Axiomathes (2022) 32:1025–1045
https://doi.org/10.1007/s10516-021-09563-6

ORIGINAL PAPER

Conceptual Learning and Local Incommensurability: A Dynamic Logic Approach

Corina Strößner[1]

Received: 2 December 2020 / Accepted: 12 May 2021 / Published online: 28 June 2021

Abstract
In recent decades, the logical study of rational belief dynamics has played an increasingly important role in philosophy. However, the dynamics of concepts such as conceptual learning received comparatively little attention within this debate. This is problematic insofar as the occurrence of conceptual change (especially in the sciences) has been an influential argument against a merely logical analysis of beliefs. Especially Kuhn's ideas about the incommensurability, i.e., untranslatability, of succeeding theories seem to stand in the way of logical reconstruction. This paper investigates conceptual change as model-changing operations similar to belief revision and relates it to the notion of incommensurability. I consider several versions of conceptual change and discuss their influences on the expressive power, translatability and the potential arising of incommensurability. The paper concludes with a discussion of animal taxonomy in Aristotle's and Linnaeus's work.

Keywords Conceptual change · Conceptual learning · Dynamic logic · Kuhn · Incommensurability

1 Introduction

Our beliefs are subject to constant change through communication, observation and examination. This concerns our everyday beliefs but also and especially scientific theories. In the recent decades, philosophers extensively researched the change of knowledge and beliefs in view of new evidence. This field of research, *formal epistemology*, uses different frameworks such as Bayesianism (Talbott 2016), belief revision theory (Alchourron et al. 1985; Hansson 2017), ranking theory (Spohn 2012), Dempster-Shafer theory (Dempster 1967; Shafer 1976; Tsiporkova et al. 1999) or dynamic epistemic logic (Baltag and Renne 2016; van Benthem 2011; van Ditmarsch et al. 2007). They provide insights into the way rational agents (or groups of them) should change

✉ Corina Strößner
corina.stroessner@ruhr-uni-bochum.de

[1] Institute for Philosophy II, Ruhr University Bochum, 44780 Bochum, Germany

their beliefs in view of empirical evidence or truthful communication. However, in most frameworks of formal epistemology, the language and concepts are assumed to be fixed. This is a useful simplification that helps to focus on merely epistemic changes. However, there are philosophical worries about such an approach, here named after their most popular representatives:

- Quine's worry: If one models how to revise and update *beliefs*, one implicitly assumes that our language is fixed and that there are *eternal* conceptual truths (analytic laws), which is an unjustified assumption.
- Kuhn's worry: Conceptual changes play an important role in our cognitive history (e.g., in scientific as well as in individual development). Belief systems before and after a severe epistemic change may lack common language. A formal epistemology that studies mere belief updates or revisions *within* a language, is mostly not applicable.

These two assumptions played a major role in the rejection of logical empiricism in the philosophy of the 20th century. Prima facie, similar accusations can be made against various theories of formal epistemology, as long as they do not address conceptual change. These worries concern all kinds of conceptual development, including the learning and evolution of natural languages. First and foremost, however, they are concerned with scientific change. In a naive view, a theory can be viewed as a set of propositions or statements that are tested and confirmed by evidence. Scientific change, then, usually means to revise this set of sentences. This assumption is exactly what Kuhn and many other researchers from a historical minded philosophy of science (e.g., Thagard 1992) reject: Scientific development also involves the generation and revision of concepts.

In order to address and analyse this criticism, I will discuss conceptual change and conceptual learning in a dynamic epistemic logic (DEL). This framework has proven to be broadly applicable to the modelling of very different types of rational dynamics including the logic of public announcements, belief revision or game theory (van Benthem 2011). Because of its flexibility, it can also be used to study conceptual dynamics in a formal language.

The paper is structured as follows: The next section makes some general remarks about the notion of analytic laws and conceptual change. It also presents prior work on the formal reconstruction of conceptual change in formal theories of belief dynamics. The subsequent section focuses on a recent framework that is based on DEL (Strößner 2020). In the following sections, I extend my previous work by explicitly considering conceptual learning. The discussion is focused on how these conceptual dynamics influence translatability of the new language and the ability to discriminate between objects. The fifth section discusses my approach with respect to a historic example, namely the development of the category of mammals.

2 Conceptual Change and Formal Epistemology

This section sets out the motivation for the study by considering philosophical worries with respect to conceptual change and formal epistemology. Special attention is given to the notion of incommensurability. I will then give a short outline of existing approaches to conceptual change in formal epistemology.

2.1 Conceptual Change: A Problem for Formal Epistemology?

Quine's influential attack on logical empiricism, and particularly on the analytic / synthetic distinction in "Two Dogmas of Empiricism" (Quine 1951) rests on the idea that analytic (i.e., conceptual) truth entails a lack of revisability. This is evident when he notes that "it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements which hold come what may" and that "no statement is immune to revision" (Quine 1951, 40). Grice and Strawson (1956) answer Quine by explicitly mentioning the possibility of conceptual change. Based "on the distinction between that kind of giving up which consists in merely admitting falsity, and that kind of giving up which involves changing or dropping a concept or set of concepts" (Grice and Strawson 1956, 157), it is well possible to concede general revisability of all statements *and* to hold on to the notion of analytic truth. Nevertheless, Quine's criticism had a large impact, especially on logical empiricism.[1] An important reason is that many formal theories of belief dynamics did and still do suppose a fixity of language.

This brings us to what I have called Kuhn's worry, namely that scientific changes often entail a lack of translatability, that is, *incommensurability*. There is hardly any other concept that has influenced the philosophy of science of the 20th century as much as that of incommensurability. The idea can be traced back to Fleck (1979) (German original: Fleck (1935)), but became more prominent within the work of Kuhn (1962) and Feyerabend (1975). At the heart of his seminal book "The structure of scientific revolutions", Kuhn (1962) analysed the history of scientific development in terms of normal science, which basically is puzzle solving within a research paradigm, and scientific revolutions, which occur out of a crisis and alter the research paradigm. In his depiction, revolutionary change involves conceptual change.

In his later writings, especially in Kuhn (1983), he extended and clarified his notion of "incommensurability" in a way which puts an even clearer focus on aspects of meaning change. He defines "incommensurability" as "no common language" (Kuhn 1983, 36). Incommensurability excludes the possibility of translation, where translatability indicates the possibility to create an equivalent text:

> Confronted with a text, written or oral, in one of these languages, the translator systematically substitutes words or strings of words in the other language for

[1] For a detailed reconstruction of the debate see Leitgeb and Carus (2020, supplement B).

words or strings of words in the text in such a way as to procude an equivalent text in the other language. (Kuhn 1983, 38)

For Kuhn, equivalency goes beyond mere truth values of sentences or the reference of the terms, but considers the meaning, that is the intension of language (cf. Kuhn 1983, 41). Languages that are not translatable to our language (e.g., the language of Aristotelian physics) are incommensurable with it. Kuhn uses the notion of *local incommensurability* if this lack of translatability is restricted to some parts of the theories. According to Kuhn, a historian of science needs to (partially) learn the older language and its way to discriminate between objects in order to understand the theory expressed in it. This is not possible if a reader only knows the concepts of the more recent theories.

The main aim of this article is to logically reconstruct the arising of untranslatability by conceptual change. Is there a logical way to recapture how conceptual change leads to an untranslatable successor language? Before focusing on this central question, let us shortly review how conceptual change has been addressed in formal epistemology and other formal frameworks so far.

2.2 The Representation of Conceptual Change in Formal Epistemology

There are in general two kind of research strategies in the formal analysis of conceptual change. The first one is the development of formal systems that are explicitly focused on concepts and conceptual change. The second one is the incorporation of concept learning within existing frameworks of formal epistemology such as Bayesianism or epistemic logic. Research within the first tradition, that is, the development of concept-based frameworks of science, are quite advanced. Many of them draw from representational models of cognitive science or are closely related to them. Notable contributions in this direction are for example coming from Thagard (1992, 2012), who devotes much of his work to the discussion of what he calls “conceptual revolutions”. Other researches have used frames, that is, attribute-value structures (Barsalou 1992; Minsky 1975), and transformed them into a formal tool for reconstructing scientific change. Such frame-based reconstructions are, for example, found in Andersen and Nersessian (2000) and in Kornmesser and Schurz (2020). Moreover, conceptual spaces theory (Gärdenfors 2000) has been applied to study scientific changes (De Benedetto 2020; Gärdenfors and Zenker 2013). All these frameworks are extremely useful, also because of their connections to related disciplines of cognitive science, such as artificial intelligence. However, they are based on deviations from or even outright rejection of formal epistemology as a framework of representing scientific developments. As such they cannot be regarded as contributions to *extend* formal epistemology into a broader framework that is capable to treat belief revisions as well as conceptual change.[2]

[2] Since scientific change often also incorporates changes in the action of scientists, dynamics of decision spaces or habitual spaces (Skulimowski 2011; Yu 1991) as well as models of epistemic groups (O’Connor 2019) are relevant for modelling scientific dynamics. However, I will not discuss such approaches here, since they are quite different from formal epistemology.

Though there is little research on conceptual dynamics *within* formal epistemology, several contributions have been made. A central aspect in this line of research is the relation between conceptual change and factual belief. To which degree does conceptual learning alter beliefs? What about the *principle of language invariance* according to which "in the absence of any change in factual knowledge, an agent's belief function should not change as her language changes" (Williamson 2003, 62)?

Huber (2015) introduces a new type of conditionalization to ranking theory (Spohn 2012) that models conceptual extensions and logical changes in the underlying language. He emphasises that the conceptual alteration is, at least formally, independent of a factual belief revision. As far as conceptual learning is accompanied by factual learning, the latter is represented as a separate step of updating the ranking function. His framework thus fully satisfies language invariance.

On the contrary, Williamson (2003) takes a Bayesian perspective and explicitly denies that conceptual changes can or even should be independent of factual assumptions. Concepts, he argues, carry knowledge or beliefs. For this reason he substitutes language invariance with a weaker principle, called *conservativity*. That means, unless there is a conceptual reason to update beliefs, one remains as close to the old beliefs as possible. Thus, according to Williamson, a belief state after a conceptual change meets the following three demands with a decreasing priority (cf. Williamson 2003, 69):

1. First, it includes beliefs that are implicitly assumed in the new concepts or which motivate conceptual change, that is, the so-called transitional knowledge.
2. Second, the new belief state resembles the old belief state as far as the transitional knowledge allows it.
3. Finally, the new belief state remains as neutral as possible. That is, it assumes conditional independence as far as compatible with the transitional and the old beliefs.

Williamson implements these rules in Bayes nets. By this means, he outlines a probabilistic theory of a rational conceptual change.[3]

Besides this formal treatment of conceptual change, Williamson justifies the idea that conceptual change should be conservative because "it is a waste of time, energy and resources to continually change our beliefs for no reason, or to change them more than the minimum amount" (Williamson 2003, 63). He highlights the demand of continuity within conceptual change. With respect to incommensurability, he emphasises that there is a transitional language, which contains the old language together with the newly learned concepts.[4]

[3] A formally similar procedure is proposed by Romeijn (2005) to model changes of partitions in the statistical testing of theories. He calls this a change of hypotheses.

[4] At least, this is to be expected for the directly succeeding language. The demand cannot be straightforwardly applied for far conceptual systems, say Aristotle's and Einstein's physics, which play a major role in Kuhn's argumentation.

In Strößner (2020) I recently proposed a theory of conceptual change that is based on DEL. The approach employs a monadic predicate logic with a domain of (logically) possible objects instead of possible worlds. These include, besides common things such as round apples, implausible, far-fetched objects such as flying wolves. An interpretation I fixes the meaning of a predicate $\boldsymbol{\Phi}$ as the set of those possible entities that count as instances of the predicate $\boldsymbol{\Phi}$ or would count as such if they existed. The term "logically possible" means a relaxed modality that includes entities that are disbelieved to exist, entities that are in contradiction to laws of nature as well as entities that are not cognitively accessible to us. What logical possibility means is further specified in the semantics of the system. For example, in Strößner (2020) every possible object is either member of a category $I(\boldsymbol{\Phi})$ or of its complement $I(-\boldsymbol{\Phi})$.

The approach closely resembles what Carnap suggests in "Meaning and synonymy in natural languages" (Carnap 1955):[5]

> All logically possible cases come into consideration for the determination of intensions. This includes also those cases that are causally impossible, i.e., excluded by the laws of nature holding in our universe, and certainly those that are excluded by laws which Karl [exemplary agent, ed. note] believes to hold. (Carnap 1955, 38)

Among the objects that determine the meaning of a predicate $\boldsymbol{\Phi}$, (2020) I distinguished plausible ones $P(\boldsymbol{\Phi}) \subseteq I(\boldsymbol{\Phi})$, which are in accordance with our empirical observations and the laws of nature. Based on this distinction, analytic laws are defined as generalisations that hold among $I(\boldsymbol{\Phi})$ while doxastic laws need only to hold for $P(\boldsymbol{\Psi})$. Analytic truth is thus modelled as truth among all logically and conceptually possible members, while doxastic laws merely need to hold for plausible members. I modelled conceptual change as an alteration of the interpretations function I, which influences the set of possible objects that belong to $I(\boldsymbol{\Phi})$ and thus the analytic laws for $\boldsymbol{\Phi}$. Not that this change does not necessarily affect $P(\boldsymbol{\Phi})$, the plausible instances of $\boldsymbol{\Phi}$. The doxastic laws of $\boldsymbol{\Phi}$ *can* remain the same after the conceptual change. In Strößner (2020) I argue that this is even expected for several forms of rationally justified predicate changes, for example if a firm belief about a concept becomes part of its definition. In response to Quine's critic against analyticity, this shows that there is a meaningful way to formally capture analytic laws and their revision. However, not too much depends on the notion of analyticity, because the status of analytic laws versus doxastic ones is only important for the classification of implausible entities (cf. Strößner 2020, 1180).

In this paper, I will build up on my framework in Strößner (2020). The following section provides a brief outline of the formal system and explains its philosophical interpretation, especially in relation to translatability and incommensurability.

[5] Chalmers (2012) notes that this, in his view underappreciated article, can be read as a direct answer to Quine's "Two dogmas of empiricism".

3 Predicate Change

Predicate change in Strößner (2020) is based on the languages AD, a monadic predicate logic with analytic and doxastic laws, and ADT, a language with additional typicality laws. In the first part of this section, I shortly recapture these systems and relate them to Kuhn's idea of incommensurability.

3.1 AD

Syntactically, the language AD consists of unary predicates (atomic and complex ones), so-called law operators, which form sentences from predicates, and the classical connectives:

Definition 1 (Syntax) The following strings are formulae of AD:

- Given a non-empty set of atomic predicates $\mathfrak{P}$, $\Phi \in \mathfrak{P} | \Phi \cup \Psi | \Phi \cap \Psi | -\Phi$ are predicates of AD.
- If Φ and Ψ are predicates of AD, then $\mathcal{A}\Phi\Psi$ and $\mathcal{D}\Phi\Psi$ are sentences of AD.
- If ϕ and ψ are sentences of AD, $\neg\phi$, $\phi \wedge \psi$, and $\phi \vee \psi$ are sentences of AD.

The sentence-building operators $\mathcal{A}$ and $\mathcal{D}$ are used to formally capture generalizations about predicates. Sentences like $\mathcal{A}\Phi\Psi$ are called analytic laws and $\mathcal{D}\Phi\Psi$ doxastic laws. A model of AD is defined as follows (cf. Strößner 2020, 1161):

Definition 2 (Model of AD) A model $\mathbf{M}$ of AD is a triple $\langle O, I, \geqslant \rangle$ where O is a finite set of logically possible objects, I is an interpretation that assigns subsets of O to predicates, and $\geqslant$ is a totally connected, reflexive, and transitive plausibility relation among the elements of O.

In Strößner (2020, 1162) I called the elements of O conceivable objects and compared them to (logically) possible worlds. Note, however, that "conceivable object" does not mean that an agent is cognitively able to imagine or describe such an object. The interpretation function I assigns meaning to the predicates by determining which of such objects are possible instances of a predicate Φ. $I(\Phi)$ is also called the category of Φ, that is, the set of logically possible instances of the concept. Other than in a standard first order predicate logic, the interpretation function is rather intensional than extensional. In this respect, AD models do not fall prey to Kuhn's criticism of merely referential reconstruction. They accord to his demand "that something from the realm of meanings, intensionalities, concepts must be invoked as well" (Kuhn 1983, 47).

It is not excluded that there are two possible objects that are in exactly the same categories, i.e., I assigns them to the same predicates. In this case, there could be another model with the same set of possible objects O but a slightly different interpretation function I that allows to discriminate between them and that has in this respect a greater expressive power.[6] This gives rise to a formal characterization of what incommensurability, in the sense of untranslatability (Kuhn 1983), means:

> *Incommensurability* Assume a set of (atomic) predicates ($\mathfrak{P}$), a set of possible objects O, and two different interpretations I_1 and I_2. I_1 and I_2 are incommensurable if there exist two $o_1, o_2 \in O$ such that o_1 and o_2 are *indiscriminable* in exactly one of the interpretations I_1 or I_2, where indiscriminable means that for all $\boldsymbol{\Phi} \in \mathfrak{P}$ it holds that $o_1 \in I(\boldsymbol{\Phi})$ if and only if $o_2 \in I(\boldsymbol{\Phi})$.

This formal characterization represents a situation in which one cannot grasp a distinction made in one language by using another incommensurable one. Moreover, the formal description can also account for Kuhn's notion of *local* incommensurability. The number of possible objects with a difference of discriminability can be a very small subset of possible objetcs.

The idea to relate the notion of incommensurability and possible worlds semantics is briefly mentioned in Kuhn (1989, 64), where he states that "only the possible worlds stipulatable in that language can be relevant to them". In a formal model of AD, however, the set of (logically) possible objects is just given and it is not necessary that the interpretation function allows to give a unique description for each object. In order to grasp Kuhn's idea of a possible world—or a possible object—as depending on predicates and their interpretation, one can introduce *atomic constructable objects*. From each predicate $\boldsymbol{\Phi}$, one either takes $\boldsymbol{\Phi}$ or its complement $-\boldsymbol{\Phi}$. One connects them by the intersection $\cap$ and thus produces a complete description of an object in one complex predicate. If the interpretation I of this predicate is non-empty, then this complex predicate refers to such a stipulatable object in this language.

What becomes apparent in my formal characterization and what is alien to Kuhn's texts is that incommensurability is not necessarily mutual. It is possible that I_1 allows to discriminate all the discriminable objects of I_2 but not vice versa. In this situation, the language of I_2 is (locally) incommensurable with the language of I_1, because a description of an object in terms of I_1 has no translation in I_2. On the other hand all objects from I_2 can be described in I_1. The I_1-language is in this sense not incommensurable with the I_2-language: I_1 has a strictly greater discriminatory power than I_2.

Besides the interpretation and the set of possible objects, an AD model contains a plausibility rating $\geqslant$. The highest ranked entities in $\geqslant$, namely the set $\{x \in O | \neg\exists y(y \in O \wedge x \ngeqslant y)\}$ (equivalently: $\{x \in O | \forall y(y \in O \rightarrow x \geqslant y)\}$), are plausible members. This set includes entities of which we strongly believe that they exist as well as those about the existence of which we suspend judgement. The

[6] In model theory, the terms "expressive power" or "discriminatory power" are commonly used to refer to the language's ability to distinguish between different models. In this paper, it refers to the ability of an interpretation function (together with the predicates) to distinguish between the elements of O, i.e., the logically possible objects.

plausibility $\geqslant$ allows for integrating belief revision, that is, the updating of beliefs in light of unexpected evidence.[7]

Based on the model, the semantics of AD are defined as follows (Strößner 2020, 1162–1164):

Definition 3 (Semantics of AD)

- Interpretation
  $I(\Phi) \subseteq O$, $I(-\Phi) = O/I(\Phi)$, $I(\Phi \cap \Psi) = I(\Phi) \cap I(\Psi)$, $I(\Phi \cup \Psi) = I(\Phi) \cup I(\Psi)$.
- Law operators
  $\mathbf{M} \vDash \mathcal{A}\Phi\Psi$ iff $I(\Phi) \subseteq I(\Psi)$.
  $\mathbf{M} \vDash \mathcal{D}\Phi\Psi$ iff $P_{\mathbf{M}}(\Phi) \subseteq I(\Psi)$, where $P_{\mathbf{M}}(\Phi) = \{x \in I(\Phi) | \neg\exists y(y \in I(\Phi) \wedge x \ngeqslant y)\}$.
- Propositional connectives
  $\mathbf{M} \vDash \neg\phi$ iff $\mathbf{M} \nvDash \phi$,
  $\mathbf{M} \vDash \phi \wedge \psi$ iff $\mathbf{M} \vDash \phi$ and $\mathbf{M} \vDash \psi$,
  $\mathbf{M} \vDash \phi \vee \psi$ iff $\mathbf{M} \vDash \phi$ or $\mathbf{M} \vDash \psi$.

The determination of the interpretation function is intuitively obvious. The predicate connectors $(-, \cup, \cap)$ are just used like the basic set theoretic notions of relative complement, union, and intersection.[8] The meaning of the connectives is the common one from propositional logics.

The law operators $\mathcal{A}, \mathcal{D}$ both stand for universal statements of the form "All Φ are Ψ". $\mathcal{A}\Phi\Psi$ means that all *logically possible* instances of Φ are Ψ. This is what characterises analytical statements.[9] Doxastic laws $\mathcal{D}\Phi\Psi$, on the other hand, merely demand that Ψ is true among the most plausible members of $I(\Phi)$. The language AD is a reformulation of a conditional logic (Lewis 1973) or a preferential logic (Kraus et al. 1990) in terms of predicates instead of propositions. The operator $\mathcal{A}$ resembles the strict conditional and $\mathcal{D}$ a variably strict conditional (cf. Strößner 2020, 1165).

[7] An implementation of belief revision in DEL that is based on such plausibility rating is provided by van Benthem (2004).

[8] This is admittedly a very simplified understanding of composing complex predicates. For example, intuitively it seems more appropriate to consider dogs as non-cats than to consider houses as non-cats because they are from a completely different domain. Moreover, the framework cannot account for cases of unclear or unknown category membership. Moreover, natural language composition will often involve more than intersection. For example, 'black female' would not be interpreted as intersection of females and black entities. The advantage of the formalism is, however, its simplicity and its intuitive relation to the logical operators of negation and conjunction.

[9] AD allows for a distinction between analytic and logical truth. There can be logically contingent statements that are analytically true. This is generally the case if there are two atomic predicates Φ and Ψ such that $I(\Phi) \subseteq I(\Psi)$.

3.2 ADT

In addition to an epistemic preference, which distinguishes objects in terms of their plausibility, I included a second kind of ordering that orders objects with respect of their expectedness for a predicate and called the resulting system ADT.

The model of ADT includes a partial prototype function, which can be used to designate a central instance of an atomic predicate, and a weakly centred comparative similarity relation (cf. Strößner 2020, 1166):

Definition 4 (Model of ADT) A model $\mathbf{M}$ of ADT is a quintuple $\langle O, I, \geqslant, \gtrsim_{\_}, Pt \rangle$ where O, I, and $\geqslant$ are as in AD models, $\gtrsim_{\_}$ is a ternary, weakly centred comparative similarity relation, and Pt is a partial function that assigns a prototype to some atomic predicates.

The similarity relation $x \gtrsim_z y$, read as "x is at least as similar to z than y", is used to define a typicality ordering in terms of the similarity to the prototype: x is at least as typical for Φ as y if and only if $x \gtrsim_{Pt(\Phi)} y$. The following definition of typicality laws extends AD to ADT (Strößner 2020, 1167):

Definition 5 (Typicality laws of ADT) ADT adds typicality laws to AD:
If Φ and Ψ are predicates, then $\mathcal{T}\Phi\Psi$ is a sentence.
$\mathbf{M} \models \mathcal{T}\Phi\Psi$ iff $T_{\mathbf{M}}(\Phi) \subseteq I(\Psi)$, where $T_{\mathbf{M}}(\Phi) = \{x \in P_{\mathbf{M}}(\Phi) | \neg\exists y (y \in P_{\mathbf{M}}(\Phi) \wedge x \not\gtrsim_{Pt(\Phi)} y \wedge y \gtrsim_{Pt(\Phi)} x)\}$.

3.3 Predicate Change

The main focus of Strößner (2020) is the discussion of predicate change. Two variants of predicate change are defined: an inclusive one, in which new members are added to an atomic predicate Φ, and an exclusive one, in which members are excluded but no new member enters the category, always with respect to another predicate Φ (cf. Strößner 2020, 1169):

Definition 6 (Predicate change)

- Let Ψ be an atomic predicate. The inclusion of Ψ with respect to Φ on $\mathbf{M}$ yields the model $\mathbf{M}|\Psi \uparrow \Phi$, where I in $\mathbf{M}$ is changed to I' in $\mathbf{M}|\Psi \uparrow \Phi$ with $I'(\Psi) = I(\Psi \cup \Phi)$ and $I'(\Xi) = I(\Xi)$ for all other atomic predicates Ξ.
- Let Ψ be an atomic predicate. The exclusion of Ψ with respect to Φ on $\mathbf{M}$ yields the model $\mathbf{M}|\Psi \downarrow \Phi$, where I in $\mathbf{M}$ is changed to I' in $\mathbf{M}|\Psi \downarrow \Phi$ with $I'(\Psi) = I(\Psi \cap \Phi)$ and $I'(\Xi) = I(\Xi)$ for all other atomic predicates Ξ.

Real historical examples of these changes are the exclusion of Pluto from the category of planets or the inclusion of cetaceans (whales and dolphins) into the class of mammals, which will be further discussed in Section 5. The exclusion of Pluto can be formally reconstructed (where P means "planet" and D means "dominates its orbit") as a change from $\mathbf{M}$ to $\mathbf{M}|P \downarrow D$ with $I'(P) = I(P \cap D)$. The inclusion of cetaceans can be captured (where M means "mammal" and C means "cetacean") as $\mathbf{M}|M \uparrow C$ with $I'(M) = I(M \cup C)$.

The languages AD/ADT are extended by change operations that express the conceptual state after the the change.

Definition 7 (Change Operators) $[\Psi \uparrow \Phi]\phi$ and $[\Psi \downarrow \Phi]\phi$ are sentences of dynamic AD / ADT iff Ψ is an atomic predicate, Φ is a (possibly complex) predicate, and ϕ is a sentence.

- $\mathbf{M} \vDash [\Psi \uparrow \Phi]\phi$ iff $\mathbf{M}|\Psi \uparrow \Phi \vDash \phi$, and
- $\mathbf{M} \vDash [\Psi \downarrow \Phi]\phi$ iff $\mathbf{M}|\Psi \downarrow \Phi \vDash \phi$

The central result of Strößner (2020) is that any AD formula with predicate change can be translated to an equivalent one without predicate change.[10]

Based on this observation, I claimed that "predicate change, while being a kind of conceptual change, [...] is far from causing incommensurability" (Strößner 2020, 1171). In view of the above given explication of incommensurability, this requires that there is no gain but also *no loss* of discriminatory power. However, the redundancy of predicate change only entails that there is *no increase* of expressive power but not that the discriminatory ability is the same. A predicate change can be redundant while still decreasing expressive abilities.

Indeed, predicate change, as outlined in Strößner (2020), can lead to such a loss of discriminatory and expressive power. Although every doxastic or analytic law of the new system (i.e., after the change) has an equivalent formula in the old system, the reverse is not true. This becomes obvious in the following example:

Example 1 Assume a simple language with two atomic predicates S and P and a model $\mathbf{M}$ with $O = \{o_1, o_2, o_3, o_4\}$, $\geqslant$ such that $o_1 = o_3 = o_4 > o_2$, and $I(S) = \{o_1, o_2\}, I(P) = \{o_1, o_3\}$. In $\mathbf{M}$, every conceivable object has a unique description: $\{o_1\} = I(S \cap P)$, $\{o_2\} = I(S \cap -P)$, $\{o_3\} = I(-S \cap P)$, $\{o_4\} = I(-S \cap -P)$. However, predicate changes undermines the expressive power:

1. In $\mathbf{M}|S \uparrow P$, I becomes I' with $I'(S) = \{o_1, o_2, o_3\}$ and $I'(P) = \{o_1, o_3\}$. The objects o_1 and o_3 are no longer distinguishable because both are in $I'(S \cap P)$.
2. In $\mathbf{M}|S \downarrow P$, I becomes I' with $I'(S) = \{o_1\}$ and $I'(P) = \{o_1, o_3\}$. Objects o_2 and o_4 are no longer distinguishable because both are in $I'(-S \cap -P)$.

[10] The translatability of typicality laws presupposes the introduction of ternary typicality laws $\mathcal{T}^{\Xi}\Phi\Psi$, where the typicality relation of Ξ is applied to a different concept Φ. These statements have the following semantics: $\mathbf{M} \vDash \mathcal{T}^{\Xi}\Phi\Psi$ iff $T^{\Xi}_{\mathbf{M}}(\Phi) \subseteq I(\Psi)$, where $T^{\Xi}_{\mathbf{M}}(\Phi) = \{x \in P_{\mathbf{M}}(\Phi) | \neg\exists y(y \in P_{\mathbf{M}}(\Phi) \wedge y \gtrsim_{Pt(\Xi)} x \wedge x \not\gtrsim_{Pt(\Xi)} y)\}$ (Strößner 2020, 1171). A possible natural language example would be: "Fish-typical mammals live in the ocean".

3. In $\mathbf{M}|S \uparrow -P$, I becomes I' with $I'(S) = \{o_1, o_2, o_4\}$ and $I'(P) = \{o_1, o_3\}$. Again o_2 and o_4 are no longer distinguishable because both are in $I'(S \cap -P)$.
4. In $\mathbf{M}|S \downarrow -P$, I becomes I' with $I'(S) = \{o_2\}$ and $I'(P) = \{o_1, o_3\}$. Again o_1 and o_3 are no longer distinguishable because both are in $I'(-S \cap P)$.

In these examples, the changes lead to a loss of discriminatory possibilities. As such they also lead to a lack of translatability. One can no longer formulate laws for these entities. For example, the sentence $\mathcal{D}SP$ in $\mathbf{M}$ expresses that $-P$ are not among the plausible members of S (o_1 is more plausible than o_2). In $\mathbf{M}|S \downarrow P$ this belief becomes analytically true and does not depend on the plausibility of o_2. Generally, there is no way left to express the implausibility of o_2.

This short example demonstrates that even supposedly mild changes of a conceptual system can entail untranslatability and, in this sense, local incommensurability.

4 Conceptual Learning

The remainder of this paper studies an issue, which I did not address in Strößner (2020), namely the adding or deletion of concepts. The main focus of my analysis is whether and to which degree they lead to untranslatability. My formal treatment of conceptual learning assumes that the language has a set of atomic predicates with an empty interpretation. Thus, rather than altering the set of predicates, conceptual learning will be described as a merely semantic change by which an empty predicate is associated with a set of logically possible objects. In analogy to natural language, empty predicates are like meaningless but well-formed and pronounceable sound patterns. When concepts are created or learned, we start to attach a meaning to them, that is, objects that might fall into the category.

By definition, analytic, doxastic and typicality laws are vacuously true for empty predicates. This is not particularly noteworthy because most formal treatments of universal laws (and conditionals) assume trivial truth for empty subjects (or antecedents, respectively). However, for predicate learning, understood as attaching meaning to a previously meaningless predicate, this entails that by learning the meaning of a predicate, one does not gain new laws but rather eliminates (vacuously true) laws. To avoid this philosophical obscurity, I introduce the following distinction:

- A sentence ϕ is a *proper sentence* in model $\mathbf{M}$ iff ϕ contains no atomic predicate Ψ with $I(\Psi) = \emptyset$.
- A sentence ϕ is a *proper* law in model $\mathbf{M}$ iff it is a proper sentence in model $\mathbf{M}$ and $\mathbf{M} \models \phi$.

With this distinction in place, it is straightforward to state that conceptual learning increases the set of *proper* sentences and thus, at least potentially, also the set of *proper* laws. In what follows, I discuss three forms of conceptual change: the definition of new concepts, the prototype-based learning of a concept, and, finally, the elimination of concepts.

4.1 Definitional Learning

A basic form of conceptual learning is to define a concept. For example, we define "bachelor" as meaning an unmarried adult male.

Definition 8 (Definitional learning) Definitional learning of Ψ as Φ on $\mathbf{M}$, where Ψ is an atomic predicate with $I(\Psi) = \emptyset$ and Φ is a (potentially complex) predicate with $I(\Phi) \neq \emptyset$, yields the model $\mathbf{M}|\Psi = \Phi$, in which $I \in \mathbf{M}$ is changed to $I' \in \mathbf{M}|\Psi = \Phi$ with $I'(\Psi) = I(\Phi)$ and $I'(\Xi) = I(\Xi)$ for all other atomic predicates Ξ.

This version of conceptual learning is a variant of an inclusive predicate change as determined above in definition 6: $\Psi = \Phi$ is a special case of $\Psi \uparrow \Phi$, where Ψ is initially meaningless. That is, we "extend" the meaning of the previously meaningless predicate by the definiens.

As a form of inclusive predicate change, definitional learning cannot enlarge expressive power. However, unlike inclusive predicate change for meaningful predicates, it also entails no loss of discriminatory abilities, since the lost meaningless predicate Ψ obviously did not contribute to the discriminatory abilities of the interpretation function I.

By definitional learning, some previously trivially true sentences become false. For example, both $\mathcal{A}\Phi - \Psi$ and $\mathcal{A}\Phi\Psi$ were trivially true as long as Φ was meaningless. However, no *proper* law of the old system is lost by the definitional learning of Ψ. The new proper laws for the defined predicate are inherited from the definiens and are, in this sense, not particularly novel.[11] Definitional learning is thus a quite mild and conservative form of conceptual learning. In particular, the new model $\mathbf{M}|\Psi = \Phi$ and the old model $\mathbf{M}$ have exactly the same discriminatory abilities. Translation is certainly possible and incommensurability (even local one) is excluded.

Given its apparent lack of creative power, one might question whether definition should be counted as a form of conceptual learning. It seems that the whole gain in defining a concept is to generate an abbreviation, e.g. "bachelor" abbreviates "unmarried ∩ adult ∩ male". However, by becoming atomic, the newly defined predicate becomes independent from the constituents that formed the complex predicates and the defined predicate can depart from them. For example, after defining "bachelor" as meaning unmarried, adult males, one can change "bachelor" by excluding Catholic priests. After that "bachelor" is no longer synonymous with "unmarried ∩ adult ∩ male".

[11] The laws of the definiens and definiendum are largely identical. However, if the definiens is an atomic predicate with a prototype, it can have typicality laws that are not inherited by the defined predicate.

4.2 Learning by Prototypes

A gain of discriminatory abilities cannot be reached by definitions. As such, conceptual learning as enlargement of expressive abilities needs to take different routes. In ADT models, such a more creative way of learning can be grounded in prototypes. A new concept is learned by a prototypical object o and the category $I(\Psi)$ is formed from all conceivable objects which are as similar to o as o is to itself:[12]

Definition 9 (Learning by prototype) The prototype learning of a concept Ψ with prototype o on $\mathbf{M}$, where $o \in O$ and Ψ is an atomic predicate with $I(\Psi) = \emptyset$, yields the model $\mathbf{M}|\Psi \approx o$, in which Pt and I in $\mathbf{M}$ are changed to I' and Pt' in $\mathbf{M}|\Psi \approx o$ with $Pt'(\Psi) = o$ and $I'(\Psi) = \{x \in O|x \gtrsim_o o\}$. For all other atomic predicates Ξ: $Pt'(\Xi) = Pt(\Xi)$ and $I'(\Xi) = I(\Xi)$.

An obvious motivation to form such a new concept is the discovery of an object that does not fit in the existing conceptual system. However, a prototype-based learning can also be based on an abstract idea. Something in this realm is mentioned in Fleck (1979, 23–24). He speaks of “proto-ideas” or “pre-ideas”. The basic point in his account is that a pre-idea is a creative starting point in a long development from a hazy foundation to a full scientific concept.

Though the prototype takes a central role in learning, the concepts learned by prototypes have no proper typicality laws in the beginning. There are only equally typical members in a category and thus all typicality laws of the learned predicate Ψ are merely trivial counterparts of the according doxastic laws ($\mathbf{M}|\Psi \approx o \vDash \mathcal{T}\Psi\Phi$ if and only if $\mathbf{M}|\Psi \approx o \vDash \mathcal{D}\Psi\Phi$). Atypical objects are not included in the initial concept learning but can be added in further conceptual development, most notably by inclusive predicate changes.

Prototype learning is a productive form of conceptual learning. It leads to proper laws concerning the learned predicate. Most notable, unlike definitions, the prototype-based learning has the potential to increase discriminatory abilities and expressive power. Formerly indistinguishable conceivable objects can be differentiated. That means, the new language has sentences that cannot be translated to equivalent ones in the old language. As a quite radical example, consider a model with only meaningless predicates. Through a prototype-based learning the first meaningful predicate can be learned.

Since it enlarges the expressive power, prototype-based learning of concepts can lead to incommensurability. However, the incommensurability is unilateral: Obviously, the learning of the new predicates entails no loss of expressive power. Everything expressible in the old system is also expressible in the new system. Moreover,

[12] The idea is somewhat reminiscent of Voronoi tessellation, in which points are classified to closest points in a space (cf. Gärdenfors 2000). However, our approach is more simplistic in that its similarity relation is not quantitative (in contrast to a geometrically based notion of closeness). Moreover, Voronoi tessellations partition the whole space of possible objects, which is not required in our definition of prototype-learning.

prototype-based learning does not undermine the truth of the proper laws of the previous system.

4.3 Eliminating Concepts

Having introduced two forms of conceptual learning, I now consider forgetting concepts. By conceptual elimination, one changes the interpretation I of a previously meaningful predicate Ψ to $I(\Psi) = \emptyset$. If the concept had a prototype, it needs to be eliminated as well. That means, a predicate which previously had a meaning (i.e., logically possible instances), becomes meaningless. The formal definition runs as follows:

Definition 10 (Conceptual elimination) Conceptual elimination of the atomic predicate Ψ on $\mathbf{M}$ yields the model $\mathbf{M}|\Psi = 0$, where I is changed to I' and Pt to Pt' such that $I'(\Psi) = \emptyset$ and $Pt'(\Psi) = \emptyset$. For all other atomic predicates Ξ: $I'(\Xi) = I(\Xi)$ and $Pt'(\Xi) = Pt(\Xi)$.

Note that conceptual forgetting can be viewed as a kind of exclusive predicate change according to definition 6, namely $\mathbf{M}|\boldsymbol{\Psi} \downarrow \boldsymbol{\Phi} \cap -\boldsymbol{\Phi}$. Unsurprisingly, this kind of predicate change exhibits no increase in expressive power but may lead to a loss of expressive power. Several statements that were expressible before the conceptual elimination are inexpressible afterwards, which seems *prima facie* disadvantageous. However, in view of our limited cognitive capacities and a general need for efficiency, conceptual elimination is necessary. But *which* concepts will be eliminated? One possible motivation to eliminate a *particular* predicate is that, according to our belief state, no plausible instances exist. Another one is redundancy, that is, there is a (potentially complex) predicate with the same meaning. Note that in the latter case, the elimination of the atomic concept will not influence the discriminatory powers of the interpretation function I and thus it will not hinder translatability. However, it might lead to a loss of typicality laws.

5 A Case Study: Cetaceans, Mammals and Fish

In this section, I aim to bring together the formal and philosophical aspects of this paper in a case study. Thomas Kuhn repeatedly reports about the difficulties he encountered when trying to understand Aristotelian physics. Many passages seemed to make no sense unless one acquired the Aristotelian conceptual background as well. Similar problems can be found in Aristotle's biological work. Aristotle devoted a considerable amount of his research to the study of animals. His work in this area is to a large extend valid up till now (e.g. Lieven and Humar 2008). It is not so clear whether the classification of animals into higher taxa was an intrinsic goal for Aristotle. However, he repeatedly builds groups of several animals: those that have hair, those with hearts, with lungs etc. and makes claims about the further universal properties of these animals. This pattern of *hosa...panta* (as many as are... all are) is characteristic of *Historia Animalium* (HA)(see Lennox 2019).[13] At the beginning of HA, Aristotle mentions some categories of animals:

> Very extensive genera of animals, into which other subdivisions fall, are the following: one, of birds; one, of fishes; and another, of cetaceans. HA 490b08-10 (Barnes 1984, 1717)

While Aristotle was a careful observer of animals and possessed enormous amount of empirical knowledge about them, one also finds passages that are obscure from a contemporary view:

> The dolphin, the whale, and all the rest of the cetacea, all, that is to say, that are provided with a blow-hole instead of gills, are viviparous. [...] That is to say, no one of all these fishes is ever seen to be supplied with eggs, but directly with an embryo from whose differentiation comes the fish, just as in the case of mankind and the viviparous quadrupeds. HA 566b1-7 (Barnes 1984, 1945-46)

While Aristotle correctly describes the mammalian features of whales and dolphins, he repeatedly calls them "fish". Modern readers, for example Romero (2012), are astonished from this seeming contradiction between empirical knowledge and an inappropriate terminology.

The example of whales as fish (or non-fish) is also known from Carnap, who argues that the exclusion of whales from fish was a conceptual change and not a belief revision: "The change which zoologists brought about in this point was not a correction in the field of factual knowledge but a change in the rules of the language" and he adds that this change "was motivated by factual discoveries" (Carnap 1962, 5-6). However, it is still intriguing that the factual knowledge behind this conceptual change was fully available to Aristotle. From a modern view, he sticks to a "wrong" terminology while giving reasons against its appropriateness.

In order to understand Aristotle's terminology, one should first consider the folk-taxonomy on which Aristotle had to rely in his description. Natural (pre-scientific) languages have usually no concept of mammals. Aristotle himself notes that there is no joint name for the many animal species that he characterizes as viviparous quadrupeds. Names for the category of fish, on the other hand, are common in natural language. As most other concepts from natural language, they are plausibly developed by what I called prototype-based learning. That means, the (pre-scientific) concept of fish means entities that are sufficiently similar to a fish-prototype. Aristotle's usage of the concept is grounded in this folk-taxonomy. On the other hand, he stated that there are larger genera of animals and names cetaceans and fish beside each other, which seems to indicate that his notion of fish did not include whales. This apparent contradiction becomes less paradox if we accept that, other than modern zoologists, he did not aim for listing contrasting, that is, analytically exclusive, categories of animals. This interpretation is also backed up by passages in *De Partibus*

[13] Lennox (2019) argues that this is not only for the sake of pointing out correlations but for "leaving the extension of the correlation open".

Animalium (PA), where Aristotle expresses some basic caveats about *defining* animal groups. By listing “fish”, “bird”, and “cetaceans”, Aristotle simply points out a basic fundament of HA, namely the description of animals by building larger groups either on the ground of existing animal categories or by forming them on the basis of some shared properties (e.g., viviparous, flying, four-legged). Aristotle’s aim, at least in *Historia Animalium*, was primarily a description of animal features and correlations among them (see also Lennox 2019). In terms of ADT, Aristotle’s work was aimed at $\mathcal{D}$-laws, that is, empirical universal laws, not so much about $\mathcal{A}$-laws, which involve exclusive and inclusive predicate changes in order to form a taxonomic tree with its principles of disjointness and inclusion of subgroups.

While the main work in HA is empirical, there is nevertheless also conceptual development. When Aristotle speaks about the viviparous quadrupeds, he does not explicitly define such a concept. However, his frequent reference to this group has nevertheless the effect of what I have called a definitional learning. Aristotle’s observations about this category prepared the concept for further development, which no longer depended on the initial properties of being viviparous and four-legged. After centuries of further revision, especially by Linnaeus (Linné, 1735, 1758), it developed into the modern concept of mammals. Nowadays, instances of this category are neither generally quadrupeds (whales) nor viviparous (platypus).

Linnaeus’s work on biological categories, including animals, was arguably much more based on analytic principles, namely the distinction of exclusive and mutually exhaustive categories and the inclusion of subcategories. His concept of quadrupdia in the first edition of *Systemae Naturae* builds on Aristotle’s category of the viviparous quadrupeds but *explicitly* includes humans. Aristotle clearly considered humans as animals and often discussed them together with the viviparous quadrupeds but he did not include them into a larger category. Linnaeus aimed at exclusive and mutually exhaustive groups (on each level). As a consequence any animal had to be included into a larger category. To not further classify an animal was not an available option for him. Aristotle was much more flexible in this regard because he mainly considered correlations between less fixed groups. At some points, cetaceans appear quite close to our modern group of mammals but he nevertheless applies the word “fish”. In comparison to Aristotle, Linnaeus forbids such constellations and thus looses some expressive flexibility. In his first edition of *Systemae Naturae*, Linnaeus decided to include cetaceans as subcategory of fish. This categorization underwent a major change in the tenth edition, when quadrupdia were renamed “mammals” and included cetaceans. As I discussed in Strößner (2020), this step was not mainly motivated by a change in the meaning of “fish” but by the development of the concept of mammals. The exclusion of cetaceans from fish was mainly owed to the background assumption that mammals and fish have to be distinct groups. An animal that has been grouped into the mammals category may no longer be regarded as fish. However, this has not led to a more scientific concept of fish in the long run. Later scientific taxonomies usually eliminate the concept of fish because it is a paraphyletic group. This corresponds to a conceptual elimination. Note that we are still able to reconstruct that category as an aquatic vertebrate that is not a mammal. However, since the members of this category are not more closely related to each other than to non-members, the concept is no longer important enough for biology to justify an atomic concept for this category.

The short example from zoology, I hope, helps to understand in which way predicate change may lead to local incommensurability. Zoology after Linnaeus is much richer in analytic laws. There are less conceivable objects and thus less permissible ways to describe animals compared to Aristotle's system and folk-taxonomies. This makes it hard to translate Aristotle to taxonomic systems after Linnaeus and provides an example of Kuhn's observation that ancient texts, if translated into the language of modern science, seem partially obscure.

Major steps in the conceptual development of the discussed animal categories, such as the definition of concepts, inclusions and exclusions of entities, and the generation of analytic laws can be reconstructed in the formal framework of predicate change and learning presented above. An advantage of the ADT approach and its closeness to epistemic logic is that it can integrate rules of belief revision and delineate them from conceptual change. For example, Aristotle claims that all animals with hair are viviparous. This claim was falsified after platypus become known to zoologists. This discovery, no matter how influential for our understanding of mammals, is best modelled as a change in the plausibility ordering. There is nothing conceptually obscure or surprising about Aristotle's claim. It just turned out to be wrong. Learning about platypus's reproduction habits brought about a major change in the study of mammals but this does not mean that it needs to be viewed as a conceptual one. In other words, not every drastic change in science is a conceptual one.

Finally, let me also use the example of biological taxonomy to shortly remark the limitations of the ADT-based analysis when it comes to larger and wide-spread scientific revolutions. One of the largest revisions of biology is certainly evolution theory. It gave rise to new systems of classifying biological groups, such as evolutionary taxonomy or phylogenetic nomenclature. This change involves the introduction of new criteria of similarity. In ADT, this would correspond to a major revision of the similarity ordering together with the interpretation function that has to capture these similarities. This might be a change too drastic to capture in a relatively simple logic model of conceptual change.

6 Conclusion and Outlook

This paper presented possible operations of conceptual learning in a DEL framework. The discussion rested on my prior work (Strößner 2020) and focused on the influence of conceptual change on translatability. I have demonstrated that inclusive and exclusive predicate change can lead to a loss of discriminatory power but never to an expansion. The introduction of new atomic predicates in terms of definitions was revealed to be a special case of an inclusive predicate change and is thus generally not applicable to cases of true conceptual expansions. Matters are only different if the learning is based on a prototype and similarity to a prototype. This prototype-based learning increases the discriminatory possibilities of the language and leads to new proper analytic and doxastic laws that were previously inexpressible. Finally, conceptual elimination is a special case of exclusive predicate change in Strößner (2020).

In which sense and to which degree do the changes considered here involve incommensurability or conceptual revolutions? First, they have consequences for a limited part of the language, namely the sentences in which the changed predicate is involved. I did not model changes with very wide-spread consequences on the conceptual system. Moreover, the discussed conceptual changes never deprive and enrich the language simultaneously. A common language that allows the expression of all the laws of both systems is guaranteed to exist: it is either the old one or the new one. In this sense, my approach emphasises aspects of continuity. Note, however, that even though the changes discussed here are barely revolutionary, a larger series of subsequent predicate changes might result in quite drastic revisions. Although the conceptual changes considered here are by no means revolutionary in themselves, they nevertheless fall under Kuhn's notion of local incommensurability.

Before concluding the article, let me reflect on its limitations. Admittedly, this short study hardly covers all or even most of the phenomena that characterise concepts, conceptual learning and category formation (be it in the individual development of agents or the collective history of a community). In particular, I did not cover cases of large-scaled revolutions as they have been addressed by many other formal approaches that are explicitly focused on drastic conceptual changes, such as Thagard (1992) or Gärdenfors and Zenker (2013). However, the advantage of my approach is that it explicates the results of minor conceptual revisions into a broader dynamic framework that includes also belief revision and changes of plausibility.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Alchourron CE, Gärdenfors P, Makinson D (1985) On the logic of theory change: partial meet contraction and revision functions. J Symb Logic 50(2):510–530. https://doi.org/10.2307/2274239

Andersen H, Nersessian NJ (2000) Nomic concepts, frames, and conceptual change. Philos Sci 67:224–241. https://doi.org/10.1086/392821

Baltag A, Renne B (2016) The stanford encyclopedia of philosophy metaphysics research lab. In: Zalta EN (ed) Dynamic epistemic logic. Stanford University, Stanford, CA

Barnes J (1984) Complete works of Aristotle: the revised Oxford translation. Princeton University Press, Princeton, NJ

Barsalou L (1992) Frames, Fields, and Contrasts. In: Lehrer A, Kittay EF (eds) Frames, concepts, and conceptual fields. Lawrence Erlbaum Associates, NewJersey, pp 21–74. https://doi.org/10.4324/9780203062821-5
Carnap R (1955) Meaning and synonymy in natural languages. Philos Stud 6(3):33–47. https://doi.org/10.2307/2268260
Carnap R (1962) Logical foundations of probability. University of Chicago Press, Chicago
Chalmers DJ (2012) Constructing the world. Oxford University Press, Oxford, UK
De Benedetto M (2020) "Explicating 'Explication' via conceptual spaces". In: Erkenntnis. https://doi.org/10.1007/s10670-020-00221-8 (online first)
Dempster AP (1967) Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 38(2):325–339. https://doi.org/10.1214/aoms/1177698950
Feyerabend P (1975) Against method. New Left Books, London, New York
Fleck L (1935) Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Benno Schwabe & Co, Basel
Fleck L (1979) Genesis and development of a scientific fact. University of Chicago Press, Chicago. https://doi.org/10.7208/chicago/9780226190341.001.0001
Grice P, Strawson P (1956) In defense of a dogma. Philos Rev 65(2):141–158. https://doi.org/10.2307/2182828
Gärdenfors P (2000) Conceptual spaces: the geometry of thought. The MIT Press, Cambridge, MA. https://doi.org/10.7551/mitpress/2076.001.0001
Gärdenfors P, Zenker F (2013) Theory change as dimensional change: conceptual spaces applied to the dynamics of empirical theories. Synthese 190(6):1039–1058. https://doi.org/10.1007/s11229-011-0060-0
Hansson SO (2017) Logic of Belief Revision. In: Zalta EN (ed) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab Stanford University, Stanford, CA
Huber F (2015) How to learn concepts, consequences, and conditionals. Analytica 1:20–36
Kornmesser S, Schurz G (2020) "Analyzing Theories in the Frame Model". In: Erkenntnis 85:1313–1346. https://doi.org/10.1007/s10670-018-0078-5
Kraus Sarit, Lehmann Daniel, Magidor Menachem (1990) Nonmonotonic reasoning, preferential models and cumulative logics. Artif Intell 44(1–2):167–207. https://doi.org/10.1016/0004-3702(90)90101-5
Kuhn T (1962) The structure of scientific revolutions. University of Chicago Press, Chicago
Kuhn T (1983) Comparability. communicability. commensurability. In Kuhn 2000:33–57. https://doi.org/10.1086/psaprocbienmeetp.1982.2.192452
Kuhn T (1989) Possible worlds in history of science. Kuhn 2000:58–89. https://doi.org/10.1515/9783110866858.9
Kuhn T (2000) The Road Since Structure. Conant J, Haugeland J (eds) University of Chicago Press, Chicago. https://doi.org/10.1007/978-1-349-25249-7_10
Leitgeb H, Carus A (2020) The stanford encyclopedia of philosophy metaphysics research lab. In: Zalta EN (ed) Rudolf Carnap. Stanford University, California
Lennox J (2019) The stanford encyclopedia of philosophy, metaphysics research lab. In: Zalta EN (ed) Aristotle's biology. Stanford University, California
Lewis D (1973) Counterfactuals. Harvard University Press, Cambridge, MA
von Lieven AF, Humar M (2008) A cladistic analysis of aristotle's animal groups in the "Historia animalium''. History Philos Life Sci 30(2):227–262
Linné C (1735) Systema Naturae 1. Biodiversity Heritage Library. https://www.biodiversitylibrary.org
Linné C (1758) Systema Naturae 10, pt. 1. Biodiversity Heritage Library. https://www.biodiversitylibrary.org
Minsky M (1975) A framework for representing knowledge. In: Henry P, Horn B (eds) The psychology of computer vision. McGraw-Hill, New York. https://doi.org/10.1016/b978-1-4832-1446-7.50018-2
O'Connor C (2019) The natural selection of conservative science. Stud History Philos Sci Part A 76:24–29. https://doi.org/10.1016/j.shpsa.2018.09.007
Quine WVO (1951) Two dogmas of empiricism. Philos Rev 60(1):20–43. https://doi.org/10.1007/978-94-010-1863-0_2
Romeijn JW (2005) Theory change and Bayesian statistical inference. Eng. In: Philos Sci 72(5):1174–1186. https://doi.org/10.1086/508963
Romero A, Keith EO (2012) New approaches to the study of marine mammals. In: Romero A, Keith EO (eds) When whales became mammals: the scientific journey of cetaceans from fish to mammals in the history of science. Intechopen, London

Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Priceton. https://doi.org/10.1515/9780691214696
Skulimowski AMJ (2011) Freedom of choice and creativity in multicriteria decision making. In: Theeramunkong T (ed) Knowledge, information, and creativity support systems. Springer, Berlin Heidelberg
Spohn W (2012) The Laws of Belief: Ranking Theory and Its Philosophical applications. Oxford University Press, Oxford, UK. https://doi.org/10.1093/acprof:oso/9780199697502.001.0001
Strößner C (2020) Predicate change. J Philos Logic 49(6):1159–1183. https://doi.org/10.1007/s10992-020-09552-x
Talbott W (2016) The stanford encyclopedia of philosophy metaphysics research lab. In: Zalta EN (ed) Bayesian epistemology. Stanford University, California
Thagard P (1992) Conceptual revolutions. Princeton University Press, Princeton. https://doi.org/10.1515/9780691186672
Thagard P (2012) The cognitive science of science: explanation, discovery, and conceptual change. The MIT Press, Cambridge, MA. https://doi.org/10.7551/mitpress/9218.001.0001
Tsiporkova E, Boeva V, De Baets B (1999) Dempster-Shafer theory framed in modal logic. Int J Approx Reason 21(2):157–175. https://doi.org/10.1016/S0888-613X(99)00011-0
van Benthem J (2004) Belief revision and dynamic logic. J Appl Non-Class Logics 14(2):1–25. https://doi.org/10.1007/978-3-319-06025-5_8
van Benthem J (2011) Logical dynamics of information and interaction. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/cbo9780511974533
van Ditmarsch H, van der Hoek W, Kooi B (2007) Dynamic epistemic logic. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5839-4
Williamson J (2003) Bayesianism and language change. J Logic, Lang Inf 12(1):53–97
Yu PL (1991) Habitual domains. Op Res 39(6):869–876