



Knowledge transfer, templates, and the spillovers

Chia-Hua Lin^{1,2}

Received: 18 June 2020 / Accepted: 19 October 2021 / Published online: 8 January 2022
© Springer Nature B.V. 2022

Abstract

Mathematical models and their modeling frameworks developed to advance knowledge in one discipline are sometimes sourced to answer questions or solve problems in another discipline. Studying this aspect of cross-disciplinary transfer of knowledge objects, philosophers of science have weighed in on the question of whether knowledge about how a mathematical model is previously applied in one discipline is necessary for the success of reapplying said model in a different discipline. However, not much has been said about whether the answer to that epistemological question applies to the reapplication of a modeling framework. More generally, regarding the nature of the production of knowledge in science, a metaphysical question remains to be explored whether historical contingencies associated with a mathematical construct have a genuine impact on the nature—as opposed to sociological practices or individual psychology—of advancing scientific knowledge with said construct. Focusing on this metaphysical question, this paper analyzes the use of mathematical logic in the development of the Chomsky hierarchy and subsequent reapplications of said hierarchy; with these examples, this paper develops the notion of “spillovers” as a way to detect cross-disciplinary justifications for better understanding the relations between reapplications of the same mathematical construct across disciplines.

Keywords Knowledge transfer · Formal templates · Model templates · The Chomsky hierarchy · Cross-disciplinary reapplications · Modeling frameworks

1 Introduction

Mathematical models and their modeling frameworks which were originally developed to advance knowledge in one scientific discipline are sometimes sourced to answer questions or solve problems in another discipline. Philosophers of science who study

✉ Chia-Hua Lin
clin.chiahua@gmail.com

¹ College of Technology Management, National Tsing Hua University, R513, TSMC Building, No.101, Section 2, Kuang-Fu Road, Hsinchu, Taiwan 30013, Republic of China

² Corcoran Department of Philosophy, University of Virginia, Charlottesville, VA, USA

how knowledge transfers across disciplines have contemplated whether knowledge about how a mathematical model is previously applied in one discipline is necessary for the successful applications of that model in a different discipline (Bradley & Thébault, 2019; Herfeld & Lisciandra, 2019; Humphreys, 2019). However, not much has been said about whether the answer to that epistemological question applies to the reapplication of a modeling framework. In terms of the production of knowledge in science, a metaphysical question remains as to whether historical contingencies associated with a mathematical construct have a genuine impact on the nature—as opposed to sociological practices or individual psychology—of advancing scientific knowledge with said construct. Focusing on this metaphysical question and using modeling frameworks as examples, this paper develops the notion of “spillovers” to better understand the relations between reapplications of the same mathematical construct across disciplines. Responses in the literature to the epistemological question, especially how they relate to the metaphysical question, will be discussed. The recent literature on model transfer includes a prominent trend to analyze the cross-disciplinary use of mathematical models either as the transfer of a modeling framework (Knuuttila & Loettgers, 2014, 2016) or as the constructing and adjusting of a “template” in a given discipline (Humphreys, 2002, 2004, 2019). Relatedly, philosophers of science in the field of knowledge transfer ask how knowledge, including “objects of knowledge” (i.e., mathematical models or theories), transfer across the sciences to solve problems or answer questions for which that they were not originally developed for (Humphreys, 2019; Houkes & Zwart, 2019; Zuchowski, 2019; Price, 2019; see Herfeld & Lisciandra, 2019 for a review).¹ In both literatures, various notions of templates that characterize different aspects of scientific modeling have generated many insights. However, because the discussion has primarily focused on models such as differential equations, it remains unclear how the same notions may be applied to study models in different formats, such as game-theoretic models (Grüne-Yanoff, 2011) or modeling frameworks, and whether relevant insights will hold in these differing formats. As a point of departure, this paper discusses Humphreys’ (2002, 2004, 2019) template-based analysis of model transfer, relating it to his argument in a current debate about the epistemology of knowledge transfer between the Kuhnian approaches, which emphasize the role of learning from exemplars, and his (2019) approach, which suggests learning through explicitly building templates. This paper shows how Humphreys’ template-based analysis may indeed be productively applied to study the reapplication of a modeling framework across disciplines, while also arguing that some but not all insights from Humphreys’ approach apply to modeling frameworks.

Emerging from my analysis of Humphreys’ argument is a conjecture that historical contingencies are irrelevant to justifying a new knowledge-claim as the epistemic output of reapplying a single mathematical construct across domains or disciplines. This ahistoricist conjecture has an epistemological consequence; if it holds, then knowledge about how a mathematical construct is previously applied is helpful but not necessary for successfully reapplying it in one’s present context. Conversely, as opponents of Humphreys’ ahistoricist conjecture may argue, if it can be shown that knowledge from a mathematical construct’s previous application in one discipline

¹ In this paper, “knowledge transfer” refers to applying objects of knowledge in this sense.

is necessary for successfully reapplying said construct in another discipline (hereafter called “cross-disciplinary knowledge”), then we may have a reason to reconsider the conjecture and, consequently, Humphreys’ epistemology of knowledge transfer. Typical candidates for cross-disciplinary knowledge include knowledge about the modeling practice in the discipline from which the model is sourced, which some argue is required for the users in a different discipline to first identify the idealizing assumptions embedded in the mathematical construct, and then determine whether they are appropriate idealizations to be made in their reapplication (Bradley & Thébault, 2019). Similarly, others contend that cross-disciplinary knowledge is needed for interpreting said construct or understanding its epistemic potential so as to replicate it in another (Knuuttila & Morgan, 2019; Herfeld & Doehne, 2019).

In contrast to current approaches, which tend to analyze cross-disciplinary knowledge *as input* to a successful reapplication, this paper analyzes cross-disciplinary knowledge from its *output*, showing that in some cases, advancing knowledge by reapplying a cross-disciplinarily sourced mathematical construct may require a knowledge-claim produced by a former (re)application of the same construct. This paper introduces the notion of “spillovers” to capture the justificatory role certain cross-disciplinary knowledge plays in knowledge transfer. A spillover is a knowledge-claim that is indispensable to the justification of another knowledge-claim whereby 1) both knowledge-claims are products of applying the same mathematical construct, and 2) the two knowledge-claims originate in different disciplinary contexts. With this notion in mind, the relations between two cross-disciplinary applications of a single mathematical construct may be analyzed as either truth-functional or non-truth-functional, a distinction that is necessary to address the metaphysical question regarding knowledge transfer.

This paper proceeds with the following five sections. Section 2 elaborates on how the ahistoricist conjecture from Humphreys’ epistemology of knowledge transfer lends support to Humphreys’ response—which I call the self-sufficient view—to epistemological debates in the knowledge transfer literature. Section 3 argues that the self-sufficient view should be understood as a description of how an ideal scientist may learn to reapply a mathematical construct; it follows that the current accounts of how reapplications actually take place may not pose effective challenges to the self-sufficient view. Section 4 examines two contrasting examples: 1) the development of the Chomsky hierarchy in linguistics, which is consistent with Humphreys’ self-sufficient view, and 2) subsequent reapplications of the Chomsky hierarchy in theoretical computer science and cognitive biology, which illustrate the presence of a spillover. Section 5 analyzes how the presence of spillovers—by exposing truth-functional dependency between episodes of knowledge transfer—addresses the metaphysical question of knowledge transfer. In the final section, I conclude that the self-sufficient view is unnecessarily strict in directing practicing scientists to intentionally avoid spillovers.

2 Humphreys’ template-based analysis and the self-sufficient view

In his paper “Knowledge Transfer Across Scientific Disciplines” (2019, 112), Humphreys asks ‘how can a single formal representation be successfully applied to multiple scientific domains that *prima facie* have very different subject matters?’ and

proceeds to show how such reapplication can in some cases be achieved without knowledge about the source domain. Pivotal to Humphreys' analysis is his notion of formal templates and the processes of constructing and evaluating such templates. A formal template is a scheme of variables, stripped of empirical content from its former application, and detached from the phenomenon of interest in the source domain. Void of empirical content, what remains is 'a purely mathematical object ... that can be carried over wholesale from domain to domain' (Humphreys, 2019, 117). With the notion of formal templates, Humphreys suggests that '[t]he practitioners of the field to which the template is currently applied do not need to know the details of how it is applied in other domains and the application does not require a relation of analogy between systems in different fields' (ibid.).

Humphreys' template-based approach to knowledge transfer contrasts with other approaches that cite perceived similarities, analogical reasoning, or tacit knowledge to account for the reapplication of mathematical models across different domains (e.g., Hesse, 1964, 1966; Kuhn, 1970; see also Knuuttila & Loettgers, 2020 for a recent discussion). For instance, consider Kuhn's (1974) account of how certain equations, such as $f=ma$ in Newton's Second Law of Motion, function 'like schematic forms' (Kuhn, 1974, 465, as quoted in Humphreys, 2019). When applying what Kuhn calls a "symbolic generalization" (i.e. a general equation) to specific scenarios, either side of the equation will be substituted with more detailed expressions. In solving problems of free fall or simple pendulum motion, for example, $f=ma$ becomes $mg=md^2s/dt^2$ and $sinh=md^2s/dt^2$, respectively; the disparity between $f=ma$ and its actual form is yet greater in 'more interesting mechanical problems' such as 'the motion of a gyroscope' as Kuhn (1974, 465) observes. According to Humphreys (2019, 113), Kuhn approached this aspect of scientific practice from the perspective of learning. In particular, there are 'two types of learning: [first] acquiring the initial knowledge of how to apply a given symbolic generalization,' which students learn by studying the use of exemplars, 'and [second] subsequently developing the skills to know that the particular symbolic generalization can be applied to other systems'. Yet, for Humphreys, what is missing in Kuhn's account is 'an analysis of how the second kind of skill, *that of knowledge transfer*, is learned and employed beyond an appeal to resemblance and similarity relations' (2019, 113, emphasis mine). Humphreys' template-based analysis is meant to address this gap. Instead of addressing reapplications of a general equation within physics as Kuhn did, Humphreys extends the discussion to cross-disciplinary reapplications. The goal of his approach is to show that, regardless of intra- or cross-disciplinary reapplication, 'the process of applying templates can in some cases be made explicit ... we do not need to rely on similarity relations and tacit knowledge' (Humphreys, 2019, 112).

Rooted in Humphreys' method of addressing the epistemological issue of knowledge transfer, I argue, is a will to the truth of the ahistoricist conjecture. Specifically, while historical contingencies may influence scientists' psychology or norms of scientific practice, they do not have a genuine impact on the nature of the production of scientific knowledge. To illustrate, consider an array of contexts. Each context consists of a subject matter, a phenomenon of interest, and a set of methods for approaching the phenomenon of interest within a particular subject matter. Consider

also that owing to historical contingencies, this array has a temporal order such that one context is the first instance where said mathematical structure was initially constructed. Moreover, in this initial instance, as well as in all subsequent contexts, one finds a mathematical construct implemented in one way or another for solving problems or answering questions about one aspect of the world with some empirical success. Furthermore, because these contexts share neither their subject matter nor their phenomenon of interest (perhaps not even the totality of their methods), the fact that all these instances share a mathematical construct in problem-solving practice is quite spectacular. About such an array, one may ask: Does its temporal order impose epistemological constraints such that advancing knowledge in a future context would be impossible without at least one prior successful occurrence?

Humphreys' stance in favor for the ahistoricist conjecture can be observed in his characterization of knowledge transfer. '[H]istorically, some successful models [may] have been introduced ... as unanalyzed representations and that subsequent applications can be and are made by analogical reasoning from previous applications' (Humphreys, 2019, 114). However, the point Humphreys stresses is that, ultimately, 'the empirical justification' for the application 'rests on the satisfaction of the construction assumptions *in the new domain*' (ibid., 114-5, emphasis added).

One consequence of Humphreys' ahistoricist conjecture is what I call the self-sufficient view. This view is, to rephrase Humphreys' own words, about how a scientist can reapply 'a single formal representation' to 'multiple scientific domains that *prima facie* have very different subject matters' without cross-disciplinary knowledge (Humphreys, 2019, 112). In my interpretation of Humphreys' position, the self-sufficient knowledge transfer occurs when scientists who source a piece of mathematical construct from a different discipline 'do not need to know the details of how it is applied in other domains' (2019, 117). To clarify this argument, consider the self-sufficient view as founded on two premises:

- Premise 1: reapplying a mathematical object of knowledge can be conceived of as independent template building events in different contexts, whereas
- Premise 2: the success of such reapplication is assessed based on the contribution of its end-product (e.g., a model) to the new context.

Premise 1 is essentially Humphreys' ahistoricist conjecture, the main idea of which is that advancing knowledge through the reapplication of mathematical objects can be conceived of as independent events of template building; hence, one needs only to consider the components of the construct rather than the complex specificities of its origination.

To unpack Humphreys' (2019, 113) epistemology of knowledge transfer, one needs to begin with his (2004) account of computational models. According to Humphreys, a computational model can be expressed through six components: <Template, Construction Assumptions, Interpretation, Initial Justification, Correction Set, Output Representation> (Humphreys, 2004, 103). Analyzing these components substantiates Premises 1 and 2, thereby lending support for the self-sufficient view. Although Humphreys speaks of the components of a model, his analysis offers a framework for studying modeling practice, such as what one does with these

components. Essentially, this conceptual framework offers a useful way to study the reapplications of a given modeling framework – a point I shall return to in Section 4. Let's first take a closer look at each component of Humphrey's account.

Best understood in pairs, the first two components in Humphreys' account, "Template" and "Construction Assumptions," refer to, obviously, a template and all the assumptions needed for building said template. A template in this sense is a mathematical scheme of variables whose syntax determines how variables, including coefficients, relate to one another. Similarly, the assumptions that go into constructing the template are concerned with formulating a phenomenon of interest into a mathematically treatable target system. For instance, ontological assumptions specify which of the properties or entities of the phenomenon of interest will be targeted for representation in the template. Idealizing assumptions are decisions about how certain properties or entities in the phenomenon will be represented but in a distorted manner. Other assumptions include abstraction, approximation, and physical constraints known to apply to the phenomenon in question. The goal of compiling these assumptions is to describe the phenomenon with a suitable ontology so that a question regarding the target system can be formulated in a mathematically tractable way. Humphreys (2004, 76) calls this assortment of assumptions the "construction assumptions" and the process of integrating them the 'process of construction.'

The third and fourth components, "Initial Justification" and "Interpretation," refer to the fact that all of the assumptions surrounding an ontology of the phenomenon of interest enable an initial justification for using the template. Conversely, this initial justification is bound, to some extent, by said ontological interpretation. For this reason, using the same template in different contexts with a radically different ontology (i.e., a radically different empirical mapping) may seem problematic (Bradley & Thébault, 2019), but this is not necessarily the case. We will come back to this point in Section 3.

The final two components, "Correction Set" and "Output Representation," refer to adjustments made to the template based on relevant empirical data and the template's output at the end of an inquiry. Having rendered a phenomenon of interest into a target system with a mathematically tractable template is a good start, but this rendering alone is insufficient for completing an inquiry; it must also speak to the available and relevant empirical data. When incongruence exists between data and the resulting template, adjustments are made, for instance, to relax or refine the construction assumptions, or to mildly revise the ontological assumptions if necessary. Humphreys (2004, 76) calls this step the 'process of adjustment.' The goal of this process is to provide an accurate representation of the phenomenon, or at least one that is superior to other approaches in the same context.

Together, the process of construction and the process of adjustment constitute the foundation of modeling as template building. Something needs to be said about the claim in Premise 1 that template building occurs independently: it amounts to saying that the content that is necessary for completing the building processes is fully contained within the construction assumptions and the correction set. To Humphreys, practicing scientists may have looked into a model's prior success for inspirations. But this does not interfere with the epistemology of knowledge transfer since in his view template-building activities are logically independent from one another. The

argument for this ahistoricist claim is that everything a template builder *needs* for the building processes is contained within knowledge specific to her field of specialty and general mathematical competence.

Moreover, regarding Premise 2, the success of repurposing a mathematical construct is determined by the construct's ability to meet the objectives of tractability and improved empirical accuracy in the new domain. Note that both tractability and empirical accuracy are sensitive only to one context, i.e., the subject matter, its associated phenomenon of interest, and the methodology (e.g., data collection, experiment) that is relevant to them. Thus, tractability and empirical accuracy do not seem to require content from contexts outside of the context at hand. Consequently, the success of reapplying a mathematical construct is assessed based on the contribution of its end-product in the new context. Hence, we have arrived at the self-sufficient view: Context-specific knowledge and general mathematical competence are sufficient for successfully repurposing a mathematical construct in a new context.

One may ask: How does the independency of template building explain the fact that some constructs have applications across multiple contexts that may or may not be related but clearly involve analogical reasoning? Humphreys' answer to this question is twofold. On the one hand, as in some cases, the same template can be (or could have been) constructed with different sets of construction assumptions concerning different phenomena of interest (see Humphreys, 2004, 73-76). On the other hand, as in other cases, a popular template can be (or could have been) constructed with mathematical content alone and thus has no interpretation beyond mathematical interpretation (as he discussed in 2019, 116). Humphreys refers to these cases, respectively, as theoretical templates and formal templates.

According to him, a theoretical template contains at least one schematic property variable which can be substituted by different predicates such that the template functions as a general representational device. Humphreys' (2004, 60) paradigmatic example is, again, the equation in Newton's Second Law: it 'describes a very general constraint on the relationship between any force, mass, and acceleration,' and when 'all of the schematic variables have been substituted for,' the equation 'can be successfully used to represent a variety of different phenomena within the domain of the theory' where the template occurs (2019, 114).

In contrast, a formal template contains only mathematical interpretation, and all of its construction assumptions are mathematical in nature. By mathematical, Humphreys (2019, 114 footnote 9) includes 'representations from mathematical logic and some programming languages.' This point concerning mathematical logic opens the possibility of analyzing a modeling framework as a template, which I will explore in Section 4. All that said, in either case, achieving tractability and improving empirical accuracy *in one context* only requires general mathematical competence and knowledge specific to said context.

For instance, consider the case in which a template user builds a template from scratch. Recall that the template builder is in charge of conjuring assumptions that are directly relevant to the subject matter during the processes of construction and adjustment. This step requires the builder's command of mathematics to identify appropriate mathematical operators with which to render a phenomenon of interest mathematically tractable. Moreover, Humphreys (2004) suggests that the builder,

having gone through the process of construction, would know which assumptions to adjust and how to adjust them when needed. Everything in the correction set is supposedly available in the context from which she drew her construction assumptions. That is, the knowledge about how to relax, refine, and revise the assumptions about the phenomenon with regard to a particular subject matter originates from the context that she is working in. Therefore, the template's prior success (i.e., through a modeling effort in a different context) is irrelevant to completing her inquiry.

Needless to say, not all template users are template builders. How do we account for the scenario in which the user does not build the template from scratch? This is where general mathematical competence and Humphreys' notion of formal templates become crucial to maintaining the self-sufficient view. Because the same template can be constructed with only mathematical content, it can be conceived of as having no interpretation beyond mathematical interpretation. Using such a template requires the user to go through the process of adjustment as usual but not the regular process of construction. Instead, she would first identify the mathematical construction assumptions and then assign the template with a new ontology. Secondly, she would check whether any adjustments are needed such that her modeling efforts produce an empirically accurate representation. This way, regardless of whether the user built the template from scratch or repurposed it, the self-sufficient view's conclusion is not affected.

3 An ideal scientist and practicing scientist' ideals

In the previous section, I unpacked Humphreys' response to the epistemological debate in the knowledge transfer literature in terms of the self-sufficient view. In this section, I discuss how scholars have responded to the epistemological question and make a claim on how Humphreys' position should be critically assessed. I argue that what Humphreys offers is a description of how an ideal scientist may learn to reapply mathematical constructs, which he extends into a prescriptive claim. I follow Burgess (1992) who makes a distinction between a prescription and a description of idealized scenarios. According to Burgess (1992, 12), a description of idealized scenarios is not meant to be an accurate description of how events took place, yet 'its results can serve as minor premises in arguments with prescriptive major premises leading to prescriptive conclusions.'

Burgess' (1992) observation fits nicely with what Humphreys aims to achieve with the self-sufficient view, i.e., to aspire toward better knowledge transfer practices. As mentioned at the beginning of this section, Humphreys' point is to show that 'the process of coming to know that a particular model applies to a system' *need not* rely on 'similarity relations and tacit knowledge' (2019, 112-3). Humphreys contends that 'it is reasonable to *prefer* explicit criteria over implicit judgments of resemblance' (115, emphasis added) such that the process of knowledge transfer may be made explicit. In particular, by identifying formal templates in knowledge transfer, Humphreys shows that one does not need to leave the process of 'coming to know that a particular model applies to a system overly psychological' (2019, 113). While a combination of analogical reasoning and statistical goodness-of-fit criteria

applied to the output of the model may be playing a heuristic role in real scenarios, in Humphreys' (2019, 114) words, this possibility is 'neither the only nor the best mode of introduction and application.' Thus, the scientist in the self-sufficient view is not meant to describe practicing scientists such as those in Kuhnian accounts of knowledge transfer. Instead, Humphreys' account is more suitably understood as how an ideal scientist would carry out the process of knowledge transfer and consequently why practicing scientists should aspire to the practices of this ideal scientist.

In the context of knowledge transfer, an ideal scientist is competent in mathematics and native to the subject matter of her study. When implementing an existing formal template, she could employ implicit strategies at will but she would not have to. Using her general mathematical competence and knowledge specific to her field of specialty, she could reconstruct the template and provide justification on demand. In other words, when it comes to transferring a formal representational device, an ideal scientist is self-reliant. No cross-disciplinary knowledge is required.

Correctly identifying the nature of the self-sufficient view is of utmost importance when it comes to appropriately assessing it. Critics charge that Humphreys' (2019) account is either incomplete or unrealistic (more on this below). My assessment of Humphreys' position is that while the self-sufficient view is possible in practice, what Humphreys thinks to be an ideal scientist may turn out to be not so ideal because it does not take spillovers into account. Supporting evidence for this claim is discussed in Sections 4.2 and 5. In this section, I discuss one way an ideal scientist can fall short that is germane to presence of a spillover in scientific practice.

Following the self-sufficient view, an ideal scientist is self-reliant only if 1) she deliberately avoids tackling problems that *require* cross-disciplinary knowledge including those that may become part of construction assumptions and hence integral to justifying the final output of the inquiry or 2) all problems or questions in science can *eventually* be solved or answered without cross-disciplinary justifications. Perhaps 2) could be the case, which might be an article of faith but which would render 1) a defensible position if we disregard any sense of urgency to have a solution or an answer before *eventuality*. I shall not discuss whether it is reasonable to presuppose this eventuality but to stress one point: if all scientists avoided tapping into cross-disciplinary justification, as the self-sufficient would advise, certain problems or questions could be left in suspense indefinitely. Indeed, Humphreys is correct in urging for an explicit method over relying on tacit knowledge, but if having any answer at all in the first place is necessary for being able to make explicit the justification, be it within or across disciplinary contexts, then it is not clear that staying away from cross-disciplinary justification is a wise advice to make. That said, many scholars of knowledge transfer seem to share the independency position of the self-sufficient view, including some opponents of the notion of formal templates. In what follows, I turn to discuss these nuances among views related to Humphreys' position.

Many authors in the current debate hold a collective position which takes the self-sufficient view to be unrealistic or incomplete, arguing that cross-disciplinary knowledge of one sort or another is at least advisable for reapplying a mathematical construct in one's present context. Moreover, some of these authors tend to focus on the sociological dimensions of knowledge transfer (e.g., Herfeld & Doehne, 2019,

and to some extent Bradley & Thébault, 2019), yet because such dimensions are entirely absent from Humphreys' view, sociological criticism may inadvertently disregard certain nuanced contrasts between Humphreys' original position and his critics. As a remedy, I shall analyze these differences according to, on the one hand, their acceptance or rejection of the two premises in the self-sufficient view (Section 2) and on the other hand, how they relate to its key implication, i.e., the call to be like an ideal scientist engaged in knowledge transfer. In particular, I demonstrate that this collective position seems to accept the ideal scientist implication, reject Premise 2, yet either presupposes Premise 1 or leaves it unquestioned. Indeed, except for Knuuttila and Loettgers' (2014, 2016, 2020) work on model templates, Premise 1 seems to be implicitly held by many philosophers of knowledge transfer. In Section 5, I shall discuss how the presence of spillovers in science challenges Premise 1. Overall, criticisms converge on the complications of "re-situating" (Morgan, 2014) a mathematical construct in actual scenarios. These scenarios involve, among other things, a new disciplinary context, a new scientific community, or a new target system, or any combinations of the three. For critics, these complications of knowledge transfer *in practice* suggest that successfully reapplying cross-disciplinarily sourced mathematical constructs requires variable kinds of cross-disciplinary knowledge.

One implication of the self-sufficient view is that practicing scientists *should* prefer explicit approaches over tacit strategies.² Learning from exemplars or implicit judgments of resemblance relations are not the best mode of reapplying an existing equation form when one 'can check directly whether the assumptions are satisfied for the system at hand' (Humphreys, 2019, 115). For Humphreys, in some cases, every scientist *can* perform knowledge transfer as an ideal scientist. For instance, general equation forms, such as Laplace's equation, the diffusion equation, Poisson's equation, or several statistical distributions 'transcend specific theories and their subject matter' (Humphreys, 2019, 115). Their construction assumptions can be found in 'the better kind of *Methods* textbooks, ... the conditions for application are laid out explicitly' (ibid., 115 emphasis original). Humphreys (2004, 154) could be arguing that because it is 'possible in practice' to reapply a general equation form like an ideal scientist, one *should* opt into doing so in practice. This argument is consistent with Bradley and Thébault (2019), whose work offers instructions for carrying it out in practice.

Bradley and Thébault (2019) point out that when reapplying a model with a radically different ontology than its new target domain, the idealizing assumptions the model inherited from its previous context need to be explicitly justified, a process they call "re-sanctioning." The reason for re-sanctioning is the uncertainty over whether the justification of those idealization assumptions in the previous context applies in the present context. It is uncertain because justification in a prior context does not automatically migrate to the next context. Moreover, according to them, re-sanctioning requires the user to first isolate the original idealizing assumptions

² Another implication may be that philosophers of science should prefer analyzing knowledge transfer as such, but Humphreys' (2019) text does not decisively exclude either interpretation.

and then justify their counterpart in the present context. In one of their case studies, for example, well-known models developed in the context of statistical mechanics are repurposed to describe wealth distribution. In physics, these models describe the exchange of kinetic energy in a gas through the basis of assuming binary dynamical interactions between the molecules through scattering processes. Although these assumptions are highly idealized, experimental data suggest that they are reasonable assumptions to be made about gases. In this sense, those assumptions about the molecules passed the “sanction” in that particular context. However, in the context of an economic theory of exchange between economic agents, it is not obvious how the same assumptions apply. Note that this is not to say that re-sanctions would certainly fail; instead, one would need to go through the assumptions explicitly, ideally in an itemized manner.

For the most part, Bradley and Thébault’s (2019) view coincides with Humphreys’ analysis. Humphreys considers idealizing assumptions to be part of the constructing assumptions. Hence, both sides emphasize the importance of satisfying the constructing assumptions and advocate for justifying constructing assumptions in an explicit manner for successful reapplication. Both sides also draw a clear boundary between contexts with regard to justifying the construction assumptions, which can be seen as presupposing Premise 1. However, Bradley and Thébault (2019, 82) arrive at the opposite of the self-sufficient view; for them, re-sanctioning requires ‘awareness of modeling practices from both the old and new contexts’ to effectively isolate idealizing assumptions in the old context and to appropriately justify their counterpart in the present context. With regard to knowledge diffusion, for practitioners in the new context to appreciate the newly demonstrated mathematical tractability of a repurposed template, the user of said template needs to convince them that the idealizing assumptions have been properly re-sanctioned. Bradley and Thébault (2019, 90) thus advise concerted cross-disciplinary efforts as one of the norms for successful model transfer. This particular aspect—being responsive to not only the epistemic norms but also the institutional norms in the new context—is entirely absent in Humphreys’ position.

Concerning knowledge diffusion but from another sociological perspective, Herfeld and Doehne (2019) find that before the epistemic potential of a mathematical construct can bear fruit in a new context, it needs to be explored and elaborated upon based on the theoretical and conceptual backdrop of each context. For instance, some effort is required for showing how a repurposed construct, in their case rational choice theories, can be used in solving problems across different sciences. This precondition includes translating related concepts between contexts. By “translation,” Herfeld and Doehne (2019, 65) refer to publications which align a ‘scientific innovation with previous research traditions’ in a way that reveals the innovative idea’s ‘potential for [solving] particular disciplinary problems and establishes the basis for its application in specialist research.’ Using a bibliometric method, they find that the role of a translator is crucial for the epistemic potential of repurposed mathematical constructs to spread within and across scientific communities.

Similar to re-sanctioning, the notion of translation is compatible with the part of Humphreys’ position that recommends explicit over tacit strategies. According to him, converting tacit strategies into explicit constructing assumptions is better than

leaving them implicit. Note that Humphreys does not assume that every practicing scientist is an ideal scientist—the idea is for practicing scientists to become one. Thus, so long as there is one ideal scientist in a research community, others can take up the role of a translator that helps with diffusing newly produced knowledge. In other words, the need for translating concepts does not directly challenge Premise 1 in the self-sufficient view.

Unlike the aforementioned philosophers who demand the template user be knowledgeable beyond the context of their field of specialty, Knuuttila and Morgan (2019) attack the very notion of formal templates. According to them, (2019, 651) relaxing the idealizations that came with ‘formal templates is problematic almost by definition.’ It may be true that tractability makes templates attractive. But to advance knowledge, as they argue and Humphreys agrees, the template needs to be tested against, and adjusted for, relevant empirical data. When incongruences arise, as Humphreys anticipates, one needs to adjust the assumptions, which among other things, include adding back factors that were assumed absent yet may or may not be causally influential. Going through those factors, which Knuuttila and Morgan (2019) call “recomposing,” entails many practical difficulties. For instance, the number of omitted factors could be very large, ‘not able to be fully specified, or dependent in complex ways on one another’ (ibid., 647). Indeed, Humphreys seems to have made the process of adjustment all too effortless for the notion of formal templates to fruitfully describe the adoption of templates in action. That said, Knuuttila and Morgan’s (2019) position does not directly challenge either Premise 1 or 2, nor does it counter Humphreys’ prescriptive point; if the process of adjustment is justified, both sides would agree that justifying it explicitly is not only possible but also preferred.

Finally, unlike all others, the notion of model templates proposed by Knuuttila and Loettgers (2014, 2016) can be interpreted as questioning Premise 1. A model template, in their words (2014, 280), ‘is an abstract conceptual idea associated with particular mathematical forms and computational methods.’ In their view, model templates manifest as ‘a conceptual framework which renders certain kinds of patterns as instances of’ phenomena of a particular type (ibid., 283). For instance, they discuss the mathematical and computational methods developed alongside the Ising model, which was initially built to investigate the phenomenon of ferro-magnetism in physics. A mathematical form resembling the Ising model appears in chemistry, biology and economics for investigating a variety of phenomena. According to them (2014, 295), the Ising model is not a mere syntactic structure. Instead, it ‘consists of such notions as cooperative phenomena, phase transitions, and long-term order embodied into the equations describing the interactions between the components of the system, the energy, and the order parameters.’ Thus, on the one hand, what motivated reapplications of the Ising model is its rich conceptual framework for interdisciplinary modeling practice, what Knuuttila and Loettgers call ‘a general mechanism that is potentially applicable to any subject or field displaying particular patterns of interaction’ (ibid., 295). On the other hand, the mathematical form of the Ising model may not be seen as detached from this conceptual framework. Instead, the Ising model, and potentially other mathematical constructs, is a mathematical structure that should be thought of as ‘*coupled with a general conceptual idea*’ ... ‘an

integrated toolbox' ... '*capable of taking on various kinds of interpretations in view of empirically observed patterns in materially different systems* (2016, 396, original emphasis; 2014, 295).

This notion of model templates, interpreted strongly, questions Premise 1. Recall that Premise 1 is about the independence of template building. Thus, if by “embodied,” “coupled,” and “integrated,” what Knuuttila and Loettgers mean to convey is “cannot be disembodied, decoupled, or disintegrated even in principle,” then there is something more than just the formal template that defines the array of contexts. For this reason, their notion of model template comes closest to rejecting Humphreys’ ahistoricist conjecture. As Humphreys may ask: Should the known first use of the Ising model have not occurred as it did in history, would other scientists in different contexts not be able to build a similar mathematical structure, coupled with the same set of concepts as Knuuttila and Loettgers so aptly analyzed? Or, to put it more generally, recall the metaphysical question regarding knowledge transfer. Does the temporal order of the array of contexts impose some epistemological constraint such that advancing knowledge in a later context would not be possible had at least one prior successful instance not occurred? I will come back to this point in Section 5.

On the whole, the current literature on knowledge transfer has generated many insights for studying the reapplication of a formal representational device in multiple systems. One such insight is to view reapplying a representational device as an analyzable *skill* in terms of template building (Humphreys, 2019), which offers an opportunity for the reapplication of a modeling framework to be studied accordingly. Surely, a model is traditionally understood as a representation of some aspect of a real-world system, and for this reason, a model is importantly different from what may be a modeling framework. However, any account that analyzes a model beyond its written symbols and is open to multiple representational functions of those symbols requires a broader elaborating framework, one to which the model’s symbols belong and one that can be employed to model real-world systems in different domains. In other words, the template building view interprets a mathematical expression as not merely standing in for the one aspect of a real-world system that it is supposed to represent, but rather as a framework for representing materially very different real-world systems. Knuuttila and Loettgers’ (2016) analysis of the Ising model is a salient advocate of this approach. The mathematical structure of the Ising model, which has been applied to study the physical phenomenon of phase transition as well as socio-economic phenomena such as segregation or opinion formation, cannot be accounted for by the traditional representational approach alone. Knuuttila and Loettgers’ (2016, 396) suggest that the Ising model provides a model template; in it, there are both mathematical structures and a general idea ‘of the kind of structure or interaction that the model exhibits.’ I shall come back to the notion of model templates at the end of Section 5. For now, let me simply note that Knuuttila and Loettgers’ (2020) epistemology of knowledge transfer prominently stresses analogical reasoning. For this reason, their view differs significantly from Humphreys’ (2019) account which, as I discussed in Section 2, aims to reduce psychological factors from the epistemology of knowledge transfer. Nonetheless, both approaches pave

the way for applying the analyses of templates to study modeling frameworks in general, a point to which I now turn.

4 The cross-disciplinary reapplications of modeling frameworks

In this section, I examine two contrasting examples of knowledge transfer. In Section 4.1, I first examine features of Chomsky's selective reapplication of mathematical logic in linguistics with the aim of showing that it is possible in practice to reapply a modeling framework as an ideal scientist. The way Chomsky modified his derivation system—which he repurposed from mathematical logic to study a different subject matter in linguistics all the while without introducing context-specific content from mathematical logic—demonstrates this aspect of the self-sufficient view. This episode also shows how the template-based approach applies to the study of modeling framework reapplication. Although Humphreys intended for the notion of templates to be an account of what a computational model *is*, his analysis of *what can be done to a template* to produce scientific knowledge (i.e., the construction and adjustment processes of a template elaborated in Section 2) applies just as well to the study of modeling frameworks.

In Section 4.2, I examine an experimental reapplication of Chomsky's modeling framework in cognitive biology with a specific focus on the presence of a spillover that occurs in the process of knowledge transfer. I shall argue in Section 5 in what sense this spillover defies the ahistoricist conjecture.

Before delving into the particulars, a few words about the spillovers are in order. Because the contrast I attempt to make in this section is between the absence and the presence of a spillover, it is worth giving the notion of spillovers a more precise definition: A spillover is a knowledge-claim that occurs when

- i. a mathematical construct, F , contributes to answering questions or solving problems in multiple contexts, A and B ,
- ii. knowledge-claims K_A and K_B are both epistemic output of applying F in A and B , respectively,
- iii. the discovery of K_A precedes the discovery of K_B , and
- iv. the justification of K_B requires K_A .

In this definition, “context” refers to the combination of a subject matter, a phenomenon of interest, and a set of methods for approaching the phenomenon of interest within a particular subject matter which yields the knowledge-claim in question. Thus, given that K_A precedes K_B , Context A is presumed to precede Context B even though Context B may belong to a discipline that has a longer history than the discipline to which Context A belongs. In other words, the institutional history of a discipline is minimized in this definition of a spillover such that logical dependence in cross-disciplinary justification is foregrounded.

4.1 Building the Chomsky hierarchy without a spillover

The Chomsky hierarchy is a scheme for classifying all formal languages into four families based on the form of the rules in the grammar that generates them. While it is typically visualized with four concentric circles (see e.g., Tecumseh Fitch et al., 2012, 1936), the Chomsky hierarchy can be expressed as a theorem (Moll et al., 1988, 26): Let L be a formal language, and $L(G)$ be the language's grammar.³ Then.

$\mathcal{L}_3 \subsetneq \mathcal{L}_2 \subsetneq \mathcal{L}_1 \subsetneq \mathcal{L}_0$ where $\mathcal{L}_i = \{L \subset X^* \mid L = L(G) \text{ for some Type } i \text{ grammar } G\}$, X is a nonempty finite set of symbols, and X^* is the set of all finite length strings over X .

To show that the building of the Chomsky hierarchy was a result of selectively reapplying mathematical constructs from mathematical logic *without introducing a spillover*, two points need to be made.

The first point to be demonstrated is that there is one mathematical construct used in two different contexts. I elaborate this point in Section 4.1.1 by showing how Chomsky (1956, 1959) constructed what he calls phrase structure grammars (PSGs) by reapplying the derivation system from mathematical logic.

The second point to be demonstrated is that supporting Chomsky's discovery does not require context-specific content from mathematical logic. I discuss this point in Section 4.1.2, starting with the correction set which Chomsky introduced to adjust his PSGs into a hierarchy of four families of formal grammars with differential expressive powers. Moreover, using this framework, Chomsky (1956) argues that the least powerful family of PSGs would not suffice for modeling the syntactic regularities of human language. Such was the subject matter in Chomsky's modeling attempt. It will become clear in the section to come that although the method in mathematical logic and the method in early Chomskyan linguistics can both be said to be formalizations, there remains a significant difference between the theoretical consequence of such derivation.⁴

4.1.1 The formal machinery of a derivation system in the construction of a phrase structure grammar

Regarding the demonstration that there is one mathematical construct used in two different contexts, the mathematical construct in question is the derivation system which logicians build within an axiomatic system to model logical properties. In mathematical logic, the subject matter is valid reasoning, and the phenomenon of interest encompasses logical properties, such as logical truth and logical consequences (Sider, 2010). In Chomskyan linguistics, the subject matter is syntactic regularities in natural languages and the phenomenon of interest includes linguistic properties, such as ambiguity and long distance dependency. Method-wise, both logicians and Chomskyan linguists build an axiomatic system to study their subject

³ In formal language theory, a formal language is defined as a set of strings, whereas a grammar of a formal language is the set of rules that describe the language.

⁴ By Chomskyan linguistics, I refer to his early work in the 1950s.

matter. However, derivations have a substantively different theoretical consequence in linguistics than in mathematical logic. I shall elaborate these in detail in the remainder of this section (see also Table 1 for a summary).

An axiomatic system in mathematical logic, such as propositional logic, typically consists of a syntax, a semantics, and a derivation system. A syntax contains a set of symbols, typically called an alphabet or vocabulary, and a definition for determining the well-formed formulas (WFFs) of the language. A semantics determines the assignments of truth values to the WFFs of the language. Finally, a derivation system specifies the “rules of inference” which allow WFFs to follow from one another in a truth-preserving, step-wise manner. A proper derivation system thus contains rules that determine what can be lawfully produced from either a set of axioms or previously derived theorems. A logician stipulates the rules of inference such that the system, together with appropriately chosen axioms, can derive *all and only* their logical consequences. This comprehensive feature is the signature property of a derivation system, which we will see again in Chomsky’s reapplication of the derivation system in linguistics.

Many readers may be familiar with propositional logic and its derivation system, but an illustration can be helpful here especially because we will be comparing it with Chomsky’s formal system shortly. Consider a formal language \mathcal{L}_P composed of WFFs. The vocabulary of \mathcal{L}_P consists of capital letters, P, Q, R as sentential variables, and the special symbols $\rightarrow, \vee, \wedge, \leftrightarrow, \neg$ are logical constants, where $(\)$ is used for grouping purposes. Given a proper definition, the formulas $(P \rightarrow Q)$ and $(\neg Q \rightarrow \neg P)$ can be WFFs of \mathcal{L}_P . Then, given a proper set of rules of inference, the WFF $(\neg Q \rightarrow \neg P)$ can be derived from (i.e., to appear at the end of a sequence that begins with) the WFF $(P \rightarrow Q)$, and vice versa, each step following one rule of inference. Metalogically speaking, the sequence in which any WFF, ϕ , is shown to “follow” from a set of formulas, Γ , in this rule-following manner is a derivation of the form $\Gamma \Rightarrow \phi$. Moreover, in any such a derivation, ϕ will be a logical consequence of Γ if the rules of inference in \mathcal{L}_P capture good reasoning (cf. Sider, 2010). In the case of mathematical reasoning, for instance, a derivation is the proof of a theorem ϕ from Γ . Note that in this sense, two different proofs of the same theorem are largely theoretically inconsequential, albeit one may be more aesthetically appealing than the other. In linguistics, however, different derivations of the same string can be explanatory.

For an illustration of how Chomsky reapplies the derivation system in linguistics, consider $\mathcal{L}_{English}$ as the set of all sentences in English. Let $\mathcal{L}_{English}(G)$ be the grammar of English. Chomsky shows that one can express the semantic ambiguity in certain English sentences using a PSG. A phrase structure grammar G is expressed as the quadruplet (V_T, V_N, S, P) where the union of V_T and V_N makes up an alphabet, S indicates the start of a derivation, and P denotes a set of rules for said derivation (cf. Moll et al., 1988). A derivation in this context is a sequence of rewrites. It starts from the symbol S and ends at a string, or more generally, $v \Rightarrow w$ where v and w are strings of symbols (Moll et al., 1988, p.2).

On the alphabet, the members of V_T refer to words or morphemes which are part of the final output of a derivation. In contrast, the members of V_N are categorical symbols (e.g., $NP, VP, Verb$) standing-in for the linguistic constituents

Table 1 The derivation system in two contexts

| Contexts | Mathematical Logic | Chomskyan Linguistics |
|--------------------------------|---|---|
| Subject Matters | Valid reasoning | The grammar of natural languages |
| Phenomena of Interest | Logical properties such as logical truth, logical consequence | Properties of natural languages such as ambiguity |
| Distinctive Features in Method | In first-order logic, two different derivations in the form $\Gamma \implies \phi$ with the same Γ and ϕ are theoretically inconsequential. | In phrase structure grammar, two different derivations in the form $v \implies w$ with the same v and w may be explanatory and in turn theoretically consequential. |

(such as noun phrase, verb phrase, verb) in a derivation. Non-terminal elements are thus called because they are designed to never appear at the end of a derivation. The rules of derivation in P , which in linguistics are called the rewrite rules or productions, are expressed in the form of $\alpha \rightarrow \beta$ where both α and β stand for any sequence of elements (terminal or non-terminal; more on this below) and the arrow refers to the direction of rewriting. When all the symbols are terminal elements, the derivation comes to an end. Thus, just as in first-order logic, only lawful, in this case grammatical, strings (i.e., sentences) may appear at the end of a derivation.

Now consider how the sentence “they are flying planes” has two distinct semantic interpretations. Chomsky shows that each interpretation can be represented by a distinctive derivation of a PSG that contains the seven production rules as shown in (1). In the first reading, one could take the third-person plural pronoun, “they,” to be referring to the same objects referred to by the word ‘planes’, as in “they - are - flying planes.” In the second reading, one could understand the pronoun “they” as referring to some pilots who are flying planes, as in “they - are flying - planes.” According to Chomsky, this semantic ambiguity can be expressed in terms of two different derivations.

(1)

1. $S \rightarrow NP VP$
2. $VP \rightarrow Verb NP VP$
3. $Verb \rightarrow are flying$
4. $Verb \rightarrow are$
5. $NP \rightarrow they$
6. $NP \rightarrow planes$
7. $NP \rightarrow flying planes$

In (1), S indicates the start of a derivation. NP is a variable which represents a noun phrase, VP a verb phrase, and $Verb$ a verb—all four, including S , are nonterminal elements. The sign ‘ \rightarrow ’ indicates that the structure on the left-hand side can be rewritten into the structure on the right-hand side. Thus, to derive the first interpretation “they - are - flying planes,” one applies the rules in the order as shown in (2) below:

(2)

1. Rule 1.1 to start from S to obtain NP and VP ,
2. Rule 1.5 to rewrite NP into *they*,
3. Rule 1.2 to rewrite VP into *Verb* and NP ,
4. Rule 1.4 to rewrite *Verb* in *are*, and finally
5. Rule 1.7 to rewrite NP into *flying planes*.

In contrast, to derive the second interpretation “they - are flying – planes,” the sequence would be decidedly different from what is shown in (2). For instance, such a derivation would replace Rule 1.4 with Rule 1.3, and Rule 1.7 with Rule 1.6. The two interpretations of the same string of words are thus represented by two different derivations. Moreover, both derivations, despite their different processes, terminate at the same string “they are flying planes.” With this expressive power, the PSG is shown to be capable of accounting for some semantic ambiguity in English. Thus, due to this explanatory power, Chomsky argues that $\mathcal{L}_{English}(G)$ must contain the production rules like those shown in (1).

The above illustration shows how Chomsky repurposes the derivation system to investigate the grammar of a natural language, which to him consists of the production rules that can generate all syntactically acceptable strings in a natural language. Chomsky modified the vocabulary of PSGs to express linguistics properties. His modifications suggest a clear shift in subject matter, as the subject-specific ingredients from the context of mathematical logic (such as notions of logical truth, consequence, or constants) do not appear in the new context of linguistics. However, despite these changes, the rule-following behavior of the derivation system in mathematical logic is replicated, and in doing so, Chomsky did not introduce a knowledge-claim from mathematical logic to justify his discovery.

As mentioned at the beginning of this section, Chomsky refined his PSG modeling framework to build what eventually became the Chomsky hierarchy. This section focused on the construction process; it showed that Chomsky replicated the formal machinery of the derivation system from mathematical logic, using what was meant to demonstrate logical consequences to instead reveal syntactic regularities in natural languages. Indeed, he was explicit about having borrowed the idea of a proof from mathematical logic. Chomsky (1956, 116) wrote, ‘[a] derivation is thus roughly analogous to a proof,’ with the set of initial strings ‘taken as the axiom system and’ the rewrite rules ‘as the rules of inference.’

In the next section, we will see a “correction set” (Humphreys, 2004) being introduced to adjust a PSG. These limiting conditions were devised to refine a given PSG for narrowing down the research space; the outcome of that adjustment became the basis of the Chomsky hierarchy. Both processes jointly complete my demonstration of how Humphreys’ template-based approach applies to studying the reapplication of a modeling framework.

4.1.2 Adding constraints as a correction set to adjust phrase structure grammars

Chomsky introduced three increasingly limiting conditions to the form of the rules as in set P follows.⁵ First limiting condition: For every production $\alpha \rightarrow \beta$ in P , $|\alpha| \leq |\beta|$, i.e., the number of symbols on the left-hand side of the production is smaller than or equal to the number of the symbols on the right-hand side of the production. That is, the resulting string β will always be longer or equal to the length of the string α before the rewrite. Thus, applying the productions of this PSG would not result in a decrease of the string length. Second limiting condition: For every production $\alpha \rightarrow \beta$ in P , (1) α consists of only one non-terminal symbol and (2) β cannot be an empty string, e.g., $S \rightarrow aSb$, $S \rightarrow ab$. Third limiting condition: For every production $\alpha \rightarrow \beta$ in P , (1) $|\alpha| = 1$, $\alpha \in V_N$, and (2) β has the form Γ or $\Gamma\Psi$, where $\Gamma \in V_T$, $\Psi \in V_N$, and $|\Gamma| = |\Psi| = 1$, e.g., $S \rightarrow a$, $S \rightarrow aS$.

Some words need to be said about the extent to which these limiting conditions act as a correction set. While those limiting conditions may seem arbitrary, the restrictions effectively allow for differential expressive powers between PSGs, which is what Chomsky needed to model English. In particular, designed in this way, the expressive power of the least restrictive PSGs will be equivalent to that of the universal Turing machine (Turing, 1936), whereas the most restrictive PSGs will be as powerful as finite state machines (Shannon, 1948 as Chomsky cites in 1956 but referring to it as the finite Markov procedure). To get a bit ahead of the story, according to Chomsky, the former is unnecessarily powerful for modeling English, whereas the latter lacks sufficient power. On the one hand, a Turing machine (Turing, 1936) consists of a finite list of instructions and a memory store that is infinite in length. The instructions specify a finite number of “states” and the conditions by which the machine transitions from one state to another. A universal Turing machine is one that can compute any function that can be computed by any Turing machine (see Shagrir, 2016 for a philosophical account of the Turing machines). Turing (1936) constructed this type of abstract computer to explore the limitations of computation. And, because *describing* a formal language entails computing a characteristic function (i.e., by giving a “yes” or “no” output based on whether any input string belongs to it), a universal Turing machine is defined as capable of describing all computable formal languages. On the other hand, at the time of Chomsky’s work, there was another kind of abstract machine called finite state machines. Like Turing machines, a finite state machine consists of a finite list of instructions that specifies a finite number of “states” and the conditions for transitioning from one state to another. However, unlike Turing machines, finite state machines do not have an explicit memory store. It has been proven that finite state machines describe only a subset of computable formal languages, which were called regular languages at the time (Kleene, 1951, 1956).

The Chomsky hierarchy came about by adjusting PSGs into a refined space of expressive powers so that Chomsky could place any natural language, e.g., English, in the right family of the hierarchy. Recall the Chomsky hierarchy expressed in a theorem:

⁵ This review is based on Levelt (2008, 10); c.f., Partee, Meulen, and Wall (1990).

$\mathcal{L}_3 \subsetneq \mathcal{L}_2 \subsetneq \mathcal{L}_1 \subsetneq \mathcal{L}_0$ where $\mathcal{L}_i = \{L \subset X^* \mid L = L(G) \text{ for some Type } i \text{ grammar } G\}$, X is a nonempty finite set of symbols, and X^* is the set of all finite length strings over X .

When a PSG is free from any restrictions (i.e., no limiting condition applies to its rules), it is called a Type 0 non-restricted grammar. Type 0 PSGs are designed to describe \mathcal{L}_0 , which includes all Turing-computable languages because they are meant to be as expressive as a universal Turing machine.⁶ In contrast, when a PSG meets all three limiting conditions, it is called a Type 3 regular grammar. Type 3 regular grammars are designed to be expressive enough to describe regular languages (Chomsky, 1959; Chomsky & Miller, 1958), i.e., the \mathcal{L}_3 in the hierarchy. Between the outermost and the innermost classes are two new grammar families. When a PSG meets the first limiting condition, Chomsky calls it a Type 1 context-sensitive grammar, whereas when a PSG meets both the first and the second limiting conditions, it is called a Type 2 context-free grammar. These Types describe the formal languages \mathcal{L}_1 and \mathcal{L}_2 in the hierarchy which are called context-sensitive languages and context-free languages, respectively. We will encounter Type 2 and Type 3 grammars in Section 4.2.

The need to refine the space of expressive power was due to the fact that, according to Chomsky (1956), finite state machines lack the necessary power to describe English. This aspect of Chomsky's argument warrants closer analysis because in defending his finding with the hierarchy, Chomsky again did not introduce a spillover.

Long distance dependency and open-endedness in English were especially problematic for finite state machines; for any finite state machine to describe English, it needs to accommodate these two features. For instance, consider the sentence pattern (3) as follows:

- (3)
- a. If S_1 , then S_2 .
 - b. Either S_3 , or S_4 .

Long distance dependency refers to the linguistic phenomenon in which if one replaces "if" in 3.a by "either," it requires a corresponding replacement of "then" by "or," as in 3.b, lest the resulting string not be an English sentence and vice versa. Moreover, as the argument goes, there is no upper limit to the lengths of S_1 and S_3 which respectively refer to a sentence clause. Thus, any abstract computer candidate for describing English needs to be able to handle indefinite lengths of symbols between "if" and "then" or "either" and "or." To handle both features, a finite state machine needs either a flexible number of states or an explicit memory store. However, by definition, a finite state machine has a fixed number of states and lacks an explicit memory store. Thus, Chomsky concludes, the features of long distance dependency and open-endedness in English will result in sentences that are admissible by native English speakers but indescribable by finite state machines; no finite state machine could describe English. As Type 3 regular grammars are designed to be as powerful as finite state machines, to describe

⁶ Unlike in computer science, in linguistics the process time required for producing such an answer was not an issue.

English, one needs a grammar with more expressive power than a Type 3 regular grammar.

I now turn to four takeaways lending support to the claim that it is *possible in practice* to carry out knowledge transfer as an ideal scientist, the second and fourth of which demonstrate how Humphreys' template-based analysis applies to modeling frameworks. First, what Chomsky took from mathematical logic is the rule-following formal feature of the derivation system; all the ingredients that went into constructing and adjusting PSGs are from within the context of linguistics. Second, the signature feature of a derivation system is preserved through the constructing process. Given a set of initial strings and a set of rules, the system derives *all and only* lawful strings. Third, however, the derivations differ in their representational function. In mathematical logic, a derivation represents a valid sequence of reasoning, where in linguistics, it represents the generation of an admissible sentence in a natural language. Moreover, while any step within a logical derivation yields a WFF, a linguistic derivation only arrives at an admissible string at the end of the derivation. This is due to how Chomsky built PSGs with linguistic entities such as verb phrases, noun phrases, etc., an ontology that is local to the context of linguistics. Fourth, and finally, Chomsky tested PSGs against the linguistic properties of English. This effort to refine PSGs was achieved by imposing increasingly strict conditions to the rules in a given PSG. At the end of the process of adjustment came the basis of the Chomsky hierarchy: a nesting hierarchy of four families of formal grammars, together with the four families of formal languages they describe, and the abstract machines that have matching expressive power to Type 0 and Type 3 grammars, respectively.

4.2 Reapplying the Chomsky hierarchy

The Chomsky hierarchy that refined the space between the universal Turing machines and finite-state machines made room for other abstract machines to be introduced. The context-free grammars were paired with a then new-found abstract machine family called pushdown automata. These new developments, which did not concern natural languages (Greibach, 1981), became key to the development of early theoretical computer science, making the Chomsky hierarchy a successful template. In Section 4.2.1, we will see the success of the Chomsky hierarchy in computer science giving rise to a knowledge-claim and, in Section 4.2.2, how that knowledge-claim contributes to experimental scientists' reapplication of the Chomsky hierarchy.

4.2.1 The success of the Chomsky hierarchy in computer science

During the formative years of what eventually became "computer science," translation between formal languages was a constant problem. The challenges involved translating between the machine language that is comprised of binary digits of ones and zeros and other programming languages. Turing machines were powerful, but in practice, both

processing time and memory store were scarce resources. It was crucial to find the most efficient algorithm to solve the translation problem. The Chomsky hierarchy's success in meeting these challenges established its theoretical status in computer science.

For instance, before a hardware computer executes a code written in a programming language, say, C++, two preliminary tasks must be completed. First the computer must check whether that code belongs to C++, i.e., whether every string in the input is a sentence. If the answer is "yes," the computer must then translate the code from the language of C++ to the machine language of the CPU. A compiler is a program that carries out these two preliminary tasks. A software engineer needs to make sure that the compiler will always *reject* code that violates the rules of the programming language (i.e., the syntactically incorrect programs) while always accepting the syntactically correct ones. This part of a compiler is called a "parser" (e.g., see Parkes, 2002).

Key breakthroughs in early computer science were due to Chomsky's notion of Type 2 context-free grammars. First, a new programming language ALGOL (Backus et al., 1960) was demonstrated to generate only Type 2 context-free languages (Ginsburg & Rice, 1962, see Hyman, 2010 for a historical account). As Chomsky's (1956) work in linguistics suggests, finite state machines cannot describe Type 2 context-free languages; they cannot parse ALGOL code. Meanwhile, a new type of state machine at the time called a pushdown automaton was introduced. Unlike the Turing machines, which have unlimited memory store, and unlike finite state machines, which have no explicit memory store, a pushdown automaton comes with a linear memory store that, figuratively speaking, pushes items down the stack and pops the latest one out before an earlier item may be retrieved. Second, the expressive power of Type 2 context free grammars was found to be equivalent with that of the pushdown automata (Chomsky & Schützenberger, 1963; Evey, 1963; cited in Greibach, 1981). Thus, combined with the finding that ALGOL generates only Type 2 context free languages, to program a compiler that can parse ALGOL code, one knew to implement at least a pushdown stack in the compiler.

The success with ALGOL rendered the Chomsky hierarchy an important object of knowledge for compiler designers. The formal machinery of the hierarchy, despite being built to study natural languages, applies to dealing with programming languages; programmers are better off knowing that to parse a Type 2 language, they need to implement at least a pushdown stack in the compiler. The repeated success of reapplying the Chomsky hierarchy in computer science gave rise to a knowledge-claim: parsing input of a Type n language requires an implementation of at least a Type n automaton. Next, we shall see this claim "spilling over" into the context of cognitive biology where the parsing powers of biological organisms are subject to experimental detection.

4.2.2 Using the Chomsky hierarchy experimentally in cognitive biology

Cognitive biologists study cognition as a collection of biological functions. One strand of cognitive biology aims to understand the evolution of and the neural substrate for human linguistic capacity. To study this subject matter, several

prerequisites need to be met. For instance, to understand the neural basis of human linguistic capacity, biologists need to first isolate those bases from the neural bases of cognitive capacities common between human and nonhuman animals (Bowling, 2014). Thus, approaching the question about species differences scientifically (Fitch, 2014) requires an experimental procedure and stimuli that are applicable across species such that testing results may be meaningfully compared.

Fitch and Hauser's (2004) solution to the species difference problem is to use context-free and regular grammars to differentiate the parsing powers of biological organisms. Fitch and Hauser (2004) argue that humans are endowed with a computational capacity beyond the class of finite state machines, and this capacity, which is at least equivalent to the pushdown stack in Type 2 automata, is not available to their cotton-top tamarin monkey (*Saguinus Oedipus*) testing subjects.

Germane to the focus of the present paper is whether advancing knowledge using the Chomsky hierarchy in cognitive biology requires a spillover. I argue that it does. For instance, consider the reconstruction of Fitch and Hauser's argument (2004).

(4)

1. If an organism recognizes a context-free language, it possesses computational resources equivalent to (at least) a pushdown automaton; if it recognizes a regular language but not a context-free language, it possesses computational resources equivalent to a finite-state automaton.
2. Tamarin monkey subjects recognize a regular language but not a (crucially similar) context-free language, whereas human subjects recognize both of these languages.
3. Therefore, tamarin monkeys lack the computational resources equivalent to a pushdown store that humans have.

The soundness of the argument in (4) depends on the truth of the empirical claim in (4.2), as well as the acceptance of premise (4.1). Let's consider these two in turn.

The truth of (4.2) must be adjudicated by analyzing the procedure in Fitch and Hauser's experiment, where they modified a standard experimental protocol called artificial grammar learning (AGL). Both the context-free language and the regular language in AGL were replaced with "local ingredients." Thus, no inter-context knowledge is required for accepting the claim in (4.2). Psychologists have been reporting results of infants as young as 7 months old using AGL (Saffran et al., 1996), but Fitch and Hauser (2004) were the first to bring the Chomsky hierarchy into this line of work (see Levelt, 2019 for discussion of the methodology).

In contrast, accepting (4.1) entails three additional assumptions, one of which comes from computer science. First, the ontological assumption that the cognitive infrastructure in different biological organisms are equivalent to the parsing powers of pre-programmed automaton; second, the methodological assumption that the parsing power of pre-programmed automata are determined based on the type of input they can recognize; and finally, third, that parsing input of a Type n language requires an implementation of

at least a Type n automaton—i.e. a *spillover*. Jointly, all three assumptions made AGL a valid empirical detector for differentiating the parsing powers of biological organisms. The knowledge-claim from computer science is an indispensable part of its justification.

5 Spillovers and truth-functional dependency in knowledge transfer

The presence of spillovers in knowledge transfer challenges the self-sufficient view as it indicates that its ahistoricist conjecture fails to generalize. In this section, I elaborate this point further, address potential objections to the example in cognitive biology, and then discuss the contribution of the notion of spillovers to the knowledge transfer literature.

Regarding the self-sufficient view's ahistoricist conjecture, recall the array of different contexts that share neither subject matter nor the entirety of their methods, and not even the phenomenon of interest. See Table 2 for a partial array of such a sequence of contexts featuring the applications the Chomsky hierarchy as discussed in Section 4. If the ahistoricist conjecture is categorically true, then knowledge production in each context is logically independent from one another. Yet, the existence of a spillover as we saw in Section 4 indicates a certain kind of logical dependency between these contexts. Because this dependency manifests in the justification in a subsequent context, the nature of such dependency can be said to be truth-functional. In this sense, the Chomsky hierarchy's reapplication in cognitive biology can be said to be truth-functionally dependent on the epistemic output from its earlier reapplication in computer science. Hence, if my analysis is sound, the self-sufficient view as an epistemology of knowledge transfer warrants reconsideration.

Those seeking to defend the ahistoricist conjecture may propose to integrate the contexts of cognitive biology and computer science into one general entity, say a domain of intelligent systems be they artificial or biological. This way, the reapplications of the Chomsky hierarchy may be analyzed as different specifications of a single theoretical template. It follows that what I analyze to be a counterexample to the ahistoricist conjecture was but the idiosyncrasy of a special science. A full treatment to this objection will take us far afield. However, it is likely that this approach shows at best that integration between contexts can be subject to ahistorical analysis. Much like Humphreys' template-based analysis, an ahistorical analysis cannot prove nor disapprove the ahistoricist conjecture.

One may try to defend the self-sufficient view by arguing that the spillover in the enhanced AGL example can be derived from general mathematical competence. This approach amounts to removing the predicates in "pre-programmed" automata and treating biological organisms simply as mathematically defined automata. There may be a branch of pure mathematics to be discovered that would allow for pairing Turing machines with differential parsing powers. But in essence, this strategy assumes that theoretical knowledge of an applied mathematics, such as computer science, can be reduced to pure mathematics. While a sharp boundary between applied mathematics and pure mathematics may be hard to come by, it is implausible that the differential "parsing powers" in biological organisms can be mapped onto mathematically defined automata—i.e., Humphreys' notion of formal

Table 2 The Chomsky hierarchy in three contexts

| | Chomskyan Linguistics | Computer Science | Cognitive Biology |
|--------------------------------|---|--|--|
| Subject Matters | The grammar of natural languages | Efficient algorithms to translate between formal languages | The difference in cognitive infrastructures between species |
| Phenomena of interest | Properties of natural languages such as long distance dependency | The memory store required for parsing programming languages | Responses to artificial grammars among different species |
| Distinctive Features in Method | The Chomsky hierarchy is used to rule out models and modeling frameworks with inadequate expressive power to describe a natural language. | The Chomsky hierarchy is used to determine which type of memory store to implement for a given task. | The Chomsky hierarchy is used to determine which type of parsing power is pre-programmed in different species. |

templates—without predicates such as being “previously engineered” or “evolutionarily programmed.” Alternatively, taking the route of building a theoretical template from scratch instead, a well-trained biologist with a perfect command of mathematics could construct a system that differentiates parsing powers in biological organisms. Yet, for the same reason that context-free grammars (see Section 4.2.1) might not have come to dominate computer science if it was not for historical contingency (Ginsburg, 1980), the hypothetical AGL protocol would unlikely be Chomskyan without Chomsky’s work in precedent.

More generally, one may interpret the ahistoricist conjecture as well as the self-sufficient view as *prescribing* independency rather than *describing* scientific knowledge production in practice. Under this interpretation, only when a given transfer of knowledge is “spillover free” does one gain genuine scientific knowledge. However, it is not immediately clear why genuine scientific knowledge needs to be free of spillover in its production. In contrast, as I argued in Section 2, Humphreys’ epistemology of knowledge transfer is more appropriately understood in terms of a description of an ideal scientist in practice. Under this interpretation, the self-sufficient view advocates replacing tacit knowledge in reapplying a mathematical construct (i.e. perceived similarities, analogy, etc.) with a template-building analysis and is thus a position consistent with the approach undertaken in this paper. What I beg to differ from Humphreys’ view is that, for him, the template-building approach could eliminate the *need* for cross-disciplinary knowledge, whereas for me this remains an open question. Surely, the presence of a spillover in a reapplication does not guarantee truth-preservation because it entails at least one assumption that is not “locally” justifiable. Nonetheless, failing to acknowledge spillovers makes the self-sufficient view unnecessarily strict. Should practicing scientists intentionally steering away from spillovers, certain research questions may otherwise be pending indefinitely for a tangible solution.

Finally, the notion of spillovers contributes to the knowledge transfer literature by allowing for a logical dependency between reapplications of a single mathematical construct to be articulated and detected in a way that can address the metaphysical

question regarding knowledge transfer. As discussed in Section 3, many scholars of knowledge transfer contend that cross-disciplinary or inter-field knowledge of some sort is at least advisable, if not necessary, to successfully reapplying an object of knowledge across disciplines. Yet among these scholars, only Knuuttila and Loettgers' (2014, 2016, 2020) notion of model templates comes closest to challenging the ahistoricist conjecture. Consequently, the existence of a spillover and that of a model template can be said to entail a dependency between applications of a single mathematical construct. However, these two notions differ significantly in at least three important ways. First, being a knowledge-claim, a spillover is an assertion, which makes it qualitatively different from being a conceptual framework which does not always, if at all, have a truth value. Second, in reapplying a knowledge object, the role of a spillover is to provide cross-context justification, whereas the role of a model template is to offer 'a general conceptual core' that motivates the reapplication (Knuuttila & Loettgers, 2020, 137). Third, and relatedly, the notion of spillovers captures truth-functional dependency in knowledge transfer, whereas the notion of model templates emphasizes a conceptual dependency between applications of a mathematical formula. It thus remains to be seen whether conceptual dependencies comprise of logical dependencies.

6 Conclusion

This paper has demonstrated that Humphreys' template-based analysis can be productively applied to study reapplications of modeling frameworks even though not all insights from his epistemology of knowledge transfer follow. I have articulated Humphreys' (2019) position on knowledge transfer as the self-sufficient view, which wagers that advancing knowledge with a repurposed mathematical construct can be conceived of as independent template-building endeavors. According to my analysis, embedded in Humphreys' (2019) response to the epistemological debates surrounding knowledge transfer is a conjecture that historical contingencies are irrelevant to the nature of knowledge production. Calling this ahistoricist conjecture into question, I developed the notion of spillovers and showed that a reapplication of a mathematical construct can be determined as either truth-functionally or non-truth-functionally dependent on a prior application of the same mathematical construct.

The conclusion is not that the self-sufficient view is false but that it is unnecessarily strict in directing practicing scientists to steer away from spillovers. Using the development of the Chomsky hierarchy and its cross-disciplinary reapplications as an illustration, I argued that it is possible in practice to reapply a cross-disciplinarily sourced mathematical construct in the way Humphreys advocates, i.e., without introducing cross-disciplinary knowledge to the template-building process. However, what the self-sufficient view fails to consider are the problems and questions in science whose solution require scientists to not only repurpose a mathematical construct but also to incorporate the construct's prior epistemic output into their reapplication. My analysis did not show whether problems or questions of this sort are prevalent in science; nevertheless,

when they do exist, following Humphreys' suggestion their solution or answers may never ascertain.

My analysis also showed that in addition to this present paper, Knuuttila and Loettgers' (2014, 2016, 2020) work on model templates have the potential to challenge the ahistoricist conjecture in Humphreys' position. I argued that while model templates may entail a conceptual dependency in knowledge transfer, future work is needed to determine whether such dependency may address the metaphysical question regarding knowledge transfer. Nonetheless, owing to their differences, the notions of spillovers and model templates may function complementarily to investigate the ways different scientific domains are related with one another to advance scientific knowledge. For instance, despite not being the kind of mathematical formula typically discussed in the literature on model templates, the Chomsky hierarchy does seem to offer an integrated set of both conceptual and mathematical tools for modeling a variety of materially different systems.⁷ The notion of model templates identifies the conceptual dependency in these reapplications, thereby revealing the extent to which analogical reasoning is prevalent in scientific modeling. In the meantime, the notion of spillovers, through logical reconstruction, picks out the truth-functional dependency within a subset of these reapplications, thus highlighting potential intertheoretical connections made through knowledge transfer. Moreover, because spillovers are by definition not native to the importing discipline, their presence may seem unintuitive, or even problematic, to some practicing scientists. For this reason, how a scientific community responds to spillovers (and the knowledge-claims they help enable) would be a complementary research topic that intersects with the philosophical investigation of knowledge transfer.

The presence of spillovers has an effect on the production of knowledge; one place to look for such effects are problems in science whose solution require a cross-disciplinarily sourced object of knowledge. The acceptance of spillovers sheds light on the diffusion of knowledge. Jointly, not only do historical contingencies have a genuine impact on template-based knowledge transfer, they also shape the image of scientific knowledge production as a whole. The notion of spillover thus contributes toward a unified conceptual framework for both practice-oriented philosophers of science and historiographers of mathematical models.

Acknowledgements The author would like to thank Paul Humphreys, Tarja Knuuttila, Mary Morgan, Colin Allen, Tecumseh Fitch, and Michael Dickson for their helpful discussions and/or comments on earlier drafts; thanks also the participants of several meetings for their feedback, including the workshop on "Transdisciplinary Model Transfer and its Interfaces," at the University of Vienna, the graduate seminar on "Scientific Ontology and the Epistemology of Science," at the University of Virginia, and NeuroTech: An Interdisciplinary Early Career Workshop on Tools and Technology

⁷ For example, context-free grammars and the linguistic derivation system have been used in molecular biology to investigate biological sequences comprised of nucleotide bases (see Searls, 2002 for a review). A similar approach to biological sequences is seen in synthetic biology where context-free grammars are used to help identify and engineer DNA sequences with certain desired biological functions (e.g., Czar et al., 2009).

in Neuroscience at the Center for Philosophy of Science, University of Pittsburgh. A special thanks goes to the two anonymous reviewers whose insightful comments and constructive criticisms helped improve and clarify this manuscript. Finally, this material is based upon work supported by the Konrad Lorenz Institute for Evolution and Cognition Research (Austria) and the National Science Foundation (United States) under grant no. 1922143. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of either funding agency.

Funding This study was funded by National Science Foundation (grant number 1922143).

Declarations

Ethical approval This paper does not contain any studies with human participants or animals performed by the author.

Informed consent N/A

Conflict of interest The author states that there is no conflict of interest.

References

- Backus, J. W., Bauer, F. L., Green, J., Katz, C., McCarthy, J., Naur, P., Perlis, A. J., Rutishauser, H., Samelson, K., & Vauquois, B. (1960). Report on the algorithmic language ALGOL 60. *Numerische Mathematik*, 2(1), 106–136.
- Bowling, D. (2014). Cognitive Theory and Brain Fact: Insights for the Future of Cognitive Neuroscience. Comment on ‘Toward a Computational Framework for Cognitive Biology: Unifying Approaches from Cognitive Neuroscience and Comparative Cognition’ by W. Tecumseh Fitch. *Physics of Life Reviews*. Elsevier. <https://doi.org/10.1016/j.plrev.2014.07.007>
- Bradley, S., & Thébault, K. P. Y. (2019). Models on the move: Migration and imperialism. *Studies in History and Philosophy of Science Part A*, 77, 81–92. <https://doi.org/10.1016/j.shpsa.2017.11.008>
- Burgess, J.P. (1992). Proofs about Proofs: A Defense of Classical Logic: Part I: The Aims of Classical Logic. In *Proof, Logic and Formalization*, edited by Michael Detlefsen, 8–23. Routledge. <https://doi.org/10.4324/9780203980255>
- Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, 2. <https://doi.org/10.1109/TIT.1956.1056813>
- Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control*, 2, 137–167. <https://doi.org/10.1075/bjil.1.08mil>
- Chomsky, N., & Miller, G. A. (1958). Finite state languages. *Information and Control*, 1(2), 91–112. [https://doi.org/10.1016/S0019-9958\(58\)90082-2](https://doi.org/10.1016/S0019-9958(58)90082-2)
- Chomsky, N., & Schützenberger, M. P. (1963). The algebraic theory of context-free languages. *Studies in Logic and the Foundations of Mathematics*, 35, 118–161. [https://doi.org/10.1016/S0049-237X\(08\)72023-8](https://doi.org/10.1016/S0049-237X(08)72023-8)
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(April 1924), 623–656. <https://doi.org/10.2307/3611062>
- Czar, M.J, Cai, Y., and Peccoud, J. (2009). Writing DNA with GenoCAD™. *Nucleic Acids Research* 37 (suppl_2): W40–47.
- Evey, R. J. (1963). *The Theory and Application of Pushdown Machines*. Cambridge, MA: Rep. No. NSF-10, Harvard Comput. Laboratory.
- Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews*, 11(3), 329–364. <https://doi.org/10.1016/j.plrev.2014.04.005>
- Fitch, W. T., & Hauser, M.D. (2004). Computational Constraints on Syntactic Processing in a Nonhuman Primate. *Science (New York, N.Y.)* 303 (January): 377–80. <https://doi.org/10.1126/science.1089401>

- Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1933–1955. <https://doi.org/10.1098/rstb.2012.0103>
- Ginsburg, S., and Rice, H. G. (1962). Two Families of Languages Related to ALGOL. *Journal of the ACM (JACM)*, 350–71. <https://doi.org/10.1145/321127.321132>
- Ginsburg, S. (1980). Methods for specifying families of formal languages - past-present-future. In *formal language theory*, 1–22. Academic Press. <https://doi.org/10.1016/B978-0-12-115350-2.50006-3>
- Greibach, S. A. (1981). Formal languages: Origins and directions. *Annals of the History of Computing*, 3(1).
- Grüne-Yanoff, T. (2011). Models as products of interdisciplinary exchange: Evidence from evolutionary game theory. *Studies in History and Philosophy of Science Part A*, 42(2), 386–397. <https://doi.org/10.1016/j.shpsa.2010.12.004>
- Herfeld, C., & Doehne, M. (2019). The diffusion of scientific innovations: A role typology. *Studies in History and Philosophy of Science Part A*, 77, 64–80. <https://doi.org/10.1016/j.shpsa.2017.12.001>
- Herfeld, C., & Liscandra, C. (2019). Knowledge transfer and its contexts. *Studies in History and Philosophy of Science Part A*, 77, 1–10. <https://doi.org/10.1016/j.shpsa.2019.06.002>
- Hesse, M. (1964). Analogy and confirmation theory. *Philosophy of Science*, 31(4), 319–327.
- Hesse, M. B. (1966). *Models and analogies in science*. University of Notre Dame Press.
- Houkes, W., & Zwart, S. D. (2019). Transfer and templates in scientific modelling. *Studies in History and Philosophy of Science Part A*, 77, 93–100. <https://doi.org/10.1016/j.shpsa.2017.11.003>
- Humphreys, P. (2002). Computational Models. *Philosophy of Science*, 69(September), 1–27. <https://doi.org/10.1093/oxfordhb/9780199675111.013.026>
- Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*. Oxford University Press.
- Humphreys, P. (2019). Knowledge transfer across scientific disciplines. *Studies in History and Philosophy of Science Part A*, 77(February), 112–119. <https://doi.org/10.1016/J.SHPSA.2017.11.001>
- Hyman, M. D. (2010). Chomsky between revolutions. *Chomskyan (R)Evolution*, 265–98. <https://doi.org/10.1075/z.154.09hym>
- Kleene, S. C. (1951). *Representation of events in nerve nets and finite automata*. U.S. Air Force Project RAND, Research Memorandum.
- Kleene, S. C. (1956). Representation of events in nerve nets and finite automata. *Automata Studies*, 34. <https://doi.org/10.1515/9781400882618-002>
- Knuutila, T., & Loettgers, A. (2020). Magnetized memories: Analogies and templates in model transfer. In S. Holm & M. Serban (Eds.), *Living machines? Philosophical perspectives on the engineering approach in biology* (pp. 123–40). Routledge.
- Knuutila, T., & Loettgers, A. (2014). Magnets, spins, and neurons: The dissemination of model templates across disciplines. *The Monist*, 97(3), 280–300.
- Knuutila, T., & Loettgers, A. (2016). Model templates within and between disciplines: From magnets to gases—and socio-economic systems. *European Journal for Philosophy of Science*, 6(3), 377–400.
- Knuutila, T., & Morgan, M. S. (2019). Deidealization: No easy reversals. *Philosophy of Science*, 86, 641–661.
- Kuhn, T. S. (1970). *The structure of scientific revolution* (2nd ed.). University of Chicago Press.
- Kuhn, T. S. (1974). Second thoughts on paradigms. *The Structure of Scientific Theories*, 2, 459–482.
- Levelt, W. J. M. (2008). *An introduction to the theory of formal languages and automata*. John Benjamins Publishing.
- Levelt, W. J. M. (2019). On empirical methodology, constraints, and hierarchy in artificial grammar learning. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12441>
- Moll, R. N., Arbib, M.A., & Kfoury, A. J. (1988). *An introduction to formal language theory*.
- Morgan, M. S. (2014). Resituating knowledge: Generic strategies and case studies. *Philosophy of Science*, 81(5), 1012–1024.
- Parkes, A. (2002). Introduction to languages, Machines and Logic : Computable Languages, Abstract Machines and Formal Logic.
- Partee, B.H., Meulen, A. G., and Wall, R. (1990). *Mathematical methods in linguistics. Studies in Linguistics and Philosophy*. Kluwer Academic Publishers Group.
- Price, J. (2019). The landing zone – Ground for model transfer in chemistry. *Studies in History and Philosophy of Science Part A*, 77, 21–28. <https://doi.org/10.1016/j.shpsa.2018.06.010>
- Saffran, J. R., Aslin, R. N., Newport, E. L., Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.
- Searls, D. B. (2002). The language of genes. *Nature*, 420, 211. <https://doi.org/10.1038/nature01255>.

- Shagrir, O. (2016). Advertisement for the Philosophy of the Computational Sciences. Edited by Paul Humphreys. *The Oxford Handbook of Philosophy of Science*, 15. <https://doi.org/10.1093/oxfordhb/9780199368815.013.3>
- Sider, T. (2010). *Logic for philosophy*. Oxford University Press.
- Turing, A. M. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. In *Proceedings of the London Mathematical Society*, 2, 230–65. <https://doi.org/10.2307/2268810>
- Zuchowski, L. (2019). Modelling and knowledge transfer in complexity science. *Studies in History and Philosophy of Science Part A*, 77, 120–129. <https://doi.org/10.1016/j.shpsa.2017.10.003>

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