Kuhn on essentialism and the causal theory of reference.
Kuukkanen, J.M.

Citation

Version: Not Applicable (or Unknown)
License: [Leiden University Non-exclusive license](https://hdl.handle.net/1887/16553)
Downloaded from: https://hdl.handle.net/1887/16553

Note: To cite this publication please use the final published version (if applicable).
Kuhn on Essentialism and the Causal Theory of Reference*

Jouni-Matti Kuukkanen†‡

The causal theory of reference is often taken to provide a solution to the problems, such as incomparability and referential discontinuity, that the meaning-change thesis raised. I show that Kuhn successfully questioned the causal theory and Putnam’s idea that reference is determined via the sameness relation of essences that holds between a sample and other members of a kind in all possible worlds. Putnam’s single ‘essential’ properties may be necessary but not sufficient to determine membership in a kind category. Kuhn argued that extension is fixed by similarity-dissimilarity relations that are liable to change in taxonomic reorganizations of science.

1. Introduction. There has been an upsurge of interest in Kuhn in recent years. Kuhn is no longer necessarily taken as an irrationalist and a radical philosopher in the same mode as the later Feyerabend. As a consequence, the old view of Kuhn has been replaced by a whole diversity of views (e.g., Hoyningen-Huene 1993; Bird 2000; Fuller 2000; Andersen 2001a; Sharrock and Read 2002). And yet, few interpretations have tried to assess Kuhn’s significance for the debates on the causal theory of reference and essentialism.¹

This neglect is surprising for two reasons. First, Kuhn took the challenge of the causal theory seriously and attempted to formulate his re-

*Received October 2009; revised January 2010.
†To contact the author, please write to: Institute of Philosophy, University of Leiden, Martin de Vrieshof 4, 2300 RS Leiden, The Netherlands; e-mail: j.kuukkanen@phil .leidenuniv.nl.
‡This article has been under work for an unusually long time. While not all were, some of the reasons that prolonged the process were beyond my control. I wish to warmly thank all those who have kindly offered their comments on different versions of this paper or otherwise helped to improve the content over the years. I have especially the following persons in my mind: Alexander Bird, John Henry, Ranjan Chaudhuri, James W. McAllister, and a number of anonymous referees.

¹The exception is Sharrock and Read (2002) and also Read and Sharrock (2002).
response to it directly and indirectly on various occasions. Second, the causal theory of reference plays a central role in the debates on the rationality and progress of science. Often, it has been taken to provide a solution to the problems of meaning change, such as incomparability and referential discontinuity, that the writings of Kuhn and other historical philosophers raised.

Kuhn thought that the causal theory was “a breakthrough with respect to proper names,” but he rejected it as a general theory of reference applicable in science (2000, 312). However, Kuhn is generally not taken to have succeeded in producing something that deserves serious philosophical attention, especially with regard to more technical aspects, such as rigidity and essentialism. Suppe already stated bluntly that “despite his sustained efforts to reply to critics and clarify, modify, or improve his position . . . increasing numbers of philosophers of science reject Kuhn’s approach as irredeemably flawed, although not as hopeless as Feyerabend’s” (1977, 647–48). Although this comment applies to Kuhn’s early papers, the response to Kuhn’s attempts to specify his philosophical position did not change over the years. For example, Bird writes that it is very hard to see the relevance of Kuhn’s critique of Putnam, although the causal theory and essentialism have been under close scrutiny by analytic philosophers. Bird thinks that Kuhn had a “tin ear” for analytic philosophical arguments (2004b, 74–75). In actuality, before Sharrock and Read, who think that Kuhn gave strong reasons to doubt whether the causal theory is “co-tenable with a decent understanding of the history and philosophy of science” (2002, 152), on the one hand, and Andersen, Barker, and Chen (Andersen, Barker, and Chen 1996; Chen, Andersen, and Barker 1998; Barker, Chen, and Andersen 2003) on the other, it is hard to find any positive references in the philosophy of science literature to Kuhn’s theory of reference and concept formation, and moreover, it is practically impossible to find any at all in the philosophy of language literature still today.

I argue that Kuhn’s criticism of the causal theory contains elements worth serious consideration. Kuhn can be seen to contribute to many central topics in contemporary analytic philosophy, specifically, discussion

2. His view is presented directly in “Possible Worlds in History of Science” (reprinted in Kuhn 2000). Note that “Dubbing and Redubbing: The Vulnerability of Rigid Designation” (Kuhn 1990) is a reduced and revised version of an early draft of “Possible Worlds.” Also “Metaphor in Science” (reprinted in 2000) is relevant. Indirectly, he developed his theory of concepts over many years (see n. 6 for literature).

on rigid reference and essentialism, with direct relevance for philosophy of science. First, he offers a principled way to question whether natural kinds can be fixed by their essences. Further, he develops a unique naturalistic theory of reference fixing by similarity and dissimilarity sets. Finally, Kuhn calls into question the arguments for progress that rely on the Putnam-Kripke causal theory, or on some close modification of it, and that understand scientific development as consisting of improving approximations of the real joints of nature.

The causal theory was designed to fix a reference of a kind term without any description. Reference is determined directly by the hidden essence shared between the instances of the kind. As a consequence, theories are comparable via common references, and reference is not sensitive to theory changes. Further, the stability of reference provides continuity against which the progress of science can be measured, which can be understood as a gradual improvement in our understanding of the fundamental structure of the world, constituted by natural kinds to which our scientific terms refer. However, in the later Kuhn’s view, scientific revolutions, understood as reorganizations of scientific taxonomies, are central phenomena in the history of science. More precisely, Kuhn’s questioning of the causal theory, essentialism, and rigid reference leads to a view that the references of natural kind terms may change in (taxonomic) scientific revolutions, casting doubt on whether the history of science can be viewed as a gradual improvement and specification of the essences of the pre-structured set of natural kinds. According to Kuhn’s theory, references of kind terms are determined through a network of similarity and dissimilarity relations, which specify what properties an entity can and cannot have in order to qualify as that kind. Similarity-dissimilarity sets do not entail any principled distinction between essential and superficial properties. Although the original formation of sets does not require any description, the descriptive content stored in the sets is subsequently needed to keep the boundaries between kind categories sharp.

In this article, I first examine recent interpretations of Kuhn, seeking to pinpoint their strengths and inadequacies. An interesting point in Kuhn’s argumentation against the causal theory is that his theorizing is inclined toward naturalism. Second, it is notable that a number of theories in cognitive science show similarity to Kuhn’s ideas on the emergence and nature of kind categories. For this reason, I flesh out in what sense Kuhn can be seen as a naturalist. Third, I study in detail Kuhn’s reasons for opposing the causal theory of reference, attempting to show that there are many points in his writings that are relevant and interesting with respect to many central topics in both philosophy of science and analytic philosophy.
2. Four Images of Kuhn. Hoyningen-Huene’s image of Kuhn as a Kantian is the oldest and arguably most influential of the new interpretations, emerging after the rationality debate of the 1960s and 1970s. He argued that Kuhn implicitly distinguished two different kinds of worlds: the world in itself and the phenomenal world. The former is unknowable and purely “object sided,” while the latter is constituted by the object-sided world in itself and by subject-sided moments originating with an epistemic subject (Hoyningen-Huene 1993, 31–42; see also Devitt 1997, 72, 156–57). In addition, several monographs since 2000 and numerous articles on Kuhn over the last 20 years have appeared, many of which attempt to reassess Kuhn’s position in philosophy somehow. A further three images of Kuhn emerge from this literature.4

First, there is Kuhn the empiricist. This view maintains that Kuhn’s philosophy, in contrast to early interpretations, either has clear parallels with logical empiricism or may even be regarded as having inherited directly a number of central tenets from that tradition. One or both of these arguments can be found in Bird’s book (2000) and in some of his many papers on Kuhn (2002, 2004a) but also in an older paper by English (1978) and in articles by Reisch (1991), Earman (1993), Irzik and Grünberg (1995), and Friedman (2003). The most recent suggestion along these lines is Gattei (2008), who takes Kuhn’s philosophy as an anti-Popperian extension of logical empiricism. The connection between Kuhn and logical empiricism can be established in various ways. An interesting description is given by Bird. According to it, Kuhn retained many of the central doctrines of logical positivism, such as intensionalism (i.e., the idea that intension determines reference), meaning holism, and skepticism toward theoretical entities. In addition, he writes that partially because of these commitments, Kuhn ignored many of the central developments in contemporary analytic philosophy: externalism in epistemology and semantics including the causal theory.

Second, there is Kuhn the precursor of cognitive science. That is the view brought forward notably in the papers by Andersen, Barker, and Chen (Andersen et al. 1996; Chen et al. 1998; Barker et al. 2003) and also by Bird (2000, 2004c, 2005) in a slightly different sense. The former focus on Kuhn’s adoption of the family resemblance theory of concepts and concept learning as well as its application to philosophy of science. This theory, which questions the classical theory of concepts that defines a concept by a set of necessary and sufficient conditions, is alleged to be corroborated by empirical research in cognitive psychology.

Finally, we find also Kuhn the Wittgensteinian therapeutic philosopher.  

4. Steve Fuller (2000) is ignored here because his portrayal of Kuhn as a conservative in the political sense is not relevant in the current context.
That is the thesis advanced by Sharrock and Read (2002; see also Read and Sharrock 2002). They argue that Kuhn did not wish to develop any positive theory of science, but his misunderstood message was that philosophy of science has to leave everything (to science) as it is. As part of this argument, they also sought to revive Kuhn’s case against the Putnam-Kripke causal theory of reference.

Bird’s (2004c, 2005) aim of interpreting Kuhn in the naturalistic framework is plausible. There are indeed elements in Kuhn that justify viewing him as a naturalist. However, I do not think that Kuhn’s refusal to advocate the causal theory of reference is regrettable and a consequence of his embracing the non-naturalistic empiricist legacy. With regard to this, Read and Sharrock (2002; cf. Bird 2004b) raise an important point. They note that a “‘taxonomic conception’ of natural kinds,” or “systematicity” (Sharrock and Read 2002, 151, 155), provides an explanation of and a reason for Kuhn’s opposition. However, they also object to portraying Kuhn as a naturalist, which is unfortunate because linking systematicity with his naturalism gives us a better understanding of Kuhn’s rejection of the causal theory. In order to see this, we need to employ the ideas of Andersen et al. on how similarity and dissimilarity relations determine concepts. Their suggestion can be used to make Kuhn’s theory of extension determination explicit. Having said this, it is important to notice that Andersen’s (2001b) attempt to characterize Kuhn’s position as an intermediate position between realism and constructivism takes her close to the Kantian image of Kuhn.

3. Kuhn’s Naturalism. Kuhn’s philosophy can be linked to naturalism via his theory of concepts. An important influence here is Wittgenstein and his notion of family resemblance, which can already be seen in Kuhn (1970). Kuhn writes that the use of terms, as ‘chair’, ‘leaf’, and ‘game’, does not require grasping some set of attributes that these and only these possess: “Natural families [are ] . . . constituted by a network of overlapping and crisscross resemblances” (45). In the 1970s, Kuhn returns to discuss characterization of concepts and the notion of family resemblance on several occasions. Kuhn asks how a child who has not had any previous experience in recognizing and naming objects, or a student who

5. If I talk generally on their account, I will use Andersen et al. The presentation of their view is based on their three collective articles: Andersen et al. (1996), Chen et al. (1998), and Barker et al. (2003).

becomes a new member of a specialist scientific community, can come to master the concepts of the group. This process of learning involves a person who has already mastered the usage of the prevailing concepts and understands the differences between the instances of different concepts, who points to objects telling the learner under what category each of them falls. A learner may make mistakes and is corrected by the teacher until she or he learns to use the new concepts correctly. The process of concept learning results in the construction of family resemblance concepts, instances of which are bound by similarities to the instances of the same concept and by dissimilarities to the instances of other concepts. A natural family “is a class whose members resemble each other more closely than they resemble the members of other natural families” (Kuhn 1977, 285).

Kuhn assumes that ostensive learning does not require any previous understanding of the criteria of categorization and that people are able to attain the relevant category-fixing criteria directly from the world without conscious mediation by the mind. Learned similarity-dissimilarity relationships “are prior . . . to a list of criteria which . . . would enable us to define our terms. . . . They are parts of a language-conditioned or language-correlated way of seeing the world. Until we have acquired them, we do not see a world at all” (2000, 171). This implies that the process is automatic and takes place prelinguistically. Kuhn does indeed refer to “neural processing,” which can be programmed differently and which preconditions the practice of object clustering (2000, 197 n. 14). The similarity sets produced in concept learning count as knowledge acquired of the world, which is subsequently stored in the mind. Later, he begins to talk of conceptual schemes, lexical taxonomies, and lexicons or mental modules that all appear to have the same function: they taxonomically store empirical information that is acquired of the world.7

Kuhn’s interest in information acquisition and its organization in the mind is remarkable because it is something that already for decades has been an object of research in cognitive science and psychology and, specifically, in a field that has become known as “knowledge representation.” A classic theory is Collins and Quillian’s (1969) suggestion that information is stored in a hierarchically organized semantic network composed of nodes representing concepts and attribute descriptions associated with each concept. This model has faced difficulties since, but the assumption that information organization, or knowledge representation, is best described in terms of multiple interconnected assumptions and relations

7. Kuhn (1957) already says that the “conceptual scheme” is a framework for the organization of knowledge, guiding future research, and helping to recall data (37–41).
appears to be widely accepted and is a result of continuous empirical testing and modeling. The most serious problem with the semantic network theory is that it does not take into account that some instances of concept are, in fact, taken to be more typical than others. Eleanor Rosch showed, for example, how the Dani tribe took even a particular color as the best example of the color concept. Since then, there have been numerous studies with tools, clothing, furniture, animals, trees, fish, and birds that all come to the same conclusion: some members of a category are considered better, more typical examples than others. Human concepts show thus graded structures, varying on the perceived goodness of example (Andersen et al. 1996, 352). The lesson appeared to be twofold. First, it was necessary to reconsider how information stored in the mind can be represented. The semantic network model does not allow variation. Being a member of a category is an all-or-nothing affair: an instance either fulfills membership criteria or not. Second, if an instance of a concept cannot be characterized by the same set of attributes, we need to consider, in light of these empirical results, how a concept could be characterized. Interestingly, Rosch explicitly called into question the classic account of concept definition by a set of necessary and sufficient conditions and proposed Wittgenstein’s theory of concepts as an alternative to it. She reasons that, because all members of a category just do not share any set of exactly the same features, the features determining membership are better described as a large set over which individual instances overlap but do not share completely, that is, as a family resemblance concept. That is also Kuhn’s view, as was already shown above.

With regard to knowledge representation, one of the most successful ideas for improvement was Marvin Minsky’s (1980) suggestion that information is stored in frames that have an internal structure, which contains a cluster of “knowledge” associated with a concept. A frame can be characterized as a collection of slots and slot fillers that describe a stereotypical item. Alternatively, we may say that a frame is composed of attributes and their values. It may incorporate a taxonomic structure because of the fact that superkind and subkind can be taken as attributes, and values can be taken as attributes of further values. For example, in

8. On Collins and Quillian’s semantic network and a philosophical introduction to knowledge representation, see Stevenson (1993, esp. chap. 2). An examination of the similarities between their illustration of the memory structure of a conceptual hierarchy and Kuhn’s description of it would constitute a fruitful subject for a more detailed study not attempted here.

9. See Andersen et al. (1996) on references to Rosch. See also Lakoff (1987, 41–42) on various experiments that were used to test and substantiate prototype effects.
the frame of dog, an attribute subkind can have different breeds of dogs as values that then may be subdivided further. The stipulation that a frame describes a stereotypical item means that a description does not function as a definition as in the classical account of concepts; that is, a description is not necessarily true of all instances that are in the extension of a concept. All that can be said is that typically such objects have the features postulated in the frame, and a frame thus induces expectations of what kinds of things objects are.10

At this point, it is worth noting what Kuhn said of lexical structures: “Imagine, for a moment, that for each individual a referring term is a node in a lexical network from which radiate labels for the criteria that he or she uses in identifying the referents of the nodal term. Those criteria will tie some terms together and distance them from others, thus building a multidimensional structure within the lexicon” (2000, 52). The description sounds very much like those of the semantic network or the frame model, if we remember that they represent information by the attribute-value structure and understand their taxonomy-building feature.

Andersen, Barker, and Chen have formulated a theory that incorporates what Kuhn says about learning and determining concepts with the frame model as further developed by Lawrence Barsalou (1992). Furthermore, they invoke Rosch’s studies, and their fundamental conclusion is that the frame model supports the existence of graded structures. Because of its flexibility, the frame model can accommodate variable representations of instances in one and the same frame. Andersen et al. argue that Kuhn’s theory of concepts is “massively supported by empirical work in cognitive psychology” (1996, 359). Yet, Barker et al.’s most intriguing claim is that cognitive psychology has provided an “empirical vindication of the family resemblance account” and that empirical findings have shown that “necessary-and-sufficient condition definition of concepts will be impossible” (Barker et al. 2003, 218, 219).

We should ask whether it is correct to claim that definition by necessary and sufficient conditions “will be impossible.” I think this claim is too strong. It is undeniable that there is a lot of empirical evidence for the view according to which “human concepts,” such as concepts of trees, birds, and furniture, cannot be defined by a set of necessary and sufficient conditions. Andersen et al. are able to cite a long, but undoubtedly not exhaustive, list of studies that advance this point (e.g., Andersen et al. 1996, 352). However, the past failure of the classical account does not mean that it will also fail in the future. The lack of progress of attempts to find the defining features of a concept thus far does not yet mean that

there are no defining features—only that we have not yet been able to find them (cf. Smith and Medin 1981, 30–31). Indeed, Kuhn remarked that some concepts seem to possess a definitional core and appear in and are defined by exceptionless laws of nature, such as ‘force’ defined by Newton’s laws of motion (1993, 316–17). Kuhn argues that these ‘nomic’ concepts cannot be learned by contrast to other related concepts, but in complex problem situations, because there are no objects that could be pointed at by ostension. 11

What then is Kuhn’s naturalism? The process that produces the similarity and dissimilarity sets can be compared to gathering empirical knowledge from one’s surroundings and results in concepts that are qualitatively similar to such knowledge. Kuhn’s theory of concept acquisition and reference fixing are thus naturalistic in the sense that they do not imply existence of non-natural concepts and need for a nonscientific method. Although Kuhn did not have an explicit aspiration to be a naturalist, his philosophy fits well the naturalistic idea of philosophy and science. 12 And significantly, Kuhn’s theory of concept learning receives empirical support both with regard to how the end result, that is, the information acquired, is stored and represented in the mind and how the concepts that emerge as a result of this process are understood. In brief, it is empirically grounded.

4. Reference of Natural Kind Terms. Kuhn formulated his critique of the causal theory of reference as represented by Putnam explicitly in two articles (1990; 2000, 58–90). 13 The first thing to notice is that Kuhn was not totally dismissive of the causal theory, but he endorsed it within a limited range of application. Kuhn says that by denying the descriptive determination of reference of proper names, the causal theory has brought “a great advantage” (1990, 309; 2000, 198, 312–13). The problems arise if the same theory is extended also to natural kind terms: “When one

11. Andersen and Nersessian (2000) have attempted to take this approach further and argued that family resemblance nevertheless can be taken to play a major role in both learning and representation of nomic concepts.

12. In addition to Kuhn theorizing on concepts and reference, his early view on the history of science was based on empirical historical research (cf. Bird 2000; Sharrock and Read 2002; Kuukkanen 2009). However, in one of his last papers, Kuhn surprisingly expressed his desire to use a priori or pure philosophical reasoning to reach the same conclusions (2000, 95, 112, 115). He was referring to what might be called naïve empiricism in history, which he and others had used to refute the foundationalist image of science that derived from logical empiricism. The later Kuhn saw that history does not necessarily yield such an unequivocal judgment.

13. In addition to Putnam (1975), Kripke (1980) is naturally another classic in the development of the causal theory of reference. However, Kuhn’s focus was on Putnam’s theory, and because our focus is on Kuhn, we will not examine here Kripke’s thinking.
makes the transition from proper names to the names of natural kinds, one loses access to the career line or lifeline which, in the case of proper names, enables one to check the correctness of different applications of the same term. The individuals which constitute natural families do have lifelines, but the natural family itself does not” (2000, 199).

Kuhn gives us an example of a case in which the causal theory works and of another in which it does not. A single act of ostension is enough to fix reference to ‘Richard Boyd’. The situation is entirely different when one is presented the deflected needle of a galvanometer and simultaneously told that the cause of the deflection was ‘electric charge’. Kuhn claims that with natural kinds “a number of acts of ostension are required.” Further, in regard of natural kind terms (e.g., ‘swan’, ‘goose’, ‘electric charge’, etc.), an establishment of reference requires exposure not only to “varied members of that kind but also to members of others—to individuals, that is, to which the term might otherwise have been mistakenly applied” (2000, 200).

What does Kuhn’s critique boil down to? At first sight, it may look as if it is similar to the so-called Qua problem, that is, the problem that reference cannot be picked out without a descriptive specification of the kind of object that is being referred to because the number of the kinds of objects that may be picked out is practically unlimited (e.g., Sterelny 1983, 121; Devitt and Sterelny 1999, 79–81, 90–93). But this is not the crux of the argument, as Kuhn is clearly not disputing the ostensive determination of reference. What he wants to point out is that a single act of ostension is not enough in such cases; reference fixing requires multiple acts of ostension, not only to members of the kind whose reference is being fixed but also to members of other kinds. As we saw in the previous section, Kuhn believed that direct contact with objects and ostensive learning results in mental categories that classify objects according to their similarities and dissimilarities, that is, in accordance with his neural-processing model.

In general, the problems with ostensive reference fixing are well known, and most discussants have suggested that some kind of descriptive component is required (e.g., Devitt and Sterelny 1999, 96–101; Stanford and Kitcher 2000). Kuhn stands out because his idea of reference baptizing arises out of his naturalistic theories of concept learning. He criticizes the ostensive determination of reference, without suggesting that we need a descriptive theory of reference fixing. His theory could be said to be a kind of naturalistic model, wherein the direct and preconscious adaptation of the mind to a large number of external outputs leads to success in recognizing and categorizing objects and produces sets of similarity and dissimilarity relations or categories in the mind of a person.

The quotation above suggests that Kuhn is concerned not only with
baptizing but also with the question of what the reference of general terms could be. As he pointed out, there is no problem with proper names, as there is normally some individual to which a name refers, but in the case of general terms, there is no such individual. Kuhn writes that “the techniques of dubbing” and “of tracing lifelines” allow us to track astronomical individuals, such as the Earth, Moon, Mars, and Venus, through episodes of theory change. There is no change in the “lifelines” of individuals in transition from the heliocentric to the geocentric theory. ‘Mars’, for example, still refers to the same individual, that is, Mars. But there was a change in natural families because, for instance, ‘the moon’ and ‘the sun’ did belong to the family of ‘the planets’, but ‘the Earth’ did not, and so on (Kuhn 2000, 205). He makes his case forcefully in his last interview: “Look, you can trace the individual planets, Mars, heavenly bodies through the Copernican revolution—what you can’t trace through it is ‘planets.’ Planets are just a different collection before and afterward. There was a sort of localized break that fitted very closely. And now it turns out that some people, to an extent that surprises me and others, simply say, ‘In the Ptolemaic systems planets go around the Earth and in the Copernican system they go around the sun.’ But that's an incoherent statement! . . . I do not think it [the causal theory] works for common nouns” (2000, 312–13).

There are two things to point out from the passage above. First, Kuhn argues that if we cannot fix reference to an individual, we seem to lose referential stability. Second, and more important, Kuhn implicitly suggests what the reference of general terms is. It is the extension of a term, that is, the set or class of individuals that are in the extension. This was already suggested above by his ideas of how concepts are learned or how extension of kind terms is fixed (see also 2000, 82, 84 n. 30, 85). This reading makes Kuhn’s exclamation above intelligible. ‘Planet’ refers to different sets of individuals in the Ptolemaic system and in the Copernican system; that is, there was a change in the extension of ‘planet’ in the transition to the Copernican system. From this perspective, Kuhn is absolutely right to insist that to say without any qualification, “In the Ptolemaic systems planets go around the Earth and in the Copernican system they go around the sun,” is absurd. For example, with the Ptolemaic reading of ‘planet’, it includes the individual called ‘sun’. Then the claim implies that it goes around another individual called ‘Earth’. But how could one then also claim that ‘planets’ also traverse around the sun, specifically because that individual is taken to be in the extension of ‘planet’, without recognizing the change in the extension (cf. Kuhn 2000, 94)?

Kuhn’s understanding of meaning change as change in extension, even after becoming familiar with the causal theory, seems justified if general terms refer to individuals or classes of them. In any case, if the reference
is not an individual or sets and classes of them, it is not at least immediately clear what it is. For Kuhn, this commitment became increasingly important in his middle and later career, as change in extension came to be central phenomena in scientific revolutions: “That sort of redistribution of individuals among natural families or kinds, . . . is, I now feel, a central (perhaps the central) feature of the episodes I have previously labelled scientific revolutions” (2000, 205).¹⁴

Having said this, it is clear that to take general terms to designate their extensions is by no means the only possibility. Further, while it is recognized that there is no obvious candidate for the reference of common nouns, it also is understood what the consequence of the commitment to extension is: terms would refer nonrigidly (e.g., Laporte 2000, 294; 2007; Schwartz 2002; Devitt 2005, 140; Haukioja 2006, 156). If ‘tiger’ refers to all the tigers in this world, it does not follow that it refers to the same set in all possible worlds. Quite clearly, some tigers might not have existed, and some others could have come into existence. A further consequence is that we would seem to lose the prospect of finding interesting, necessary, a posteriori truths by way of theoretical identifications, such as ‘Water = H₂O’. That is because both designators in an identity statement have to refer rigidly in order to express a necessary truth (e.g., Laporte 2007; cf. Soames 2002, chap. 10).

A popular response is to say that natural kind terms refer to some kind of abstract or universal objects (Donellan 1983; Laporte 2000). ‘Tiger’ would refer to the abstract kind tiger in all possible worlds and ‘water’ similarly to the abstract kind water in all possible worlds. The problem with this suggestion is that it threatens to trivialize rigidity because we could construct kinds as we wish. ‘Bachelor’ would refer to the kind bachelor, and so on. As a consequence, rigidity could not be used to distinguish between natural and non-natural kind terms. For some, this is enough to make kind rigidity unacceptable (Schwartz 2002; Soames 2002, 249–50, 260–61; Devitt 2005), while some continue to pursue this line of argument (Laporte 2000). It also is worth noting that the proposal that explains rigidity by taking general terms to designate abstract kinds appears to commit one to a substantive metaphysical position, that is, to rejecting nominalism (Salmon 1982, 43, 53; 2005, 133–34; Devitt 2005; 2007).

¹⁴. The later Kuhn associated the concept of incommensurability with the kind of taxonomic reorganization and with the speciation of scientific fields that lead to the violations of the no-overlap principle between the old and the new taxonomy. The principle says that no two kind terms may overlap in their references, unless they are related as species to genus. Further, interestingly, Kuhn thought that incommensurability is a precondition, not an impediment, for scientific progress (e.g., 2000, 99).
556 JOUNI-MATTI KUUKKANEN

Laporte 2007; see also Macbeth 1995). This observation may offer us a further explanation of Kuhn’s position on the matter (see sec. 7).

It is unlikely that Kuhn spent much time ruminating on rigidity, although in one of his papers (1990) it appears even in the title. In this paper, Kuhn suggests that scientific development involves interrelated acts of redubbing sets of scientific terms. The reorganization of heavenly bodies in the Copernican revolution could be a case in point. Kuhn goes on to claim that between the acts of redubbing, dubbing results in rigid designation, a view positively commented on by Laporte (Kuhn 1990, 298; Laporte 2000, 309). However, it is difficult to square the idea that general terms designate their extensions with rigid designation. It is not clear how we could preserve the set of individuals across all possible worlds, even when the taxonomic system remains invariant. And Kuhn nowhere even hints of the possibility of referring to abstract entities or universals.

5. Sameness Relation and Essentialism. Putnam suggested that rigid designation may be secured via a sameness relation between samples of the kind. Salmon has usefully reconstructed Putnam’s theory of natural kind terms. I reproduce part of it here (Salmon’s T3 and T5):

(T1). One may “explain the meaning” of a natural kind term such as ‘water’ in an ostensive definition that has the intended logical force as follows:
For every possible world w, and every individual x in w, x is water in w if and only if x in w is the same liquid as this in the actual world where the demonstrative ‘this’ rigidly denotes a sample of water in the actual world.

(T2). Natural kind terms such as ‘water’ are rigid designators of natural kinds, as a consequence of (T1). (Salmon 1982, 97)

The thesis T1 and Putnam’s philosophy offer one explanation why natural kind terms would refer rigidly. All samples of a kind have to bear a sameness relation to the actual sample baptized and referred to. The crucial question is therefore what an appropriate sameness relation is. What kind of sameness relation between samples would make them fall under the kind water? There has to be something that guarantees membership in that category. Further, a failure to have it would disqualify a sample from that category. Putnam refers here to “important physical properties,” “microstructure,” “nature,” “hidden structure,” or most tellingly “essence” (1975, 232–35). And, in the case of water, the obvious candidate is the chemical structure of H2O. As Putnam puts it, “Once we have discovered the nature of water, nothing counts as a possible world in which water doesn’t have that nature. Once we have discovered that
water (in the actual world) is \( \text{H}_2\text{O} \), *nothing counts as a possible world in which water isn’t H\(_2\)O* (233).

The suggestion above is thus that water is \( \text{H}_2\text{O} \), and it is necessary for a sample to be \( \text{H}_2\text{O} \) in order to be water; \( \text{H}_2\text{O} \) is the essence of water, and in general, essence is something that a substance cannot fail to have and still be that substance. Therefore, if water is \( \text{H}_2\text{O} \), it is necessarily so. However, in order for this to work, we must be satisfied that we have found an essence of substance. In other words, we need to prove the theoretical identification statement ‘Water = \( \text{H}_2\text{O} \)’. This is where Kuhn enters the picture. Kuhn writes that if only one single essential property is required by each natural kind, then the causal theory has promise because we may be able to establish an identity relation between each substance and its essential property. And in fact, Kuhn admits that if ‘gold’ can be “paired” with ‘having atomic number 79’ and ‘water’ with ‘being \( \text{H}_2\text{O} \)’, then there is no problem (Kuhn 2000, 82). According to Kuhn, the problem is that, in general, this strategy does not work.

Kuhn questions the theoretical identity statement ‘Water is \( \text{H}_2\text{O} \)’. He remarks that water is not identical with \( \text{H}_2\text{O} \), nor is the reference of ‘water’ \( \text{H}_2\text{O} \). Kuhn says that \( \text{H}_2\text{O} \) can be in three states of aggregation: solid, liquid, and gaseous. Consequently, \( \text{H}_2\text{O} \) picks out not only water but also ice and steam. Therefore, the reference of ‘water’ cannot be equated with that of ‘\( \text{H}_2\text{O} \)’ because the latter covers a much wider domain than we want (i.e., liquid, ice, and steam). According to Kuhn, a more proper description would be that ‘water’ refers to liquid \( \text{H}_2\text{O} \) or close-packed \( \text{H}_2\text{O} \) particles in relatively rapid motion.

Bird suggests that we may not need to claim as much as Kuhn implies. Namely, we do not need to draw the conclusion from Putnam’s story that water and \( \text{H}_2\text{O} \) are identical, only that in all possible worlds water consists of \( \text{H}_2\text{O} \) (Bird 2000, 183). This sounds a promising way to go, as attempts to prove that water is literally identical with \( \text{H}_2\text{O} \) are hampered by great difficulties (see Johnston 1997, 564–73). Similar to Bird’s, Johnston’s preferred understanding is that the identity statement should be taken as saying that water is constituted by \( \text{H}_2\text{O} \), which can be used to explain its manifest properties (582). It is indeed important to consider how the alleged theoretical identification should be interpreted. Salmon suggested that “is water” in T1 is not the ‘is’ of identity but the ‘is’ of predication (1982, 99). How should we then read the statement ‘Water = \( \text{H}_2\text{O} \)’? A good suggestion is that it be interpreted as a conditional in the following way: if \( x \) is water, then \( x \) is \( \text{H}_2\text{O} \). Further, it may be even better to use a universal quantification: \( \forall x \) (\( x \) is water \( \supset \) \( x \) is \( \text{H}_2\text{O} \)), or in general, \( \forall x \) (Ax \( \supset \) Bx). Alternatively, the sentence can be interpreted as a universally quantified biconditional in a similar fashion (see Soames 2002, 250–59).

Putnam’s essentialism thus requires that, if a sample is water, it also is
H₂O. It is important to notice that, although Kuhn criticizes essentialism, he does not seem to question this claim. This becomes apparent from Kuhn’s disagreement with Putnam’s Twin Earth story. Kuhn cannot accept that visitors to Twin Earth would report back, “On Twin Earth, the word ‘water’ means XYZ.” The reason is that modern chemistry does not accept a substance like XYZ, which has the same superficial properties as the stuff we call ‘water’ but nonetheless is a totally different chemical structure from our water. For example, it would be too heavy to evaporate at normal terrestrial temperatures. Therefore, the visitors would report, “Back to the drawing board! Something is badly wrong with chemical theory” (Kuhn 2000, 80). Quite clearly, Kuhn thinks that water has to be H₂O, although the reason for this belief does not derive from any kind of metaphysical necessity but from the fact that the lexicon of modern chemistry does not allow any other kind of description (cf. Read and Sharrock 2002, 154).

Jackson writes that we resist saying that the waterlike stuff on Twin Earth is water because it does not have an important property that we associate with water. We have never come across waterlike stuff XYZ. Just like Kuhn, he says that this shows that the chemical structure H₂O is one important property in the scientific lexicon that we associate with ‘water’. It is part of the description associated with ‘water’ used to determine its extension. Further, similarly to Kuhn, Jackson says that if we were acquainted with XYZ, this discovery would lead to a more or less radical scientific readjustment but not to the refutation of the idea that natural kinds objects are associated with and individuated by their assumed properties (1998, 213).

However, essentialism does not only say that having the essential property is necessary for something being a certain substance. It also says that having it is sufficient for being that substance (cf. Soames 2002, 15). A substance or an entity can lack any superficial property or stereotype that is associated with a kind and still be the kind: a stripeless tiger is tiger, if it has a certain genetic structure; a sample of stuff is water even if it is pink, provided it is H₂O, and so on. However, the very possibility of finding such a sufficient essential property is what Kuhn questions. In other words, Kuhn wonders whether it is sufficient for x to be water if x is H₂O.

Kuhn’s argument is that if something is to qualify as water, it needs to have some further qualities, such as ‘liquidity’. That can be seen if you consider the following situation. We may say that ‘water consists of or is H₂O’ but also that ‘steam consists of or is H₂O’ and further that ‘ice consists of or is H₂O’. If this is all the information we have, we are left
wondering what on earth the difference between these is. Why do we have three terms instead of one? Therefore, it seems that, if x is H2O, it is not yet enough for its being water because it could be steam or ice as well. Bird correctly reminds us that the causal theory is a thesis about words, not a metaphysical thesis about essences (2004b, 63). Indeed, ‘water’, ‘steam’, and ‘ice’ are all ordinary words, and the differences in their usage require an explanation, even if their references share an essence. Of course, we can stretch the boundaries of language, but it seems rather difficult to deny that the use of these terms is crucially related to a distinction between liquid, gaseous, and solid substances (cf. Johnston 1997; Soames 2002, 294–96).

Interestingly, Putnam’s sameness relation, which decides whether two or more samples are water, is specifically the same liquid relation (1975, 232). This implies that in addition to being H2O, the stuff has to have another property in order to be water. It is not enough that the stuff is H2O; it also has to be liquid, which is (surprisingly) a similar position to that of Kuhn. Kuhn says that when two or more properties are needed, such as liquidity and H2O, then each predicate denotes a larger class than the conjunction does. Naming properties that limit the class becomes crucial: “For if two properties are required, why not three or four? Are we not back to the standard set of problems that causal theory was intended to resolve: which properties are essential, which accidental; which properties belong to a kind by definition, which are only contingent? Has the transition to a developed scientific vocabulary really helped at all? I think it has not” (Kuhn 2000, 82–83).

Kuhn claims that, except perhaps in rare circumstances, we need to identify more than one property that an instance must possess in order to be a member of the kind category. Kuhn thinks, as we have seen, that categories are formed by recognizing similarities and dissimilarities between instances. Samples of ‘water’ may be thought to be both H2O and liquid. Liquidity is needed to separate water, for instance, from steam and ice. We need to make a difference by naming some properties that samples in each category have and use in this task, such properties as liquid, gaseous, and solid; perhaps also some other properties, such as boiling and freezing points, optical wavelengths, and so on, can be used. Further, it is not self-evident which ones to choose. Kuhn writes, “The so-called superficial properties are no less necessary than their apparently essential successors. To say that water is liquid H2O is to locate it within an elaborate lexical and theoretical system. . . . If water is liquid H2O, then these

15. Johnston has pointed out that the statement ‘water = H2O’ takes us easily to absurd conclusions. If that statement is acceptable, then so is ‘ice = H2O’. But then we should conclude that ‘water = ice’ (1997; cf. Soames 2002, chap. 11).
properties are necessary to it. If they were not realized in practice that would be a reason to doubt that water really was H₂O” (2000, 83). In other words, the so-called superficial properties have an indispensable role to play in the prevailing lexical system, in separating instances of kinds from each other and assigning expectations to substances.

6. Kuhn’s Descriptivism. It is obvious that Kuhn denies the validity of the direct theory of reference as a general theory of reference.¹⁶ He suggests that it may work with proper names, and even with some exceptional natural kind terms, but notes that it does not work with scientific terms in general. Although Kuhn does not require the application of description in baptizing, because direct contact with the members of a natural kind gives rise to a set of kind discriminating predicates, the similarity and dissimilarity relations are subsequently used to determine and adjust reference. Underlying Kuhn’s critique of the causal theory is, hence, his own suggestion that conceptual categories are formed by observing similarities and dissimilarities between objects, which results in “natural families,” that is, family resemblance concepts. This theory is something that Sharrock and Read seemed to be looking for in their allusion to “systematicity,” the functioning of which is characterized by Andersen et al. with the help of recent theories in cognitive science.

Because Kuhn’s theory specifies not only what properties a kind can have but also what it cannot have, it is a rather peculiar sort of descriptivism (cf. Bird 2000, 163–68). As we saw in section 3, the set of properties that characterize a stereotypical member of a kind is not a definition of the kind. This means that a description does not have to be fully satisfied in order for a term to refer. The function of a description is to determine to what category/kind a particular individual belongs. Therefore, the objection that description is synonymous with reference does not have a bite against Kuhn’s theory. The essential claim is that reference is determined by a certain mechanism, which is the possession (and exclusion) of associated properties up to a satisfactory degree (see Jackson 1998).

Kuhn’s account may deliver at least a partial answer to the so-called objection from error, that is, the claim that people may have most of the relevant properties wrong and yet still manage to refer to an object by a term (see Jackson 1998). Because Kuhn’s set characterizes what a member of the kind typically is and is not, it is indeed possible that one is wrong about most of the associated properties. Something in the description has to be satisfied, and moreover, there are properties that an entity cannot

¹⁶. The direct theory of reference says that terms refer directly to their references, not via any form of description. The Putnam-Kripke causal theory of reference is a kind of direct theory of reference.
have, but it is less important which ones. The only requirement for this theory is that it is successful in the categorizing of kinds so that their boundaries are sharp. If the latter is achieved by an erroneous set of properties, it does not matter. That Kuhn’s theory is able to limit kind categories is shown by Andersen (2000). Kuhn’s reliance on dissimilarity or the contrasting sets of others kinds, and not only to similarity, manages to limit what falls in a kind category.

A further consequence of Kuhn’s descriptivism is that it does not matter whether people associate different sets of properties with a term. Kuhn explicitly denies that the sameness of description is a requirement for successful reference. Homogenous categorizing of objects in a community does not require that the criteria or “a set of learned expectations about the similarities and differences between the objects” that picks out objects in the world are the same. Only the taxonomic structure needs to be the same (Kuhn 2000, 239, 242).

7. Concluding Remarks. I have attempted to show that Kuhn’s work has many interesting connections with contemporary debates in analytic philosophy, with direct relevance for philosophy of science. Kuhn took, implicitly or explicitly, a stand on ostensive baptizing, the designation of general terms, rigidity, and essentialism. Perhaps the most interesting of these is his argument, contra Putnam, that the establishment of an essence by way of theoretical identifications is not sufficient to determine the extension of a kind. Also Kuhn’s theory of reference, in which extension is determined by similarity and dissimilarity relations, is a notable contribution.

The debates on Kuhn’s philosophy have triggered strong rhetorical reactions from many commentators (see sec. 1). Some have argued (e.g., Hoyningen-Huene, Oberheim, and Andersen 1996; Hoyningen-Huene and Oberheim 2009) that meta-incommensurability between the realist and the nonrealist makes rational dialogue extremely difficult between the participants on the different sides of the dispute. It is a valuable observation that the debate is preconditioned by radically different metaphysical assumptions and argumentative intentions, the settlement of which is not easy through philosophical discourse. Yet I would hesitate to claim that the disagreements derive from incommunicable differences in the meanings of some basic terms. It strikes as being futile to insist that no rational argument can have compelling force or that the arguments are necessarily circular, as this position really becomes an “argument stopper” (Sankey 2009, 212). My view is that the undisputable tendency to argue within and from one’s own framework does not make rational and fruitful philosophical debate impossible. It only means that one needs to show heightened sensitivity and attention to the details from which the disagreements
are expected to arise. Indeed, I hope to have demonstrated that even if Kuhn’s arguments were tentative and in need of reconstruction, all the themes mentioned above place him in a fruitful dialogue with contemporary philosophers in the analytic tradition.

Kuhn’s philosophy challenges the view according to which the history of science converges on the true description of the real joints of nature. According to Kuhn, the history of science comprises discontinuity in the form of taxonomic reorganizations of kinds, and “there is no basis for singling out one or the other juxtaposed set as the more real” (2000, 249). In this article, I have explicated the philosophical grounding of this argument and shown that it can be reached on a principled basis. An interesting further question is whether there is a more general framework behind this line of reasoning. There are two or three fruitful suggestions to explain Kuhn’s antiessentialism and commitment to individuals. Kuhn can be taken as a metaphysical nominalist (Hacking 1993), an epistemological nominalist (Bird 2003), or a Kantian (Hoyningen-Huene 1993; see also Andersen 2001b). Perhaps he presupposed that universals, including essences, do not literally exist and cannot be taken to underlie any actual or potential kind category. Or maybe he only assumed that they are, in the world in itself, veiled from us, which forces us to study and categorize observed similarities and dissimilarities in the phenomenal world. This is a topic that cannot be taken further here but opens an interesting theme for future research.

REFERENCES


