

**PHD THESIS SUMMARY:
A Model-Based and Mechanistic Approach to Social
Coordination**

MATTI SARKIA

PhD in Philosophy, September 2022

University of Helsinki

Human beings live in group sizes that outscale any other animal society, with the exception of genetically closely related superorganisms, such as ants and beehives. The average proximate group size (involving regular contact and name recognition) of most humans is on the scale of 150 individuals, thousands of individuals cross each other on a daily basis when commuting, and millions of individuals may occupy the same urban territory. Such large-scale social life demands coordination: in order to avoid hostile confrontations and accidental collisions, individuals need to be able to coordinate their attitudes and behaviors in a manner that is mutually beneficial, or minimally, that avoids unnecessary harm. How do we achieve this remarkable feat?

The human capacity for social coordination, as well as the cognitive and social mechanisms underlying its manifestations in various environments, have become prominent topics of research in many different scientific disciplines during recent years. These disciplines include analytical philosophy, cognitive science, developmental psychology, experimental economics, game theory, and evolutionary anthropology. They engage with human social coordination on multiple spatial and temporal scales, from the phylogenetic time scale of gene-culture coevolutionary processes (consider the evolution of symbolic signaling systems) through the ontogenetic time scale of individual development (for example, learning natural language during infancy) to the synchronic time scale of two physically co-present individuals coordinating their motor behaviors with one another (such as the tendency to adjust one's tone of voice to that of one's conversation partner). However, genuinely interdisciplinary investigations of social coordination have remained rare, and there have been few attempts to integrate different disciplinary perspectives into a more comprehensive, multidisciplinary theory of social coordination. Of

course, institutional incentives may play some role in hindering interdisciplinary collaboration. However, my dissertation is premised on the idea that one important obstacle to more extensive interdisciplinary integration is the absence of shared methodological frameworks for formulating multilevel explanations of complex phenomena that occur at multiple temporal and spatial scales, involve a mix of sub-personal and personal-level cognitive processes, and are influenced in myriad ways by the cultural and social environments that they occur in.

Two such methodological frameworks from contemporary philosophy of science play an especially central role in my dissertation. First, the mechanistic approach to explanation, which was originally developed in the biological and life sciences, enables me to frame questions about how investigations of social coordination in the behavioral, cognitive, and social sciences relate to one another (Sarkia et al. 2020; Sarkia 2016; 2022b). The central idea behind mechanistic explanation is to explain *why* some phenomenon occurs by describing *how* it occurs, or in the words of one classic paper, by describing "a structure performing a function by virtue of its component parts and component operations and their organization" (Bechtel and Abrahamsen 2005, 423). While mechanistic explanations might be viewed as *ontologically reductive*, I argue in my dissertation that they are nevertheless *epistemically anti-reductive*, because the parts and operations that make a difference are individuated relative to the phenomenon to be explained. This is so especially in the case of complex behavioral phenomena, such as human aggression, altruism, coordination, or cooperation. In these cases, the *phenomenal decomposition* of the *explanandum* into different manifestations (consider the distinction between pure coordination games and mixed-motive games in game theory) serves as an important prerequisite for its *mechanistic decomposition* into entities and activities that may produce these different manifestations (such as psychological mechanisms for altruism or empathetic preferences as explanations of cooperation in public goods and dictator games). Although not all phenomena are mechanistically decomposable, many are, and much cutting-edge research at the intersection of the behavioral, cognitive, and social sciences is premised on the assumption that complex human behaviors do not resist mechanistic explanation in principle, even if mechanistic explanation is a more interdisciplinary and multifaceted endeavor than is often imagined.

The second methodological framework that plays a central role in my dissertation is theoretical modeling, understood as a distinctive approach

to scientific investigation, which is based on the construction and elaboration of hypothetical surrogate systems that are used to stand in for some system in the real world. The central idea behind model-based science is that instead of studying the world directly, by means of observation and experimentation, it is sometimes a more feasible strategy to study the world indirectly, by first constructing a theoretical model, and then using that model to draw inferences about its target. Theoretical models abound in sciences that deal with complex phenomena, which resist experimental isolation and control: consider climate models in environmental science, models of rational action in economics and artificial intelligence, models of predator-prey interactions in ecology and evolutionary biology, and model organisms in medicine. However, even when they engage with the same subject matter, there can be significant differences between the types of models that are developed in different academic disciplines, which in turn makes it difficult to draw comparisons and adjudicate among them. For example, game theorists typically model social coordination as equilibrium behavior that is based on best response-reasoning, where no individual has an incentive to change her behavior provided that the others do not change their behaviors. On the other hand, philosophers have often discussed social coordination under the rubric of collective intentionality (referring to the sharing of goals or other mental states with other individuals), while cognitive scientists have also discussed non-intentional forms of coordination (such as two individuals unconsciously aligning their bodily postures, mannerisms, and facial expressions). Given their contrasting disciplinary conventions and perspectives, the integration of theoretical models from different scientific fields gives rise to numerous conceptual and methodological challenges. My article-based dissertation aims to answer such challenges in a manner that is piecemeal and problem-oriented, in accordance with a naturalistic approach to the philosophy of social science.

The first three articles of my dissertation deal with questions of general philosophical interest, philosophical methodology, and social ontology. In “Modeling Intentional Agency: A Neo-Gricean Framework” (Sarkia 2021), I distinguish three alternative strategies for modeling intentional agency in analytic philosophy of mind and compare them to similar strategies of scientific modeling. The bottom-up strategy of Gricean modeling is compared to the use of computer simulations and agent-based models, while the horizontal strategy of analogical modeling is elucidated by reference to research on analogical inference in cognitive science, and the

top-down strategy of theoretical modeling is compared to the uses of abstract and domain-general mathematical frameworks in fields like economics and ecology. In “A Family of Models of Shared Intentionality” (*under review*), I describe different philosophical analyses of shared intentionality (by Bratman, Gilbert, Searle, and others) as alternative elaborations of a common folk psychological framework, which share a set of core assumptions about how different types of mental states (such as goals and intentions) are related to one another, but which make contrasting elaborations and extensions to this framework thus leading to contrasting accounts of shared intentionality. And in “A Model-Based Approach to Social Ontology” (Sarkia 2022a), I provide an in-depth analysis of Raimo Tuomela’s (2013) philosophical account of the we-perspective as a highly idealized theoretical model, which deliberately abstracts away from certain aspects of social reality in order to bring others into sharper relief.

The latter three articles of my dissertation engage with questions of interdisciplinary integration and division of labor. In “Minimalism and Maximalism in the Study of Shared Intentional Action” (Sarkia 2016), I argue that cognitively undemanding accounts of social coordination have been defended by means of two contrasting strategies of argumentation, which I describe as ‘complementarist’ and ‘constitutionalist’. In “Mechanistic Explanation, Interdisciplinary Integration, and Interpersonal Social Coordination” (*under review*), I apply the distinction between phenomenal decomposition and mechanistic decomposition to carve out a schematic division of labor between behavioral and cognitive scientists studying social coordination. And in “Mechanistic Explanations in the Cognitive Social Sciences: Lessons from Three Case Studies” (Sarkia et al. 2020), I (together with my co-authors) present the mechanistic approach to explanation as a general template for integrating models and theories across the behavioral and social sciences, using cognitively inspired research on ethnicity, social coordination, and transactive memory systems to make our case.

Taken together, the six original articles in my dissertation demonstrate how contemporary approaches related to mechanistic explanation and theoretical modeling in the philosophy of science can play an important role in adjudicating interdisciplinary debates related to social coordination, and indicate how philosophical studies of social coordination interlock with empirical and theoretical studies of social coordination in other scientific disciplines.

REFERENCES

- Bechtel, William, and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36 (2): 421–441.
- Sarkia (né Heinonen), Matti. 2016. "Minimalism and Maximalism in the Study of Shared Intentional Action." *Philosophy of the Social Sciences* 46 (2): 168–188.
- Sarkia, Matti. 2021. "Modeling Intentional Agency: A Neo-Gricean Framework." *Synthese* 199: 7003–7030.
- Sarkia, Matti. 2022a. "A Model-based Approach to Social Ontology." *Philosophy of the Social Sciences* 52 (3): 175–203.
- Sarkia, Matti. 2022b. *A Model-Based and Mechanistic Approach to Social Coordination*. Philosophical Studies from the University of Helsinki.
- Sarkia, Matti, Tuukka Kaidesoja, and Mikko Hyryläinen. 2020. "Mechanistic Explanations in the Cognitive Social Sciences: Lessons From Three Case Studies." *Social Science Information* 59 (4): 580–603.
- Tuomela, Raimo. 2013. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press.

Matti Sarkia obtained his PhD from the University of Helsinki in September 2022. He is currently working as a post-doctoral researcher in the Kone Foundation-funded research project: Cognition in Social Interaction: Towards a Mechanistic Integration of the Cognitive and Social Sciences. Contact e-mail: <matti.sarkia@helsinki.fi>