1), 23–46.
- MacLeod, M., & Nagatsu, M. (2016). Model coupling in resource economics: Conditions for effective interdisciplinary collaboration. *Philosophy of Science*, 83(3), 412–433.
- Marengo, L. (2015). Representation, search, and the evolution of routines in problem solving. *Industrial and Corporate Change*, 24(5), 951–980.
- Marengo, L., & Dosi, G. (2005). Division of labor, organizational coordination and market mechanisms in collective problem-solving. *Journal of Economic Behavior & Organization*, 58(2), 303–326.
- Marengo, L., Dosi, G., Legrenzi, P., & Pasquali, C. (2000). The structure of problem-solving knowledge and the structure of organizations. *Industrial and Corporate Change*, 9(4), 757–788.
- Mennes, J. (2018). SenseDisclosure: A new procedure for dealing with problematically ambiguous terms in cross-disciplinary communication. *Language Sciences*, 69, 57–67.
- Mennes, J. (2020). Putting multidisciplinary (back) on the map. *European Journal for Philosophy of Science*, 10(2), 1–23.
- Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104). Prentice-Hall.
- Nickles, T. (1980). Scientific discovery, logic, and rationality. In *Boston studies in the philosophy of science* (Vol. 56). Reidel.
- Nickles, T. (1981). What is a problem that we may solve it? *Synthese*, 85–118 .
- Nicolescu, B. (2002). *Manifesto of transdisciplinarity*. Suny Press.
- Page, S. E. (1994). Covers: A theory of Boolean function decomposition. *Complex Systems*, 8(1), 1–24.
- Page, S. E. (1996). Two measures of difficulty. *Economic Theory*, 8(2), 321–346.
- Reijula, S. & Kuorikoski, J. (2019). Modeling epistemic communities. In: Fricker, M., Graham, P., Henderson, D., & Pedersen, N. (Eds.), *The routledge handbook of social epistemology* (pp. 240–249). Routledge.
- Reijula, S., & Kuorikoski, J. (2021). The diversity-ability trade-off in scientific problem solving. *Philosophy of Science*, 88(5), 894–905.
- Reijula, S. & Kuorikoski, J. (2022). Modeling cognitive diversity in group problem solving. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44(44) .
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), 155–169. <https://doi.org/10.1007/BF011405730>
- Salmela, M., MacLeod, M., & af Rosenschöld, J.M. (2021). Internally Incentivized Interdisciplinarity: Organizational restructuring of research and emerging tensions. *Minerva*, 1–23 .
- Simon, H. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482.
- Simon, H. A. (1989). The scientist as problem solver. *Complex information processing: The impact of Herbert A. Simon*, 375–398 .
- Sobieszczanski-Sobieski, J., Morris, A., & Van Tooren, M. (2015). *Multidisciplinary design optimization supported by knowledge based engineering*. Wiley.
- Strasser, U., Vilsmairer, U., Prettenhaler, F., Marke, T., Steiger, R., Damm, A., Hanzer, F., Wilcke, R. A., & Stötter, J. (2014). Coupled component modelling for inter-and transdisciplinary climate change impact research: Dimensions of integration and examples of interface design. *Environmental modelling & software*, 60, 180–187.

- von Hippel, E. (1990). Task partitioning: An innovation process variable. *Research Policy*, 19(5), 407–418.
- Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Harvard University Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.