Habits and Interpretation: defending the pragmatist maxim

Christopher Hookway

1. Strategies for proving the pragmatist maxim

Peirce's pragmatic maxim was introduced as a methodological tool for clarifying the content of what in later years he referred to as 'intellectual concepts' (EP 2:401). By enabling us to obtain explicit formulations of the contents of such concepts, it provides information that we can use in carrying out inquiries into hypotheses containing those concepts. It also enables us to recognize when expressions that purport to express coherent intellectual concepts fail to do so. We can thus abandon a priori 'ontological metaphysics' as either 'meaningless gibberish' or as 'downright absurd' (EP2: 338); and we also obtain guidance in how to carry out scientific inquiries efficiently. The core idea behind Peirce's maxim is that we can make the content of a concept fully explicit by identifying the effects that the objects of the concept have 'which might conceivably have practical bearings' (EP1: 132) As Peirce put it in 1878, 'our conception of these effects is the whole of our conception of the object.' (EP1: 132, Hookway 2004)

Why should we accept the pragmatist maxim? As was explained in Hookway 2005, Peirce claimed that it was a merit of *his* version of pragmatism that its correctness could be proved and, from 1902 onward, he made a series of attempts to provide this proof. In earlier papers I have identified the strategies of proof that he employed in 1903 and in writings around 1905 and 1906. The current paper is concerned with a series of related and overlapping manuscripts dating from 1907 in which he presented what

appears to be an alternative strategy for proving the correctness of pragmatism. Although, in 1903, Peirce made bold claims for just how 'mathematical' and rigorous his proof would be, we should probably characterize his goal in less loaded terms. His aim is to provide an explanation of the correctness of the pragmatist maxim making use of a theoretical framework whose acceptance does not itself depend upon the correctness of pragmatism. This theoretical framework is provided by his semeiotic or logic, his general theory of signs and their understanding or interpretation and their role in cognition. The correctness of the pragmatist maxim is thus to be understood as a theorem that can be established within Peirce's theory of signs or meaning.

How can such a proof be constructed, and what is distinctive about the strategy of proof that Pierce used in the 1907 manuscripts? In order to answer this question we should compare the 1907 strategy with another that was used in the 'Lectures on Pragmatism' than he delivered at Harvard in 1903. The earlier strategy exploited a supposedly exhaustive classification of all the sorts of arguments that are employed in scientific inquiry, together with the claim that *all* cognitive activity involved the use of arguments of the kinds contained in this classification. Inspecting each of these argument forms in turn, Peirce undertook to show that the pragmatist maxim contained all the information that was relevant to efficiently using arguments of any of these kinds (see Hookway 2005 and EP2: 234f). Thus the argument was designed to show that use of the pragmatist maxim provided all the information that was relevant to using concepts and hypotheses in deductive, inductive, or abductive reasoning. Peirce subsequently came to doubt that he could demonstrate that his classification of arguments was exhaustive and thus looked to supplement the arguments he had employed in 1903.

The 1907 argument for pragmatism does not depend upon a classification of arguments or any other list of the elements of scientific

investigation. Instead, it relies upon concepts that are yet more fundamental to the theory of signs. Peirce's theory of signs is primarily a theory of sign interpretation, a theory of how signs can be understood as having particular objects. Hence the strategy is to derive the pragmatist maxim from an account of how it is possible to provide 'logical interpretants' for intellectual concepts. Although he does not explicitly make this point, we could say that such a proof would establish that, whatever additional forms of arguments could be identified which have a role in the use of intellectual concepts, they will not provide counter-examples to the pragmatist maxim. The proof of pragmatism relies upon facts about the possibility of understanding, rather than being based on specific claims about the cognitive activities that are involved in the method of science. Thus he seems to argue that the pragmatist maxim guides us to the 'ultimate logical interpretant' of a concept or hypothesis. The following sections trace how he understands these concepts and how he gets to this conclusion.

2. Signs, objects, interpretants.

In order to deal with this argument for pragmatism we shall need to discuss some of the concepts used in Peirce's theory of signs. Before we examine the 'definition' of a sign that we find in the 1907 manuscript, it will be best to use a simple example to introduce the three most important concepts. Suppose that I treat a rash on someone's skin as a sign that they have measles. If we ask what the rash is a sign *of*, we are asking what the *object* of the sign is. My treating the rash as a sign consists in my judging that the victim probably has measles or expecting other symptoms, or my taking care to avoid infection. This reaction, which constitutes my understanding the rash as a sign is what Peirce calls the 'interpretant'. Although this example is of a simple natural

sign, Peirce takes it that all signs, including linguistic expressions and thoughts, can be understood as having *objects* and *interpretants*.

My understanding of the rash as a sign on measles can take different forms. It may be manifested in the thought that this person has measles; it may take the form of my judging, with the aid of background knowledge of other cases, that this looks like a measles epidemic; it may even take the form of my moving away in order to avoid infection or a feeling of sympathy towards the potential victim. In all these cases, my reaction makes sense against the background that I understand what the rash signifies. Moreover it is important that the sign puts me into indirect contact with the victim's ailment – the reaction is a response to the object as well as a response to the sign. Indeed the interpretant need not be a judgment about the sign, for example the thought that 'this rash is a sign of measles'. As Murray Murphey has made clear, the interpretant of a sign is typically 'a second sign which is determined by the primary sign to refer to the same object to which the primary sign refers and which translates the meaning of the primary sign.' (Murphey 1961: 313, *CP* 1.339, 1.553, 2.303).

The 1907 definition is as follows:

> I will say that a sign is anything, of whatsoever mode of being, which mediates between an object and an interpretant; since it is both determined by the object *relatively to an interpretant* and determines the interpretant *with reference to the object* in such a wise as to as to cause the interpretant to be determined by the object through the mediation of this "sign". (EP2: 410)

He continues by describing the object and the interpretant as 'the two correlates of the sign; the one being antecedent the other consequent of the sign' (ibid). The two italicised phrases can be best understood as follows: the rash is a sign of measles because it has the power to be understood or

interpreted in that way; and the sign can be understood as an indication of measles because the object (the measles) has produced the sign. The measles affect my understanding because the sign (the rash) mediates between the measles and my thoughts about them.

As was noted above, Peirce is primarily concerned with a particular kind of sign: 'intellectual concepts'. Intellectual concepts are 'those upon the structure of which arguments concerning objective fact may hinge', and he adds that they are 'the only sign-burdens that are properly denominated "concepts"' (EP2: 401).

In these writings he is concerned primarily with the distinctive sorts of *interpretants* that propositions, hypotheses and concept expressions typically have when we are concerned with matters of 'objective fact'. And he wants to show that the application of the pragmatist maxim will enable us to identify what he calls the 'ultimate logical interpretants' of such expressions.

The application of Peirce's theory of signs to natural languages would require us to provide a detailed account of how we understand a wide range of complex constructions, and there is no space to do that here. One complexity should be mentioned. The pragmatist maxim is presented as a tool for clarifying 'intellectual concepts' and 'hard words' (CP 5.464), but, for the most part, Peirce's applications of his maxim appear to be to propositions and beliefs, things that correspond to whole sentences. Thus when Peirce clarifies a concept such as *hard*, he does so by describing the 'practical consequences' of some particular object being hard. We clarify hardness by clarifying propositions of the form 'That object is hard', for example. Peirce's later writings contain complex classifications of signs that reflect his system of categories and enable us to identify some different elements of *objects* and *interpretants*. Although we cannot become embroiled in the controversies that

surround these classifications, it will be important to identify some of the kinds of objects and, especially, interpretants that he discusses.

We shall be concerned with some examples of propositions that contain 'intellectual concepts, propositions such as:

Cicero was a Roman orator.

Diamonds are hard.

And we shall take it that linguistic items normally function as signs when their tokens are used in particular contexts. Our first example is about (has as its object) a particular person, the property of being an orator, the possible state of affairs of that person being a roman orator, and so on. The simplest kind of interpretant here involves recognizing and endorsing what was said, or perhaps having such thoughts as *Yes, Cicero was a Roman orator*. In most cases, this is unlikely to happen. If it is already known that Cicero was Tully, then the interpretant may take the form of 'So Tully was an orator'. Or, relying upon further background knowledge, the interpretant may express surprise that an Academic Sceptic could be an orator, or wonder whether oratorical skills contributed to his ability to expose the conspiracy of Catiline. The interpretant that actually occurs, on a particular occasions, when the sign is understood is called the 'dynamic interpretant'. We have already seen how varied dynamic interpretants can be and how they 'develop' the sign by relying upon background information. Another factor which influences how a sign is interpreted is the goals and interests of the interpreter, the reasons he or she has for seeking information by attending to the utterance or other sign. We can see over time the agent may learn more about the object and thus associate ever-richer intrepretants with the utterance.

Of course, something can be a sign even if it has never been interpreted, even if no one has attended to it. Peirce suggests that in such

cases the sign still possesses a kind of potential interpretant, a capacity to determine an interpretant. He called this possible interpretant the 'immediate interpretant', observing that it 'is an abstraction, consisting in a Possibility.' He wrote 'My Immediate Interpretant is implied in the fact that each sign must have its peculiar Interpretability before it gets any interpreter.' (LW:111, 1909) As Short writes, the immediate interpretant 'is something against which to measure any interpretant actually formed. Is it or is it not properly grounded?' (Short 2007: 56). The immediate interpretant which 'is ordinarily called the *meaning* of a sign'(CP4.536) is something that constrains interpretation: if we do not recognize that it would be legitimate to interpret our sentence as a sign that Cicero was an orator, we would be *wrong*.[1]

Peirce also speaks of another idealized kind of interpretant, the 'final interpretant': 'there is certainly a third kind of Interpretant, which I call the Final Interpretant, because it is that which *would* finally be decided to be the true interpretation if consideration of the matter were carried so far that an ultimate opinion were reached' (Letter to William James, 1909, EP 2:496-7). The final interpretant 'consists in a truth which might be expressed in a conditional proposition of this type: "if so and so were to happen to any mind this sign would determine that mind to such and such *conduct*." (Letter to William James, 1909, CP 8.315)

> My Final Interpretant is… the effect the Sign *would* produce upon any mind upon which the circumstances should permit it to work out its full effect … The Final Interpretant is the ne interpretative result to which every interpreter is destined to come if the Sign is sufficiently

[1] Another interesting formulation is provided by "The Immediate Interpretant consists in the *Quality* of the Impression that a sign is fit to produce, not to any actual reaction." (Letter to William James, CP 8.315, 1909)

> considered. … The Final Interpretant is that towards which the actual tends. (Letter to Lady Welby, SS 110-1, 1909)

Very broadly speaking, the distinction concerns what we can learn from signs and utterances. The final interpretant constitutes the totality of what a competent interpreter could eventually learn from the sign or utterance, taking into account thei goals and their growing body of background knowledge. The Immediate interpretation constitutes what we can learn if we can make use of no information beyond what it is provided by the *meaning* of the sentence uttered. In practice, our actual interpretation (the dynamic interpretant) falls between thsee two, presenting what can, in practice, be learned from the meaning of the proposition together with the body of knowledge we actually happen to have acquired (or could acquire) and our ability to make use of it.[2]

3. *Logical Interpretants: three arguments*

In the 1907 manuscript (MS 318) Peirce is concerned with how we understand 'intellectual concepts'. The primary aim of the paper is to formulate and defend the pragmatist maxim, but he immediately tells us that rather than deal with the maxim directly 'a more serious question … concerns the nature of that logical interpretant, the conveyed thought, which we easily assure ourselves that some signs have.' (EP 2:410). This may be the only occasion on which he uses the prhase 'logical interpretant', and he reminds us that 'not all signs have logical interpretants, but only intellectual concepts and the like' (ibid). And right at the beginning, he identifies some formal features that a logical interpretant must have: the proposition that serves as logical interpretant must be in the future conditional tense (EP 2:410). Thus it is

[2] This is not quite right: the dynamic interpretant, unlike both the immediate interpretant and the final interpretant can, presumably, contain error

concerned with a 'would-be's or 'would-acts' (EP2: 402): when we judge that something is *hard*, for example, the logical interpretant would claim that if we were to hit it with a hammer or attack it with a knife, it would not be likely to be damaged or affected. So one task of a proof of pragmatism is to show why this should be so. He offers three arguments for this claim, expressing dissatisfaction with the first two. The third is commonly recognized as a major advance in his theory of signs and a distinctive clue to how the pragmatist maxim can be established.

Even if these arguments led to a conclusion about the logical or grammatical forms of logical interpretants, that would not be enough to establish the correctness of the pragmatist maxim. The pragmatist maxim holds that any meaningful proposition is potentially action guiding. And since, when we are reflecting upon how to act, the action itself is in the future, it may follow that having a formulation in the conditional future tense may be a *necessary condition* for being in accord with the pragmatist maxim. But it is not a sufficient condition: it does not follow that every proposition in the conditional future tense is potentially action guiding; and it is certainly not true that every such proposition entails that actions of a particular kind would have distinctive *sensible* consequences. A proof of the pragmatist maxim needs to forge a link between cognition, action and sensory experience.

His first argument for the conclusion about the grammatical form of logical interpretants depends upon some distinctive features of both interpretants and intellectual concepts.

1) The first stage in the arguments exploits a general thesis about interpretants: there is 'an essential difference … between the nature of an object and that of an interpretant; which difference is that the former antecedes, while the latter succeeds a sign.' It follows from

this that 'the logical interpretant must … be in a relatively future tense.' (EP 2:410)

2) The second stage examines a special feature of 'intellectual concepts and the like': they are all either general or intimately connected with generals, as it seems to me'. And he concludes from this 'that the species of future tense of the logical interpretant is that of the conditional mood, the "*would-be*" (ibid).

It is not surprising that Peirce was dissatisfied with this argument. He admits that the current state of his discussion was 'in a quandary' and looked for other arguments which would show him 'how and why the logical interpretant should in all cases be a conditional future'. The first stage of his argument slides from the premise that the interpretant is in the future when the sign occurs to the conclusion that the content of the interpretant concerns what is in the future with respect to the logical interpretant. There may be an argument for the latter conclusion, but the argument Peirce uses on this occasion is not it.

Peirce then adopts a second strategy that depends upon examples:

> … being in a quandary, it occurred to me that if I only could find a moderate number of concepts which should be at once highly abstract and abstruse, and yet the whole nature of whose meanings should be quite unquestionable, a study of them would go far toward showing me how and why the logical interpretant should in all cases be a conditional future. (ibid)

Having realized that mathematics provided many such examples, Peirce observed that they had a common form. Each concept is associated with a rule or operation, and 'if such and such a concept is applicable to such and such an object, the operation will have such and such a general result; and

conversely' (EP 2:411). Examining 'a score of intellectual concepts, only a few of them being mathematical', he was persuaded that 'any such concept of a real or imaginary object is equivalent to declaring that a certain operation, corresponding to the concept, if performed upon that object, would … be followed by a result of a definite general description.' (EP 2:411)

But Peirce is still dissatisfied. This, again, is unsurprising: at best it provides inductive support, based upon a limited sample, for the claim that all logical interpretants should be of this form. And Peirce presumably wants his semeiotic theory to explain why this *must* be the case. Thus he concludes that this argument 'does not quite tell us just what the nature is of the essential effect upon the interpreter, brought about by the semeiosis of the sign, which constitutes the logical interpretant.' (EP 2:411). We need to derive the account of the logical interpretant from a general account of sign action, one that takes seriously the interactions and interdependencies of objects and interpretants.

The third argument that Peirce then sketches relies on the 'provisional assumption', soon to be recognized as an oversimplification, that interpretants are typically 'modifications of consciousness' (EP 2:411-12). It also relies on the assumption that signs that possess logical interpretants are 'either general or closely connected with generals', which is 'not a scientific result, but only a strong impression due to a life-long study of the nature of signs' (EP 2:413). So, in looking for the logical interpretant, he provides an exhaustive list of those 'categories of mental facts there be that are of general reference.' Presumably, the logical interpretant must belong to one of those categories. There are four:

- Conceptions
- Desires (including hopes, fears, etc)

- Expectations.
- Habits

The core of Peirce's argument is that to use conceptions, desires or expectations to explain the nature of the logical intepretant would involve circularity; and this circularity can be avoided only if we explain it in terms of some kind of *habit* or *habit change*. This is supposed to vindicate the claim that the logical interpretant is in the future conditional mood: habits determine how things would behave in various circumstances, and how we should act in the light of experience.

A normal understanding of the pragmatist maxim will involve several different components:

1. Pragmatist clarifications involve 'would-be's. This appears to be a claim about the logical form of logical interpretants, they are typically formulated in the future conditional form.
2. Such 'would-be's' typically involve *habits*.
3. Logical interpretants typically involve the results of actions, experiments or other operations that can be carried out on things. So they make reference to a) agency and b) experiential results in contexts that involve substantial background information.

The argument from examples is concerned with both 1) and 3), apparently providing empirical evidence that both are true of intellectual concepts, both mathematical and non-mathematical ones. The argument to the conclusion that only a habit change can serve as the ultimate logical interpretant of an intellectual concept addresses 1) and 2), but it does not appear to take into account the claim, fundamental to the pragmatist maxim, that its clarifications identify *practical consequences* or experiential results. Hence, although this argument can *contribute* to a demonstration of the correctness of pragmatism,

it requires considerable supplementation of the sort provided in the 1903 argument by the claim that the inductive method, or the 'method of science' is fundamental to cognition.

4. The charge of circularity

Peirce dismisses the first of the four candidates by saying that 'it is no explanation of the nature of the logical interpretant (which, we already know, is a concept) to say that it is a concept.' (EP 2:412) This is undeniable, but it may be uncertain whether using the idea of a concept in explaining what logical interpretants are has to take this form. For example, suppose that I want to provide a logical interpretant of the concept of a bachelor and do so by explaining that a bachelor is an adult unmarried male person. This can be informative because we clarify one concept by showing that it is related to other concepts. Suppose then that someone defended the view that *any* intellectual concept can be interpreted by tracing its connections to other concepts. Peirce's argument as given above does not seem to be effective against this proposal. However, what he has to say about expectations and desires indicates that he would be unhappy with this. Both desire and expectation are useless as explanations of logical interpretants because 'neither of these is general otherwise than through connection with a concept.' (EP 2:412) Peirce apparently holds that the kind of generality that is characteristic of concepts must be explained in terms of something wholly non-conceptual. Rather than explaining particular concepts, their similarities and differences, he seeks a (reductive?) explanation of *what it is to be a concept* and how applicable concepts are possible.

It is useful to compare these arguments with those that he employed thirty years earlier in his paper 'How to make our ideas clear'. Peirce describes three grades of clarity that we may aspire to (or achieve) with

respect to our ideas; we can take it that possession of a logical interpretant is related to the achievement of the highest, third grade of clarity. He followed Leibnitz's use of the distinction between clear and distinct ideas, 'in making familiarity with a notion the first step towards clearness of apprehension, and the defining of it the second (CP5.392). Provision of an abstract definition clarifies a complex concept by analysing its relations to other, possibly more abstract or simpler, ones. This occurs when we define *reality* as 'that whose characters are independent of what anyone thinks them to be' (CP5.405). And useful as it may be, 'it would be a great mistake to suppose that it makes the idea of reality perfectly clear' (CP5.406). This is deemed insufficient because it provides a conditional clarification: it clarifies the concept only if we are already clear about the concepts used in giving the clarification. The degree of clarity such a definition provides depends upon how clear we are about the concepts used in the definiens. We can use a concept in definitions without fully understanding it. Full clarity will be achieved only when we have a full description of how the expression is actually used, one which leaves nothing open to further interpretation. We clarify a concept fully when we know in detail what is involved in having beliefs that contain it; and, since a belief is described as a 'habit of action' clarifying a concept involves identifying the habits involved in such beliefs. So in 1878, we achieve maximal clarity about concepts or propositions by identifying the habits of actions that belief in them involve; and in 1907 we obtain a logical interpretant for an intellectual concept by identifying a habit.

Why should habits provide a way out of these supposed problems? An important concept here is that of an 'ultimate' interpretant: only a habit can serve as the ultimate interpretant of an intellectual concept. What this involves is clear from a version of MS 318. Peirce acknowledges that a 'mental sign' might function as a logical interpretant for an intellectual

concept. But, he observes, since what is interpreted is 'of an intellectual kind – as it would have to be- it must itself have a logical interpretant; so that it cannot be the *ultimate* logical interpretant of the concept.' (CP5.475, 1907) To be an ultimate logical interpretant, it would have to be something general that did not need, or could not receive a logical interpretant in its turn. A crucial passage runs:

> It can be proved that the only mental effect that can be so produced and that is not a sign but is of a general application is a *habit-change*; meaning by a habit-change a modification of a person's tendencies toward action … (CP5.475)

This is all very murky, and, in order to understand and evaluate it, we need a better grasp of this all-important concept of a *habit*.

6. Ultimate interpretants

Peirce's argument rests on the assumption that every intellectual concept must receive (or be capable of receiving) what he called an 'ultimate interpretant'. And the argument is intended to show that an ultimate interpretant of a concept must be a *habit-change* of some distinctive kind. The distinctive feature of an 'ultimate interpretant' is that it is a response to a sign which can function as an interpretant but is not itself a sign. It is an immediate consequence of this that an ultimate interpretant does not itself require to be interpreted, or, indeed, cannot be interpreted. How can there be interpretants which are not themselves signs? And why is it indispensable that intellectual concepts can receive such interpretants? Peirce does little to answer these questions. So I shall begin the discussion by making some observations about just how we can manifest our understanding of concepts.

Our mastery of a concept, or our understanding of a general word or related kind of sign, can be manifested in a variety of different ways. One possibility is that we display our mastery of a concept by *specifying* it, by saying what it means. This may take the form of a verbal definition, perhaps a set of necessary and sufficient conditions. Thus I can display my mastery of *bachelor* by saying that a bachelor is an unmarried adult male. Specifications may take other forms too. They can be partial, perhaps just identifying a contextually salient sufficient condition. And, in the context of a particular conversation or discussion, they may be relative to a body of background knowledge or to judgments of salience that reflect the cognitive or practical goals of those whose understanding is being described.

Our mastery of a concept can also be manifested, not by our ability to specifiy its content but by our ability to *use* the concept appropriately in thought and discourse. For example, my understanding of *dog* may be manifested in an ability to recognize does when I see them. Or, my understand of *bachelor* may involve a disposition to infer that people are unmarried when I learn that they are bachelors. And my inferences of logical concepts (the conditional, conjunction, disjunction etc) may be displayed by the inferences I accept and those that I reject. A capacity to endorse steps of modus ponens and modus tollens shows is inseparable from an understanding of conditionals and negation. It is easy to see how, in such cases, we may be able to specify the content of a concept without being able to use it effectively in reasoning and inquiry, or we may be able to use a complex concept without being able to specify its content.

Even if we restrict our attention to cases where we possess both abilities, to specify the content of a concept and to use it in reasoning inquiry, we can distinguish some different kinds of case. The differences concern the relations between the two capacities, which possesses a certain kind of priority. One possibility is that the specification possesses priority: for

example, our ability to *use* the concept may depend upon our ability to reflect upon the information contained in our specification and using the concept in the light of the guidance that the specification provides. Another possibility is that the use is fundamental: although we can provide a specification, this is offered as a description of the content that is manifested in our use.

The force of this distinction emerges when the specification and the use are not in harmony. In the first sort of case, if our use is not in accord with our specification, then, ceteris paribus, our use should be revised so that it corresponds to the specification. In the other case, by contrast, the specification is answerable to the use. We should revise our specification so that it is in harmony with our evidence about how the expression or concept is used. In this case, the specification has no regulative authority in determining correct use. Indeed it is easy to see how our attempts to describe our confident use of a concept can be fallible, and recognition of the erroneous character of our specification may have no tendency to weaken our confident in the habits that guide our use. And in such cases when we provide a specification of some sign, this *need* not determine its significance.

The manifestation of understanding that consists in *use* is more fundamental than that which is revealed in a specification. When I provide a *specification* of a concept, this only provides conditional evidence of my understanding: it manifests understanding of the concept only if I already understanding the concepts made us of in my interpretation. Someone can parrot the definitions of *bachelor* or *complex number* without understanding what they are saying. Indeed, they can do this even if they believe they understand what they are saying. But if I can use the concept successfully, then there is no further question of whether I really understand what I am saying. So the interpretation of a sign that consists in providing a specification is only conditional; but the in interpretation that consists in unreflective successful use appears to provide an unconditional interpretation.

Some passages from another draft of MS 318 supports these claims. Although a concept can be a logical interpretant, it is 'only imperfectly so'. This is because it resembles a verbal definition, 'and is as inferior to the habit, and much in the same way, as a verbal definition is inferior to the real definition.'

> The deliberately formed self-analyzing habit … is the living definition, the veritable and final logical interpretant. (CP 5.491

We are now getting close to the pragmatist maxim. Peirce concludes that 'the most perfect account of a concept that words can convey will consist in a description of the habit what that concept is calculated to produce. CP 5.491).

7. Habits

Peirce's use of the concept of a *habit* leads some readers to accuse him of a commitment to a kind of reductive behaviourism, especially when it appears that he thinks that habits of action and inference do not need any sort of interpretant. In this section, we shall explain how Peirce understood habits and why it was so important for him to distinguish them from dispositions, and we shall explore what sort of content habits of actions and inference have. Bt the end of the section, I hope we will be able to understand how it is possible for habits to serve as ultimate interpretants. The concept of habit was used extensively in Peirce's writings from the 1860s until his final writings after 1900. It was used both to describe the nature of laws of nature and related metaphysical aspects of reality and as a tool for explaining cognition. We shall concentrate upon the latter, but the metaphysical use of *habit* will also be relevant to our discussions.

According to Peirce, 'every reasoner … since he approves certain habits, and consequently methods, of reasoning, accepts a logical doctrine, called his *logica utens*.' He will thus rely on 'a general habit of thought which

he may not be able precisely to formulate, but which he approves as conducive to true knowledge' (CP2.773) This *logica utens* is the product of experience rather than something learned from the 'logica docens' produced by theorists. This is not just a claim about primitive unsophisticated thinking. Thus 'Mathematics performs its reasonings by a *logica utens* which it develops for itself, and has no need for a *logica docens*; for no disputes about reasoning arise in mathematics which need to be submitted to the philosophy of thought for decision.' (CP1.417). We trust our habits of inference and see no need to defend them unless given reason to doubt their reliability. The role of habits in reasoning is reinforced when we become reflective about our practice. The normative sciences are concerned with 'the purposive formation of habit, as common sense tells us they are in measure controllable' (MS655, 1910, see Kent 1987: 148-9). In this spirit, Ethics is 'the theory of the formation of habits of action that are consistent with the deliberately adopted aim'(Kent 1987: 133). Aesthetics fosters the 'Esthetic ideal through the cultivation of habits of feeling' (Kent 1987: 162), and logic aims to 'discover the habits of inference that will lead to positive knowledge if there should be reality' (CP2.60-4??, Kent 1987: 66)). This concern with the habits we *ought* to have shows that the role of habits in reasoning is not restricted simply to the reliance on vague instinctive patterns of reasoning that characterizes the starting point for reasoning. Self-control involves perfecting our habits; it does not involve escaping from dependence upon habits at all.

In the 1907 manuscript on Pragmatism (MS 318) Peirce is anxious to distinguish habits from mere dispositions:

> Habits differ from dispositions in having been acquired as a consequence of the principle, virtually well known even to those whose powers of reflection are insufficient to its formulation, that multiply reiterative behaviour of the same kind, under similar

> circumstances of percepts and fancies, produces a tendency – the habit – actually to behave in a similar way under similar circumstances in the future.'
>
> Moreover – *here is the point* – every man exercises more or less control over himself by means of controlling his own habits; and the way in which he goes to work to bring this effect about in those cases in which circumstances will not permit him to practice reiterations of the desired kind of conduct in the outer world shows that he is well acquainted with the important principle that *reiteration in the inner world – fancied reiterations – if well-intensified by direct effort, produce habits,* just as do reiterations in the outer world; and *these habits will have power to influence actual behaviour in the outer world*; especially if each reiteration be accompanied by a peculiar strong effort that is usually likened to issuing a command to one's future self. (EP 2:413)

A very clear exposition of the role of habits is found in a manuscript from 1895 called 'Of reasoning in general'. Peirce explains *belief* by saying that it is 'a state of mind of the nature of a habit, of which a person is aware, and which, if he acts deliberately on a suitable occasion, would induce him to act in a way different from what he might act in the absence of such a habit.' (EP 2:12) This is supported by two examples. If someone believes that a straight line is the shortest distance between two points, then if he wants to take the shortest route between one place and another, and he thinks that he can travel in a straight line, then he will do so. The second example concerns someone who believes that drinking alcohol is harmful but also desires not to harm herself. Were she to continue to drink alcohol 'for the sake of the momentary satisfaction, then she would not be 'acting deliberately'.

When we speak of a drug addict's *habit*, the habit is evidently something that limits the addict's freedom; it is something that imposes itself

upon the addict's deliberations and prevents him from exercising *self-control*. It should be clear that this is not how Peirce understands habits. Our habits are subject to control and, presumably, we can interfere to stop ourselves acting habitually on particular occasions. But we can act 'automatically' in accord with these habits – we respond appropriately without the mediation of deliberate reflection. As discussed earlier, this means that we can describe our habits – 'make them explicit' in a familiar terminology; and we can reflect upon them and evaluate them, they provide a way of acting in accord with our beliefs and standards without needing the mediation of interpretants which take the form of signs.

6. Some problems facing the interpretation and assessment of these views

Considered as an argument for the pragmatist maxim, the arguments we have examined seems to be incomplete for several reasons.

a) Peirce simply lists the four candidates for logical interpretants (desires, expectations, concepts, and habits). He does not appear to give any reasons for thinking that there are no further possibilities.
b) Much of the time, he appears to be concerned with formal, grammatical or logical features of logical interpretants: they are in the conditional future tense. This says nothing about the *content* that the interpretant has. Thus how can we rule out formulations which do not present material which has practical bearings or relations to experience. Peirce talks about 'would-be's in this regard, yet elsewhere in discussing the pragmatic maxim, he talks about 'would-acts' (EP 2:402)

The passages we have considered so far, prompt another worry. Habits are of many kinds. There are indications that Peirce wants the ultimate

logical interpretant to be a habit of *action* but, in this manuscript he does not rule out the possibility that these interpretants could be habits of some other kind. For example they could be habits of inference which are cut of from the determination of action. Other drafts of MS 318 suggest how Peirce would respond to this concern (see CP 5.467-496 and especially CP 5.491). Describing what occurs when we carry out mathematical reasoning, Peirce writes that 'the activity takes the form of experimentation in the inner world; and the conclusion (if it comes to a definite conclusion), is that under given conditions, the interpreter will have formed the habit of acting in a given way whenever he may desire a given kind of result.' And this habit of action is 'the real and living logical conclusion; the verbal formulation merely describes it'. This emphasis upon action is reinforced when he adds that 'the habit conjoined with the motive and the conditions has the action for its energetic interpretant; but action cannot be a logical interpretant because it lacks generality'. So it seems clear that what Peirce has in mind when he talks about habits in this connection is habits of action. But this may strengthen an earlier worry: is Peirce's list of candidate logical interpretants complete? In order to show this, he would have to establish that there are no other candidates. Can we eliminate the possibility that there are mental or intellectual habits which are general yet do not require interpretation and which are not habits of *action*.

These observations relate to a single general concern. Most of Peirce's characterizations of the pragmatist maxim emphasize that in clarifying a concept we need only attend to considerations that relate to the determination of action or indicate sensible effects that we would expect the object of the concept to have when acted upon appropriately. This may be in the background of the arguments found in MS 318, but there does not appear to be any explicit discussion of how these elements are to be introduced. The earlier arguments, which appeal to the fundamental role of concepts as

belonging to 'inductive method' appear to fill this gap. Why doesn't Peirce now see any need to argue for the view that the habits which serve as ultimate logical interpretants are habits of action. If all he can show is that logical interpretants must be in the conditional future tense, he has yet to vindicate the pragmatist maxim as commonly understood. And the claim that we are concerned with habits of action is what needs to be established.

The pragmatist maxim is often presented as a tool that can be used to dismiss *a priorism* in science and philosophy; and Peirce saw *a priorism* as a genuine intellectual threat. But, the arguments contained in MS 318 do not seem to support this aspect of Peirce's pragmatism. The final section of this paper makes some conjectures of why he came to think that he did not need to address this issue.

8. The proof of pragmatism

So how should we understand the relations between Peirce's discussion of logical interpretants in MS 318 and his defence of the maxim of pragmatism? Here is one conjecture. First we should distinguish two elements in Peirce's formulations of the maxim:

- There is the anti-apriorist flavour which is captured in remarks such as:
- There is the pragmatist flavour which links meaning to *practical* consequences and agency.

In his writings from 1878, the two are connected by Peirce's views about the role of experience in the guidance of action: sensible experience always acts as the stimulus for action and the consequent of action (EP1: 129-30). As my labels have suggested, it is the second of these that is most characteristic of *pragmatism* properly so-called. And it is conceptually possible that someone should adopt a position that shared the verificationist character of

pragmatism while rejecting the additional thesis that this was inseparable from a concern with agency.

In section 6, we noted some passages in which Peirce seems to assume that the only kind of mental habit that we need take account of is a habit of action. If that is accepted, there may be an easy route to establishing the correctness of the maxim of pragmatism. In the draft of MS 318 that we have referred to above, he asserts that 'the most perfect account of a concept that words will convey will consist in a description of the habit what that concept is calculated to produce' (CP 5.491). When we take account of his assumption that habits are habits of action, it is no surprise that he raises the rhetorical question:

> 'But how otherwise can a habit be described than by a description of the kind of action to which it gives rise, with the specification of the conditions and of the motive?' (CP 5.491)

If the answer to this question is that such a description is all is required, then, we should be able to conclude that the application of the pragmatist maxim leads to 'the most perfect account' of intellectual concepts. But we have seen that the claim that logical interpretants are all *habits of action* is in need of more defence than Peirce provides in these manuscripts. How might Peirce fill this gap in his argument?

We need to be careful in identifying what the inquirer can already know when he or she raises the question of whether and why we should endorse the maxim of pragmatism. In 1903, Peirce's attempt to prove the correctness of pragmatism rested on the assumption that there were just three kinds of argument, and his subsequent anxieties of his 1903 proof appear to have arisen from doubts about that assumption. What does Peirce think he

can take for granted in 1907? The maxim of pragmatism is supposed to be a logical principle, but, since he took it to belong to the third of the logical disciplines, 'methodeutic', he is entitled to rely upon results obtained from the other normative sciences, esthetics, ethics, speculative grammar, and critic. We have already seen the use he makes of information from speculative grammar, and, although he is entitled to make use of results from critic, we have seen that he had doubts about just what information was available from that source. So what information could he make use of which might help to close the gaps in his proof of the maxim that we have identified?

John J Fitzgerald suggests an answer to this question. Commenting on the passage in which Peirce argues that the ultimate interpretant should be a change of habit rather than an expectation, desire or concept, he wrote that 'It seems that the narrowing down of the field to a habit or a habit-change cannot be accomplished merely from the theory of signs, but requires further that one know what is the intended purpose of the sign-user' (1966: 163). The choice of habit is appropriate because the sign user 'is working towards his ultimate aim through the pursuit of inquiry.' (ibid) Ethics can give 'the decisive reason in favor of habits' because it reveals that 'the ultimate purpose which specifies our use of intellectual signs is the apprehension of thirdness in the world and a corresponding growth in rationality of the scientist.' (1966: 165, and cf. Murphey 1961: 361ff).

So, one possibility is that the maxim of pragmatism is evaluated as a means to achieving some ultimate cognitive goal which is identified by the other normative sciences. We might even suppose that the anti-a priorist element is established once we have identified the 'ultimate' cognitive goal. Why should we then need a proof for the pragmatist maxim? Maybe to demonstrate the importance of the claim that meaningful propositions must have *practical* consequences. Perhaps the argument has to show that agency

(in the form of habits of action) is involved in the application of empirical concepts to experienced objects. In that case, an important component of the proof Peirce seeks would be an account of this connection between agency and the application of concepts.

Unfortunately, there is not much textual support for such an interpretation of Peirce's proof. The manuscripts we have been discussing do not have much to say about our ultimate ends, and the discussion of *habits* is fairly cursory, providing few details about the relations between agency and habit and about the role of these ideas in explaining how we can apply empirical concepts to external things. Indeed, the discussion of 'ultimate interpretants' is also not particularly clear. If we are to understand this material, we shall need to explore Peirce's views about the objects of perception and about how we are able to apply general terms to experienced objects.

When we bring together Peirce's claims about ultimate logical interpretants and his concern with the pragmatist maxim, it would be easy to conclude that there is a tension between them. Compare these two claims:

1) The real meaning of a concept consists in a habit of action which serves as an intepretant for the concept but which, since it is not a sign, cannot receive a logical interpretant of its own.

2) Applying the pragmatist maxim to the concept leads to a description of this habit of action, something which makes explicit what is implicit in our employment of the habit of action which constitutes the meaning.

It is natural to conclude that what we obtain from the use of the pragmatist maxim is itself a logical interpretant of that habit which, we have seen, can have no logical interpretant.

T.L. Short has recently expressed views about the significance of these 1907 writings for the development of Peirce's views about signs which may help us to explain why these views do not involve any inconsistency (Short 2007: 58f). The 'fundamental change in doctrine that occurred in 1907', he holds 'was to have recognized that it is the habit itself, and not the concept of it, that is the ultimate interpretant.' What might we learn from this? It seems to follow that the pragmatist clarification is a *description* of the ultimate interpretant, and its correctness depends upon whether it correctly describes the habit which constitutes the meaning of the concept. If the habit and the verbal clarification do not match, then it is the pragmatist clarification that has to be revised. It may be compatible with this that the pragmatist clarification *is* an interpretant, albeit not an ultimate one. If so, it is an interpretant of the original concept, not of the habit, because the habit is not, itself, a sign.

References

Boler, John. 1964. 'Habits of thought', in Moore and Robins (ed) *Studies in the Philosophy of Charles Sanders Peirce: second series*. Amherst: University of Massachusetts Press: 165-90.

Fitzgerald, John J. 1966. *Pragmatism and the Theory of Signs*. Paris: Mouton.

Gentry, George. 1952. 'Habit and logical interpretant', in Wiener and Young (eds) *Studies in the Philosophy of Charles Sanders Peirce*. Cambridge MA: Harvard University Press: 75-90.

Hookway, Christopher 2004. 'The principle of pragmatism: Peirce's formulations and examples', *Midwest Studies in Philosophy*, 119-136.

------- 2005 'The Pragmatist Maxim and the Proof of Pragmatism', *Cognitio*, v. 9: 25-42

-------- 2008 'The Pragmatist Maxim and the Proof of Pragmatism (2) after 1903', *Cognitio*,v. 9: 57-72.

Kent, Beverley. 1987. *Charles S. Peirce: Logic and the Classification of the Sciences*. Kingston and Montreal: McGill-Queen's University Press

Murphey, Murray. 1961. *The Development of Peirce's Philosophy*. Cambridge MA: Harvard University Press.

Peirce, Charles Sanders (1931-60) *Papers of Charles Sanders Peirce,* eight volumes, eds. C. Hartshorne, P. Weiss and A. Burks. Cambridge: Harvard University Press. References are of the form *CP* n.m - to paragraph m of volume n.

------- (1998) *Essential Writings, volume 2*. Eds. The Peirce Edition Project. Indianapolis: Indiana University Press. References are of the form EP *2*, followed by page number.

Short, T.L. 2007. *Peirce's Theory of Signs*. Cambridge: Cambridge University Press.