Objectives

The objective of the project is to go beyond the generic idea of the unity of science and ask questions of practical value for today's science: What unifies a field of research? Why various studies are considered contributions to the *same* topic? What makes cognitive sciences integrated? Is there anything that makes it a unified perspective for research on cognition and related phenomena? Although there is little doubt that cognitive science has always been conceived as an interdisciplinary unity (Arbib et al., 1978), integration and unity of cognitive science (or any interdisciplinary field) is still poorly understood. There is obvious danger in disintegrating the cognitive sciences, which may lead to isolation and duplication of effort in research on similar topics; additionally, there are parallel dangers of illusory integration by simply ignoring the input of some disciplines and monopolization of the field by one discipline.

These latter dangers are usually referred to as *greedy reductionism* (Dennett, 1995). The recent controversy over the Human Brain Project (where the funding for the cognitive architectures and behavioral studies is expected to be cut [Frégnac & Laurent, 2014]) shows that the neuroscientific community disapproves greedy reductionism. Similarly, theorists of cognitive science have criticized premature reduction of cognitive research to biological research (Fodor, 1974) and advocated for the autonomy of cognitive science and psychology from brain sciences. However, contemporary psychology does not proceed in splendid isolation from biological sciences. For example, psychological theories of attention, such as (Posner & Reichle, 1994), have been rejected because of neuroscientific evidence (Francuz, 2012, p. 44; Lauwereyns, 2011). No wonder that greedy antireductionism defended by Jerry Fodor in the 1970s usually falls on deaf ears today. In brief, psychologists realize that biological evidence can help constrain their own theorizing. However, there is no satisfactory account of such constraints.

How best to avoid the Scylla of greedy reductionism and Charybdis of extreme autonomy? To answer this question, we need to develop a theory of unity and integration in science. The received view on theories (wherein the unity of science was framed in terms of inter-theoretical relationships, usually reduction) has fallen out of favor (Dupré, 1993; Fodor, 1974; Paprzycka, 2005). A reason for this is that reduction was conceived by logical empiricism in terms of deductive relationships, and the progress in reducing various special theories to a fundamental physical theory has been extremely slow, to say the least. Even if there are still defenders of the received view, their advice rarely goes beyond a programmatic manifesto. Such is the case with the idea of *consilience* defended by E. O. Wilson (1998).

In this project, a mechanistic account of integration and unification in cognitive sciences, with a possible application in other special sciences, will be developed. The account will be based on the study of extant cases of successful collaboration and theoretical cross-pollination in cognitive sciences. In other words, the objective is *not* to find a single unified theory of cognition that would replace or eliminate all other contenders. Just the opposite. It is how this proposal is different from the received view in the philosophy of science.

This consideration means that integrating the discipline is not to be confused with replacing all other theories with a single one, as long as sound, alternative theories can be found. For our purposes, it means that the account of mechanistic unification and integration will not aim at eliminating multiple models or theories. So how should such integration proceed?

The hypothesis defended in the project is that integration and unification can be understood in terms of constraints between models of mechanisms, and that building multi-level models of mechanisms underlying cognition is the most promising way of integrating the field and the best candidate tool for theoretical unification.

The mechanistic account of explanation is particularly sensitive to issues of inter-field research (Craver & Darden, 2013; Darden & Maull, 1977). According to new mechanism (which is one of the most successful approaches to explanation in contemporary philosophy of science), to explain a phenomenon φ is to elucidate the causal structure of the mechanism that gives rise to φ. While mechanisms are defined variously, the core idea is that they are organized systems, comprising causally relevant component parts and operations (or activities) thereof (for a recent review, see, e.g., [Illari & Williamson, 2011]). Component parts of the mechanism interact, and their organized operation contributes to the capacity of the mechanism to exhibit φ.

Another important notion to be elucidated is the one of the *inter-field theory*. The inter-field theories are ones that relate at least two fields of study. By a "field of study", Darden and Maull understand, for example, cytology or genetics rather than biology; in other words, a field has a more restricted scope than a theory or a discipline. Two fields may appeal to the same spatiotemporal locations, entities, or activities, and one of them may provide a better understanding of the spatiotemporal relationships, causal relationships, physical nature, structure, or function thereof. In the case of cognition, it is quite clear that cognitive processes relationships may be explained in various ways by various disciplines.

There are at least three ways fields may become integrated mechanistically: by *simple integration*, when the models of mechanisms can be considered pieces of puzzle that fit together; by *inter-level relationship*, when another level of organization is added to make explanation more complete; and by *inter-temporal integration* (Craver & Darden, 2013, Chapter 10). In the case of simple integration, two fields may simply study cognition in a similar way but with a slightly different stress.

The inter-level relationship is much more complex, as it may be confused with reduction. The need to introduce multiple levels of explanation is related to the nature of the mechanism under study. As Herbert Simon (Simon, 1996) has argued on theoretical grounds, complex systems are likely to be *near-decomposable* or composed of subsystems whose interactions are weak but not negligible. One facet of near-decomposability is hierarchical organization, in which different levels can be discerned with interactions at a different order of magnitude. Such systems (as long as their capacities to be explained are identified) can be subject to mechanistic constitutive explanation, in which lower levels of organization explain higher levels. Organization levels are understood spatiotemporally, and the relationship between them is proper part-whole relationship (Craver, 2007) (for a longer account, see [Wimsatt, 2007]; for a more deflationary one, [Eronen, 2014]). Note that the existence of a lower level explanation does not make the higher level disposable in this framework. The higher level is explained by the interaction of the components and activities on the lower level, and the lower level contributes to the capacity of the higher level. In such a case, one could say that our knowledge of the higher level has been deepened, which also leads more empirical credentials to our previous beliefs about the higher level (Thagard, 2007).

As long as explanations are integrated in an inter-level fashion, they are not only truth-constrained. The model of the lower level of a mechanism elucidates the activities and entities of the mechanism on the higher level. Such mechanistic explanations, called *constitutive*, cover at least three levels of organization: the *bottom* (-1) level, which is the lowest level in the given analysis and describes the internals of mechanism parts and their interactions; an *isolated* (0) level, at which the parts of the mechanism are specified along with their interactions (activities or operations); and the *contextual* (+1) level, at which the function of the mechanism is seen in a broader context. Depending on the shared scientific practice, the bottom level in the explanation will vary (Machamer, Darden, & Craver, 2000), as well as the upper levels. Note that one can easily introduce a further level if needed. Let's take an explanation that accounts for reproduction of bacteria. The reproduction of bacteria in a given environment (contextual level) is explained in terms of division (isolated level that ignores the environment) and division in terms of cellular mechanisms (the bottom level). The cellular-level mechanism can be further explained by its molecular parts, which would introduce a fourth level in this explanation.

The main, though fallible, heuristics in such explanations are localization and decomposition (Bechtel & Richardson, 1993). For example, social capacities of a human being can be explained by a psychological model, the psychological capacities with a neuroscientific model, and many neuroscientific explanations proceed from neuroimaging studies (though some of them are not reliable; see [Carp, 2012; Trout, 2008]). Many models that try to explain economical behavior only in neurophysiological terms fail to cite relevant factors known from psychological studies; for example, neuroeconomics has lost touch with the rest of neuroscience by failing to integrate behavioral studies combined with physiological, pharmacological, or anatomical techniques that rely on animal models (Lauwereyns, 2011, p. 89).

The inter-level nature of constitutive mechanistic explanations makes new mechanism a natural ally of explanatory pluralism (Dale, Dietrich, & Chemero, 2009; Gervais, 2014; Miłkowski, 2013). Instead of suggesting that there should be just a bottom-level causal explanation of all levels of organization of a given complex system, new mechanism insists that explanations at all levels are needed. These different explanations need not belong to the same discipline. They may be shared among various fields and disciplines. In other words, constitutive mechanistic explanations don't require different disciplines to become completely lumped together, as long as they can provide input for a common body of knowledge about a given mechanism.

The inter-temporal integration applies to phenomena that can be analyzed (due to their hierarchical organization) on multiple temporal scales. Different explanations of behavior will require researchers to appeal to mechanisms operating at different time scales. For example, timing considerations at different scales are critical in consciousness studies (Varela, 1999) and in multiple other cognitive processes such as language use (Rączaszek-Leonardi & Kelso, 2008) or imagination (Dominiak & Nosal, 2008).

By framing explanations of cognition in mechanistic terms, one can not only understand current scientific practice but also articulate certain norms of explanation, which is useful for integrating cognitive sciences. In brief, they will be unified as long as it will study the complex mechanisms underlying cognition and use knowledge about mechanisms operating at different levels of organization. This includes the environment broadly construed (Lewicka, 2012) and various time scales to constrain hypotheses about the overall structure of the mechanism (for another similar proposal, see [Castelfranchi, 2014]).

The important task for the project at hand is to study different modes of mechanistic integration in detail by performing case studies and develop a richer taxonomy of possible relationships between mechanistic models. These studies will be used to justify normative considerations for cognitive modeling.

Instead of suggesting that there is a single, privileged theory of cognition to which all other theories or models should be reduced, new mechanism stresses that understanding complex phenomena requires rich, multilevel models operating at multiple time scales. As such, it can help establish common research topics and identify the core sub-mechanisms of cognition. One important open issue is how to develop mechanistic models in the case of multi-model idealization; in other words, the objective is an alternative account to (Irvine, 2014).

The deliverable of the project will be at least **20** papers in English on the topic, authored by the team members. There will also be a call for proposals under a special research topic in an international open access journal (such as Frontiers in Theoretical and Philosophical Psychology) and one special issue of an international journal devoted to the research topic. In an effort to make science more open and for better outreach of the project, a project blog in English will host discussions on the topic of the project.

Significance

There are several existing accounts of integration and unification via: (1) shared concepts; (2) common methodology; (3) reduction to the basic level theory; (4) (specifically in cognitive science) appeal to cognitive or robotic architectures; (5) vertical integration; (6) multiple constraints. We will elaborate on these in turn.

The account of unification and integration as based on *shared concepts* could, for example, claim that the notion of cognition is what unifies cognitive sciences.[1] The notion of cognition is, however, notoriously controversial. In the recent discussion over the mark of the cognitive (Fred Adams & Garrison, 2012; Frederick Adams & Aizawa, 2008; Frederick Adams, 2010; Rowlands, 2009), most proponents discuss the claim that the mark of the cognitive is the notion of genuine, or underived, mental representation. However, the notion of representation seems to be even more controversial, with numerous theorists rejecting its value for cognitive research (Chemero, 2000, 2009; Garzon, 2008; Hutto & Myin, 2013). Even if the notion of representation is valuable for cognitive science, it cannot play a pivotal role for inter-field integration and theoretical unification.

There is also a deeper meta-theoretical reason to think that basic notions of a field research such as *cognition* cannot fulfill the integrative role. They need not refer to the same entity. As John Dupré (1993) has argued, inter-theoretic identifications of entities are not straightforward. Different biological disciplines idealize their entities in various ways. A lynx for ecological biology is an extremely idealized entity, especially when it comes to mathematical predator-prey models such as the Lotka and Volterra equation (for an analysis of the model, see [Weisberg, 2013]). In this equation there is no role for genes, so as far as this model is concerned, lynx might have no genes. It just has to play the predator role. But for molecular biology, predation might have no importance at all, whereas genetic structure is crucial. Hence, different fields of study may focus simply on a different entity when they use the notion of cognition.

However, Dupré's conclusion that the lynx under study is a different entity in different fields does not follow from his arguments. Even if particular models consider only some of properties of an entity, the very applicability of the Lotka-Volterra model to a lynx presupposes a larger body of knowledge about the animal. For example, we have to know that it is predatory because it is a species of wild cats. This knowledge is necessary for the application of the predator-prey models but not contained therein. In other words, Dupré's argument fails, even if he is right that we may idealize lynx differently for different explanatory answers. Still, those different idealizations may be so disparate that they do not constitute a consistent field of research. Physics and theology can study the same entity, such as an ancient scripture, but they do not constitute the same field of research. There is little reason to think that the mention of the same entity (even if it is the same natural kind) makes different disciplines unified. Various disciplines may frame cognition in different terms, idealizing it for its own uses, just like life sciences frame the notion of life quite differently (Miłkowski, 2013).

In other words, even if cognition is a natural kind, it need not unify theoretically different sciences of cognition. However, there is a recent proposal that cognition is best understood in terms of a cluster of properties (Buckner, 2013). The claim is based on the analysis of criteria used in comparative psychology to test whether a behavior was generated by a cognitive or a non-cognitive process. Buckner, however, does not claim that cognitive science can be unified just because it uses the notion of cognition. Just the opposite. The claim is based on

[1] This is just one example; some suggest that the unifying notion should be *affordance* or *scaffolding* (Estany & Martínez, 2013). They are, however, equally controversial (Frederick Adams & Aizawa, 2008; Rupert, 2009) as cognition and largely for the same reason.

the hypothesis that there are underlying mechanisms responsible for cognition. In other words, it is the mechanisms, not the theoretical notion that does the integrative work in this proposal. That seems to be very close to the idea in this project.

A second account of unification would claim that cognitive science is and should be integrated by its *methodology*. For example, 90% of theoretical papers in top cognitive science journals are based on computational models (Busemeyer & Diederich, 2010), so one could claim that computationalism is the methodological unifier of the field. However, computational modeling is used in various ways by different theoretical approaches (Miłkowski, 2013) and in other disciplines, such as sociology (Sawyer, 2004). In other words, it does not guarantee any kind of deeper theoretical unity. Even more restricted computational methodologies (such as the Neural Engineering Framework [Eliasmith & Anderson, 2003]) do not guarantee that models have the same set of theoretical assumptions, even if they can be used to link the neural and the cognitive level of explanations (Miłkowski, 2012). They do that, however, only when accompanied with a hypothesis about underlying cognitive mechanisms.

According to the third view listed above (which is the *received view on the unity of science*), the goal of the unification is to create a single, universal theory. Hence, in this view, the most important relationships are inter-theoretical, reduction being the most prominent. However, it is neither realistic nor desirable to build a single theory in fields dealing with complex phenomena (Levins, 1966; Weisberg, 2007; Wimsatt, 2001, 2007). Building multiple independent, usually highly idealized, models of phenomena may be much more useful than replacing them prematurely with a single theory. The core of the argument is that more robust results can be expected when they are produced independently by multiple statistically-independent models than when they are generated with just one of them, as long as they draw from the same evidence base. Moreover, by testing how multiple theories or models match available evidence, researchers can compensate their confirmation bias, which makes all people prone to making a mistake of prematurely discarding alternative hypotheses (Farrell & Lewandowsky, 2010). Additionally, for special sciences such as biology or geography, it is simply neither realistic nor practical to reduce them to a fundamental physical theory, such as quantum mechanics. In brief, there are genuine advantages in having multiple theories dealing with the same phenomena to be explained.

The fourth view is associated with unification via *cognitive or robotic architectures*. Allen Newell and Herbert Simon saw psychology as offering micro-theories, or theories of very limited scope, that did not contribute to a common view of the human mind (Newell, 1973). Newell, as an alternative, proposed to unify the research program in cognitive science with a notion of a cognitive architecture (Newell, 1990). In his opinion, a cognitive architecture can be used for creating multiple micro-theories and offers a unifying perspective on how the mind works. In a more contemporary context, researchers from the field of cognitive robotics suggested that unified cognitive-robotic architectures could be used to unify research efforts (Morse, Herrera, Clowes, Montebelli, & Ziemke, 2011).

However, unified theories of cognition are not ways to unify *fields* of research. They may broaden the scope of theories but need not cross-fertilize the field as such, and, as Herbert Simon has stressed already in 1998 (Taatgen & Anderson, 2010), they usually stress the systems as a whole instead of mechanisms in the systems that make cognition possible. To theoretically integrate the whole field, there must be stress on such mechanisms.

The fifth view is that behavioral sciences should proceed by acknowledging *vertical integration*: "The natural sciences are already mutually consistent: the laws of chemistry are compatible with the laws of physics, even though they are not reducible to them … the natural sciences are understood to be continuous" (Barkow, Cosmides, & Tooby, 1992, p. 4). Barkow et al. complain that in behavioral and social sciences such is not the case. These fields proceed in splendid mutual isolation. But, Barkow et al. do not specify exactly how one should understand this vertical integration or compatibility of sciences. For this reason, the view is not entirely satisfactory.

The most recent proposal of a general account of inter-theoretic and inter-model relationship is to use the notion of a *constraint* (Danks, 2014). The weakest kind of constraint is a truth-constraint: two bodies of knowledge satisfy a truth-constraint just in case they can be both true at the same time. The notion of truth-constraint can then be used to precisely spell out the notion of vertical integration; namely, two bodies of knowledge are vertically integrated if and only if one body of knowledge truth-constraints the other and the first body cannot (in some weak sense) be false. (Notice that this definition accounts for *vertical integration* in a somewhat deflationary manner. The relationship is asymmetric in such a case, when normally, truth-constraining is horizontal, i.e., no body of knowledge is presupposed to be definitely true or more reliable than the other.)

However, truth-constraining is a weak relation of logical coherence. The wave theory of light does not exclude the particle theory of light, so they satisfy the (horizontal) truth-constraint, even if they propose a completely different account of the basic nature of light. One stronger constraint concerns the nature of entities and processes (activities) presupposed by both theories. Both theories of light no longer satisfy such a constraint, unless a unifying theory is proposed; one that holds that light has both the nature of a particle and wave at the same time.

The account proposed under this project can be understood as a developed and thoroughly mechanistic elaboration of the idea of constraining. It is also pioneering in that it will strive to understand multiple model idealization (Levins, 1966; Weisberg, 2007; Wimsatt, 2001) in a mechanistic manner. It is useful to distinguish mechanistic integration from non-mechanistic accounts of inter-model integration. As applied to cognitive modeling, the weak truth-constraints may apply merely to the model's output or products of cognitive processes. Such modeling has been defended as offering a kind of non-causal and non-mechanistic explanation (Irvine, 2014). However, as most models of cognition aim at explaining cognitive processes, the kind of integration offered by non-mechanistic explanations is shallow, and it does not allow finding common entities and processes in various models. For this reason, such integrations and unifications are not entirely satisfactory, and from the mechanistic point of view, they are incomplete. This project will try to offer ways to integrate incompatible models and offer advice in face of conflicting assumptions.

The mechanistic approach to integration has not been applied to cognitive science, and our current understanding of unification is mostly based on practices of life sciences. However, cognitive science is a diverse field, and methodological controversies are still far from being resolved. For this reason, the project may arrive at normative advice that is sought in some of the discussions over the proper level of explanation of cognitive phenomena. A mechanistic account may also help understand that simplistic greedy reductionism is detrimental for the purposes of interdisciplinary collaboration.

In terms of the growing Polish research network in cognitive science, the significance of the project lies in that it will strive to offer advice to properly balance research at different levels of organization of cognitive systems. It will also help realize the full research potential of young scholars interested in the philosophy of cognitive sciences in Poland. The growing number of researchers involved in the cognitive science program at the University of Warsaw (many of which come from other institutions) is a prime example of successful interdisciplinary and inter-institutional collaboration. Having a strong team of theorists of cognitive science may also help establish the Warsaw program internationally.

More generally, the project will elucidate theoretical and methodological controversies and offer a new mechanistic alternative to slightly outdated advice given by the received view in the philosophy of science, which still remains the basis of the curriculum for cognitive sciences, even if it is disconnected from their research practice. Integration and unification are vital issues for any interdisciplinary field (Chmielewski, Dudzikowa, & Grobler, 2012) but still remain poorly understood, and disintegration may lead to waste of time, energy, and overall duplication of effort in multiple disciplines tackling the same problem in different ways; for this reason, an account of such issues informed by current research is significant both for philosophy and for cognitive sciences in general.

Work plan

The focus of the project will be on studies of unification and integration within cognitive science, also in the case when integration is lacking. The starting point will be the neo-mechanistic account of modeling; it will be applied to explanations in cognitive science that had not been analyzed this way before, in particular when the research in question is integrative, e.g., in computational, embodied, and distributed accounts of cognition (Hutchins, 1995). New proposals for integrations will also be studied, for example the recent account of cognition in terms of predictive coding (Clark, 2013; Hohwy, 2013).

The principal investigator is an external faculty member in the Cognitive Science program at the University of Warsaw, and the project team will collaborate with other members of the faculty. At the same time, co-investigators will be appointed based on an open competition.

After appointing the team (2015), case studies will be planned to be performed by the project members. In the first competition, one post-doc collaborator and one PhD student will be appointed.

In the second competition (2016), two post-doc collaborators will be appointed as well as a PhD student with a documented experience in empirical research. A project workshop is planned, with speakers invited from abroad. It will be in the series of Kazimierz Naturalist Workshops (KNEW), which have been successfully established internationally for the past ten years, and co-organized by the principal investigator. The workshop should help to find contributions to the research topic in an open access journal such as Frontiers in Theoretical and Philosophical Psychology, and should develop further international collaboration. The work on the special issue of another journal will start.

In the third year (2017), there will be an effort to organize a symposium at an international conference devoted to integration and unification within cognitive science. The research topic will be published.

In the fourth year (2018), the special issue will be published. In the final year (2019 / 2020) the focus will be on project outreach and international collaboration.

Throughout the duration of the project, there will be project seminars for team members every month to facilitate collaboration, as well as meetings with invited guests from different disciplines. There will be regular contributions to the project blog, with invited contributions from other researchers, including members of the group blog Brains (http://philosophyofbrains.com). (The principal investigator is one of the contributors to Brains). Project team members will participate in conferences, workshops, and short research visits that will enable both better project outreach and more "networking" opportunities.

Methodology

The project team will analyze explanatory models and frameworks in cognitive sciences, both contemporary and historical. In particular, specific case studies of unification and integration in cognitive sciences will be performed. Based on such cases, some programmatic principles of unification will be sought.

The method used in the project differs from the conventional rational reconstruction in that it accounts for the context of discovery of science, i.e., it accounts for the heuristic value of some explanatory strategies; apart from conceptual analysis, facts from the history of cognitive sciences will be important. The primary research perspective employed in the project is the new mechanism in the philosophy of science. The methodology has already been successfully applied to the study of neuroscience (Craver, 2007), cognitive science (Bechtel, 2008), and life sciences (Craver & Darden, 2013), as well as computational explanation in cognitive sciences (Miłkowski, 2013).

Appointing a new scientific team

Performing multiple case studies requires not only philosophical and historical skills but also some scientific competence in the field under study, which makes the proposed project interdisciplinary. For this reason, the appointment of a team will make it easier to further the goals of the project. The Institute of Philosophy and Sociology of the Polish Academy of Sciences is numbered among the top philosophical research institutions in Poland, categorized in the latest evaluation as "A+". It offers a good milieu for top quality research work.

Another important reason for appointing a team is to build a strong scientific group that will be acknowledged internationally. There are already some grounds for this group; soon, Dr. Paweł Gładziejewski, a philosopher of cognitive science, will become a new member of the Section of Logic and Cognitive Science in the post-doc program FUGA of the National Science Centre (under the supervision of the principal investigator in this project). Although there are Polish researchers in cognitive science and in the philosophy of cognitive science who have had their work published in important international journals, their efforts are usually individual, which makes the career of talented junior researchers unnecessarily difficult. The aim of appointing the team is to prevent a waste of talent, and to build a strong team of cognitive science theorists. The members of the team will be appointed on a competitive basis.

Literature references

Adams, F. (2010). Why we still need a mark of the cognitive. *Cognitive Systems Research*, *11*(4), 324–331. doi:10.1016/j.cogsys.2010.03.001

Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Blackwell Pub.

Adams, F., & Garrison, R. (2012). The Mark of the Cognitive. *Minds and Machines*, *23*(3), 339–352. doi:10.1007/s11023-012-9291-1

Arbib, M. A., Baker, C. L., Bresnan, J., D'Andrade, R. G., Kaplan, R., Keyser, S. J., … Zurif, E. (1978). *COGNITIVE SCIENCE, 1978*.

Barkow, J. H., Cosmides, L., & Tooby, J. (1992). *The Adapted Mind. Evolutionary Psychology and The Generation of Culture*. (J. H. Barkow, L. Cosmides, & J. Tooby, Eds.). New York and London: Oxford University Press.

Bechtel, W. (2008). *Mental Mechanisms*. New York: Routledge (Taylor & Francis Group).

Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press.

Buckner, C. (2013). A property cluster theory of cognition. *Philosophical Psychology*, (September 2014), 1–30. doi:10.1080/09515089.2013.843274

Busemeyer, J. R., & Diederich, A. (2010). *Cognitive modeling*. Los Angeles: Sage.

Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *NeuroImage*, *63*(1), 289–300. doi:10.1016/j.neuroimage.2012.07.004

Castelfranchi, C. (2014). For a science of layered mechanisms: beyond laws, statistics, and correlations. *Frontiers in Psychology*, *5*, 536. doi:10.3389/fpsyg.2014.00536

Chemero, A. (2000). Anti-Representationalism and the Dynamical Stance. *Philosophy of Science*, *67*(4), 625 – 647.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, Mass.: MIT Press.

Chmielewski, A., Dudzikowa, M., & Grobler, A. (Eds.). (2012). *Interdyscyplinarnie o interdyscyplinarności: między ideą a praktyką*. Kraków: Oficyna Wydawnicza Impuls.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*(3), 181–204. doi:10.1017/S0140525X12000477

Craver, C. F. (2007). *Explaining the Brain. Mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.

Craver, C. F., & Darden, L. (2013). *In search of mechanisms: discoveries across the life sciences*.

Dale, R., Dietrich, E., & Chemero, A. (2009). Explanatory Pluralism in Cognitive Science. *Cognitive Science*, *33*(5), 739–742. doi:10.1111/j.1551-6709.2009.01042.x

Danks, D. (2014). *Unifying the mind: cognitive representations as graphical models.* Cambridge, Mass.: MIT Press.

Darden, L., & Maull, N. (1977). Interfield Theories. *Philosophy of Science*, *44*(1), 43–64.

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. New York: Simon & Schuster.

Dominiak, A., & Nosal, C. S. (2008). Rola mechanizmu temporalnej integracji poznawczej, afektu i wyobraźni w procesie przewidywania zdarzeń przyszłych. *Przegląd Psychologiczny*, *51*(2), 215–233.

Dupré, J. (1993). *The disorder of things: metaphysical foundations of the disunity of science*. Cambridge Mass.: Harvard University Press.

Eliasmith, C., & Anderson, C. H. (2003). *Neural Engineering. Computation, Representation, and Dynamics in Neurobiological Systems*. Cambridge, Mass.: MIT Press.

Eronen, M. I. (2014). Levels of organization: a deflationary account. *Biology & Philosophy*, (January). doi:10.1007/s10539-014-9461-z

Estany, A., & Martínez, S. (2013). "Scaffolding" and "affordance" as integrative concepts in the cognitive sciences. *Philosophical Psychology*, (November), 1–14. doi:10.1080/09515089.2013.828569

Farrell, S., & Lewandowsky, S. (2010). Computational Models as Aids to Better Reasoning in Psychology. *Current Directions in Psychological Science*, *19*(5), 329–335. doi:10.1177/0963721410386677

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, *28*(2), 97–115. doi:10.1007/BF00485230

Francuz, P. (2012). W jakim zakresie normatywne podejście do metodologii badań naukowych w psychologii jest otwarte na eksplorację terra incognita? *Roczniki Psychologiczne*, *XV*(3), 41–46.

Frégnac, Y., & Laurent, G. (2014). Neuroscience: Where is the brain in the Human Brain Project? *Nature*, *513*(7516), 27–29. doi:10.1038/513027a

Garzon, F. C. (2008). Towards a General Theory of Antirepresentationalism. *The British Journal for the Philosophy of Science*, *59*(3), 259–292. doi:10.1093/bjps/axl007

Gervais, R. (2014). A framework for inter-level explanations: Outlines for a new explanatory pluralism. *Studies in History and Philosophy of Science Part A*, *48*, 1–9. doi:10.1016/j.shpsa.2014.07.002

Hohwy, J. (2013). *The predictive mind*. New York: Oxford University Press.

Hutchins, E. (1995). How a Cockpit Remembers Its Speeds. *Cognitive Science*, *19*(3), 265–288. doi:10.1207/s15516709cog1903_1

Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: basic minds without content*. Cambridge Mass.: MIT Press.

Illari, P. M., & Williamson, J. (2011). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, *2*(1), 119–135. doi:10.1007/s13194-011-0038-2

Irvine, E. (2014). Models, robustness, and non-causal explanation: a foray into cognitive science and biology. *Synthese*, n/a–n/a. doi:10.1007/s11229-014-0524-0

Lauwereyns, J. (2011). *The anatomy of bias: how neural circuits weigh the options*. Cambridge Mass. ;;London: MIT Press.

Levins, R. (1966). The Strategy of Model Building in Population Biology. *American Scientist*, *54*(4), 421–431.

Lewicka, M. (2012). *Psychologia miejsca*. Warszawa: Wydawnictwo Naukowe Scholar.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science*, *67*(1), 1–25.

Miłkowski, M. (2012). Theoretical Unification and the Neural Engineering Framework. In J. Stelmach, B. Brożek, & Ł. Kurek (Eds.), *Philosophy in Neuroscience* (pp. 85–112). Kraków: Copernicus Center Press.

Miłkowski, M. (2013). *Explaining the Computational Mind*. Cambridge, Mass.: MIT Press.

Morse, A. F., Herrera, C., Clowes, R., Montebelli, A., & Ziemke, T. (2011). The role of robotic modelling in cognitive science. *New Ideas in Psychology*, *29*(3), 312–324. doi:10.1016/j.newideapsych.2011.02.001

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, Mass. and London: Harvard University Press.

Paprzycka, K. (2005). *O możliwości antyredukcjonizmu*. Warszawa: Semper.

Posner, M., & Reichle, M. E. (1994). *Images of mind*. New York: Scientific American Library.

Rączaszek-Leonardi, J., & Kelso, J. A. S. (2008). Reconciling symbolic and dynamic aspects of language: Toward a dynamic psycholinguistics. *New Ideas in Psychology*, *26*(2), 193–207. doi:10.1016/j.newideapsych.2007.07.003

Rowlands, M. (2009). Extended cognition and the mark of the cognitive. *Philosophical Psychology*, *22*(1), 1–19. doi:10.1080/09515080802703620

Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Oxford: Oxford University Press.

Sawyer, K. (2004). Social explanation and computational simulation. *Philosophical Explorations*, *7*(3), 219–231. doi:10.1080/1386979042000258321

Simon, H. A. (1996). *The sciences of the artificial*. Cambridge, USA, MA: MIT Press.

Taatgen, N., & Anderson, J. R. (2010). The Past, Present, and Future of Cognitive Architectures. *Topics in Cognitive Science*, *2*(4), 693–704. doi:10.1111/j.1756-8765.2009.01063.x

Thagard, P. (2007). Coherence, Truth, and the Development of Scientific Knowledge. *Philosophy of Science*, *74*, 28–47.

Trout, J. D. (2008). Seduction without cause: uncovering explanatory neurophilia. *Trends in Cognitive Sciences*, *12*(8), 281–2. doi:10.1016/j.tics.2008.05.004

Varela, F. J. (1999). The specious present: A neurophenomenology of time consciousness. In J. Petitot, F. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 266–314). Stanford, CA: Stanford University Press.

Weisberg, M. (2007). Forty Years of “The Strategy”: Levins on Model Building and Idealization. *Biology & Philosophy*, *21*(5), 623–645. doi:10.1007/s10539-006-9051-9

Weisberg, M. (2013). *Simulation and similarity: using models to understand the world*. New York: Oxford University Press.

Wilson, E. (1998). *Consilience: the unity of knowledge*. New York: Knopf ;Distributed by Random House.

Wimsatt, W. C. (2001). Richard Levins as philosophical revolutionary. *Biology and Philosophy*, *16*(1), 103–108.

Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: piecewise approximations to reality*. Cambridge Mass.: Harvard University Press.