Leibniz on Consciousness and Self-Consciousness

Rocco J. Gennaro

[Final version in NEW ESSAYS ON THE RATIONALISTS, Oxford U Press, 1999]

In this paper I discuss the so-called "higher-order thought theory of consciousness" (the HOT theory) with special attention to how Leibnizian theses can help support it and how it can shed light on Leibniz's theory of perception, apperception, and consciousness. It will become clear how treating Leibniz as a HOT theorist can solve some of the problems he faced and some of the puzzles posed by commentators, e.g. animal mentality and the role of reason and memory in self-consciousness. I do not hold Leibniz's metaphysic of immaterial simple substances (i.e. monads), but even a contemporary materialist can learn a great deal from him.

1. What is the HOT Theory?

In the absence of any plausible reductionist account of consciousness in nonmentalistic terms, the HOT theory says that the best explanation for what makes a mental state conscious is that it is accompanied by a thought (or awareness) that one is in that state.\(^1\) The sense of 'conscious state' I have in mind is the same as Nagel's sense, i.e. there is 'something it is like to be in that state' from a subjective or first-person point of view.\(^2\) Now, when the conscious mental state is a first-order world-directed state the HOT is not itself conscious; otherwise, circularity and an infinite regress would follow. Moreover, when the higher-order thought (HOT) is itself conscious, there is a yet higher-order (or third-order) thought directed at the second-order state. In this case, we have introspection which involves a conscious HOT directed at an inner state. When one introspects, one's attention is directed 'back into' one's mind.

For example, what makes my desire to finish this paper a conscious first-order desire is that there is a (nonconscious) HOT directed at the desire. In such a case, my conscious focus is directed at the paper. If I am introspecting my desire, however, then I have a conscious HOT directed at the desire itself.\(^3\)
It is also helpful to distinguish between momentary focused introspection which only involves a brief conscious HOT, and deliberate introspection which involves the use of reason and a more sustained inner-directed conscious thinking over time. Sometimes we consciously think to ourselves in a deliberate manner, e.g. in doing philosophy or in planning a vacation. We are also often engaged in deliberate activities (in the sense of 'voluntary' or 'purposeful') directed at the external world, e.g. when our conscious attention is absorbed in building a model airplane. Although not all deliberate activity involves introspection, clearly some does involve sustained conscious thinking directed at one's inner states. But there is also a more modest type of introspection. One might consciously think about a mental state without deliberating in any way, e.g. momentarily think about a memory or briefly consciously focus on a pain or emotion. In these cases, one is not engaged in deliberation or reasoning. Some animals seem capable of this kind of introspection even if they cannot deliberate. Like deliberate introspection, such 'momentary focused introspection,' as I will call it, involves having conscious HOTs.

I suggest that self-consciousness is simply having meta-psychological or higher-order thoughts, even when the HOT is not itself conscious. I have therefore argued at length that consciousness entails self-consciousness, but more important here is that there are degrees or levels of self-consciousness, with introspection as its more complex form. All introspection involves self-conscious states, but not necessarily vice versa. Thinking about one's own mental states is definitive of self-consciousness, but nothing requires that the HOT itself be conscious. Just as one might have nonconscious thoughts directed at the world, one might have them directed at one's own mental states.

Some might still wonder why self-consciousness need not be consciousness of something. I offer two reasons here: (1) Few, if any, philosophers hold that self-consciousness is literally "consciousness of a self," especially since Hume's observation that we are not aware of an unchanging or underlying self but only a succession of mental states. Thus the "ordinary
meaning” of 'self-consciousness' is up for grabs since the term does not wear its meaning on its sleeve. It seems that we are somewhat free to stipulate a meaning (though not entirely arbitrarily, of course).

(2) Other philosophers have put forth even weaker criteria for what counts as 'self-consciousness'. For example, Van Gulick urges that it is simply the possession of meta-psychological information. While I believe that his notion is too weak, my point here is that my definition is not the weakest in the literature. Owen Flanagan also recognizes a "weaker" kind of self-consciousness: "all subjective experience is self-conscious in the weak sense that there is something it is like for the subject to have that experience. This involves a sense that the experience is the subject's experience, that it happens to her, occurs in her stream." (p. 194)

2. Leibniz and the HOT Theory

A. Unconscious Mentality and Petites Perceptions

Unlike Descartes, Leibniz held there to be degrees of awareness or perception. Moreover, Leibniz did not believe (as Descartes and Locke apparently did) that consciousness is essential to mentality or, at least, to each episode of thought or perception. There are petites perceptions or nonconscious perceptions. Indeed, Leibniz held that all monads, the ultimate constituents of reality, perceive in the sense that they 'represent' external things. We might say that all monads always have some "perceptual" or informational states with representational content, but they need not be conscious. Leibniz therefore recognized that there must be an answer to the question: what makes a perception a conscious perception? Indeed, we should ask more generally: what makes a mental state a conscious mental state? (Of course, if 'perception' is used as a generic term to cover all mental states, then the questions are equivalent.) This is the fundamental question that should be answered by any theory of consciousness. The HOT theory says that what makes a mental state conscious is the presence of a suitable higher-order thought directed at it. I believe that Leibniz also held this view.
In connection with the idea that there are petites perceptions, Leibniz often speaks of perceptions which are not apperceived and says that the Cartesians made the great mistake of not accounting for them (PNG 4; AG p. 208). Consider the statement that "...since on being awakened from a stupor, we apperceive our perceptions, it must be the case that we had some perceptions immediately before, even though we did not apperceive them." (Mon. 23; AG p. 216) The implication is that the perception is there nonconsciously and the apperception of it makes it conscious. This sounds like the HOT theory: a higher-order 'apperception' of the lower-order perception makes it conscious; otherwise, it would remain a nonconscious mental state. Of course, nonconscious mental states can still play a role in the production of behavior: "[i]t would not be adding much...if I said that it is these minute perceptions which determine our behavior in many situations without our thinking of them..." (RB pp. 55-6) But, again, it is the "thinking of them" that makes them conscious.

B. The Infinite Regress Objection

As I mentioned earlier, the HOT theorist must avoid definitional circularity and an infinite regress by explaining that the HOT need not itself be conscious when one has a first-order conscious state. Otherwise, we would be answering our basic question by appealing to consciousness, which is circular. Moreover, we would have an infinite regress because for every conscious state there would have to be a higher-order conscious state and so on ad infinitum. Leibniz cleverly noticed this problem and so made room for nonconscious second-order awareness when he said that

it is impossible that we should always reflect explicitly on all our thoughts; [otherwise] the mind would reflect on each reflection ad infinitum, without ever being able to move on to a new thought. For example, in being aware of some present feeling, I should have always to think that I think about that feeling, and further to think that I think of thinking about it, and so on ad infinitum. It must be
that I stop reflecting on all these reflections, and that eventually some thought is allowed to occur without being thought about; otherwise I would dwell for ever on the same thing. (RB p. 118)

Presumably, the first part of this ('it is impossible that we reflect explicitly upon all of our thoughts') should be taken to mean that what makes our mental states conscious cannot always be higher-order conscious thoughts, and so instead must sometimes be higher-order nonconscious thoughts or, we might say, unconscious apperceptions. Otherwise there would be an infinite regress and we would never 'be able to move on to a new thought.' However, it can be the case that we always reflect implicitly (i.e. nonconsciously) on our thoughts or perceptions. As Rescher\(^{12}\) puts it, "[t]he iterative piling-up of reflective awareness of reflective awareness must stop somewhere, and the conception of unconscious apperception provides a convenient means of termination." (p. 127)

I suggest that this is of a piece with the HOT theory and particularly with the idea of nonconscious second-order thoughts rendering lower-order states conscious. It seems reasonable to suppose that Leibniz was saying that "what makes a perception a conscious perception is the presence of a higher-order perception directed at it. However, the higher-order perception need not itself be conscious; otherwise, an infinite regress would follow." We might then use the notion of 'unconscious apperception' as our lowest form of self-consciousness, i.e. nonconscious meta-psychological thoughts. If we think of all apperception as higher-order inner-directed perception, then it is natural to distinguish between conscious and nonconscious apperception in the same way that Leibniz distinguishes between conscious and nonconscious first-order perception. It is therefore also reasonable to understand apperception as coming in degrees in the way that I described self-consciousness above.

**C. Memory**

Leibniz sometimes says that what makes a mental state or perception conscious is that the
subject remembers it. He speaks of 'sensation' as "perception accompanied by memory" (PNG 4; AG p. 208) and says that "a present or immediate memory, the memory of what was taking place immediately before -- or in other words, the consciousness or reflection which accompanies inner activity -- cannot naturally deceive us." (RB p. 238) Notice that the mental state is already present for a short time and then it is remembered shortly thereafter. The HOT theorist can similarly say that the HOT occurs shortly after the lower-order state. This echoes Leibniz's claim in the Monadology which is worth quoting again: "...since on being awakened from a stupor, we apperceive our perceptions, it must be the case that we had some perceptions immediately before, even though we did not apperceive them." (Mon. 23; AG p. 216)

But, most importantly, Leibniz is saying that what makes the perception conscious is a memory of it, which clearly entails that there is a higher-order state directed at the perception. If a perception occurs without the accompanying memory or higher-order state, then it will be one of our many nonconscious perceptions. On the other hand, the second-order state is a memory or 'record' of a first-order state which has occurred immediately prior to it. Thus, 'the immediate memory of a perception' sounds very much like 'the apperception of a perception' and thus a kind of self-consciousness that makes the lower-order perception conscious.  

But we must be careful not to fall into the trap of the infinite regress by requiring that the memory is itself conscious. Leibniz should not be taken to hold the untenable view that the accompanying memory of every conscious perception is always itself conscious. There are, of course, various kinds of memory; some more sophisticated than others. For example, there is the rather robust 'episodic' memory which involves consciously thinking about oneself experiencing something in the past. But this kind of memory cannot be the 'memory-states' Leibniz has in mind here. In this context, Leibniz seems to be using 'memory' as a very unsophisticated kind of nonconscious 'registering' or 'tracking'.

D. Some support from commentators
Mark Kulstad did not have the HOT theory in mind, but says the following: a perception which becomes...distinct enough to arouse a second-level mental activity, becomes thereby a sensation, or, if you like, apperceived. Whether this activity is called memory, apperception, consciousness, or reflection, is perhaps not so important, for in the end all refer to the mind's second-level or reflective awareness of its own perceptions. (p. 39)

This is clearly in the spirit of the HOT theory, although it is important not to put 'consciousness' into his list because otherwise we face the infinite regress problem. As we have seen, however, the terms 'apperception' and 'memory' can have a nonconscious sense. The terminological problems understandably often lead Kulstad to use the more neutral term 'awareness,' which clearly has both a conscious and nonconscious sense. Kulstad (p. 146) also agrees with Rescher and says that "apperception always involves an inner directed perception: Leibniz...holds that consciousness is a two-tiered affair, with the thought of which we are conscious being itself perceived." Again, the HOT theorist should be very sympathetic with this position. Rescher explains that

Apperception...is not consciousness as such...but self-consciousness or self-perception generally, involving the capacity for reflexive self-revealing perception of the workings of one's own mind. The procedure of some commentators in equating apperception with conscious perception in general is not faithful to Leibniz' own equation: apperception = inner-directed perception. Leibniz' distinction between perception and apperception is in strict parallel to Locke's distinction between sensation and reflection, with consciousness present on both sides of the boundary. (pp. 119-20)

The first part of this seems right, but I do not know what to make of the Lockean analogy. Rescher cannot mean to equate Locke's sensation with Leibniz' perception per se since, according
to Locke, sensation is always conscious whereas, as Rescher knows, Leibniz makes room for unconscious perceptions. Rescher makes it sound as if 'unconscious perception' is a contradiction for Leibniz, which it clearly is not.

In any case, "consciousness as such" presumably means first-order world-directed conscious states. Such states are not themselves apperceptions because they are directed at the outer world. However, first-order conscious states do entail apperceptions which render them conscious. Apperception is necessary for conscious perception, but we should not equate them.

3. The Linguistic Issue

The French verb for 'to perceive' was in Leibniz's time and still is 'apercevoir' and the verb for 'to be aware of' was and is 's'apercevoir de'. The noun for the former is 'perception', but there was no corresponding noun for the latter and so Leibniz coined the technical term 'apperception'. Since Leibniz had the noun for (first-order) perception, it would have been unnecessary for him to create another term for it (cf. KUL pp. 21-3). This suggests that apperception is designed to cover higher-order perception and, just as first-order perceptions can be both unconscious and conscious, it is again reasonable to suppose that Leibniz allowed for both unconscious and conscious apperceptions.

In their translation of the New Essays, Remnant and Bennett do not use 'apperception', but instead translate Leibniz's invented term as 'awareness'. 'To apperceive' is 'to be aware of'. But if apperception is only higher-order perception, then their use of 'aware' can be misleading since one can also be aware of external objects. Indeed, in their note on 'aware' (p. xxvii), they recognize that there are places where certain contrasts are lost. In any case, given that apperception is always higher-order, we must again note that 'aware' is ambiguous between 'nonconscious higher-order awareness' and 'conscious higher-order awareness'.

The French word 'réflexion' is typically translated as 'reflection' and is ambiguous between various forms of self-consciousness. The subtle distinctions we have made were
perhaps not distinguished properly in the dictionary or the minds of Leibniz's time. There is a sense in which 'reflexion' means something closest to introspection and probably even the more sophisticated 'deliberate introspection'. Indeed, there is a cross reference to deliberation and meditation. One standard French dictionary translates 'reflexion' as: Retour de la pensee sur elle-meme en vue d'examiner plus a fond une idee, une situation, un probleme.\textsuperscript{16} This can be translated as the "return of thought onto itself in order to examine more deeply an idea, a situation or problem." The latter part suggests deliberate introspection. Moreover, Leibniz usually uses 'reflexion' when discussing reason and our ability to discover in us necessary truths and innate ideas. Thus, it is wisest to identify Leibniz's 'reflexion' with our deliberate introspection and with a very sophisticated form of apperception.\textsuperscript{17}

4. The Taxonomy

So we have two kinds of perception for Leibniz: nonconscious and conscious. If I am right thus far, a nonconscious perception is a world-directed state which is not apperceived, i.e. not accompanied by a HOT. On the other hand, a conscious perception is apperceived and so is accompanied by a HOT. Thus we have:

World-Directed Perceptions: (a) nonconscious, and (b) conscious.

Moreover, Leibniz uses 'apperception' in the way that I have described 'self-consciousness' or 'self-perception' generally. It comes in degrees or levels of 'self-awareness' and, for the reasons given in section 3, reflection (which involves the use of reason) is best understood as the 'deliberate introspection' first discussed in section 1. Recalling the other two levels of self-consciousness, we can thus label our three forms of self-consciousness as follows:

Nonconscious Meta-psychological Thoughts = Apperception\textsubscript{1}

Momentary Focused Introspection = Apperception\textsubscript{2}
Deliberate Introspection = Apperception$_3$ = Reflection

Understanding Leibniz's view about the relation between consciousness, apperception and reflection will help us to become clearer about many of the philosophical problems he addressed.

Kulstad does distinguish between two kinds of 'reflection': simple (or mere) and focused. The former is any reflection "which does not involve a focusing of the mind's attention on what may properly be said to be in us." (p. 24) If we read this as 'mere reflection is the kind of apperception where the mind's attention is not consciously focused on a mental state,' then we would have apperception$_1$. In such a case, one's conscious attention is focused on external objects, which is precisely what the HOT theory demands. Kulstad rightly allows for this when he says that mere reflection involves attention, but "attention [is] directed towards other objects, most typically external objects." (p. 24) This is perhaps not the best fit it could be, but it seems close enough for our purposes.

However, Kulstad's 'focused reflection' is often ambiguous between apperception$_2$ and apperception$_3$ since his primary description is simply "when one's attention is directed at something in oneself." Presumably, then, Kulstad's focused reflection is meant to capture the notion of conscious HOTs, but, as I urged earlier, having conscious HOTs can involve two kinds of ability: momentary focused introspection or the more sophisticated deliberation or reasoning ability. When discussing Leibniz's theory of innate ideas and necessary truths, Kulstad is rightly concerned with apperception$_3$. But it is unclear whether his 'focused reflection' is meant to be identified with apperception$_3$. If not, then we have an ambiguity between apperception$_2$ and apperception$_3$. If so, then he does not properly account for apperception$_2$.

5. Two Troublesome Passages

Some commentators (e.g. McRae) have tried to support the idea that 'apperception' and even 'reflection' just is 'consciousness' for Leibniz. As the reader will gather from the above
discussion, I believe that this is far off the mark. But even those who are more sympathetic to my position must acknowledge that some passages pose problems for the interpretation of apperception as higher-order mentality. Consider the following:

So it is well to make a distinction between perception, which is the inner state of the monad representing external things, and apperception, which is consciousness or the reflective knowledge of this inner state itself and which is not given to all souls or to any soul all the time. (PNG 4; L p. 637)

It seems that some commentators see "...apperception, which is consciousness..." and stop there in an effort to equate apperception and consciousness. But this ignores what follows where the 'or' can plausibly be read as 'that is' or 'in other words' so that apperception is really being identified with "reflective knowledge". My understanding is that the French gives us little help here since the word 'ou' can have many meanings. But, even if I am wrong, my opponents should at most have Leibniz making a disjunctive claim such that apperception can either be consciousness or some kind of self-consciousness (i.e. 'reflective knowledge of this inner state').

On the other hand, Rescher (p. 119) is so biased in my direction that he inserts '[self-]' before "consciousness" in his translation. However, this move is not justified in the original French text. We should still acknowledge the ambiguity in the notion of 'reflective' knowledge of the inner state. We must admit that Leibniz is simply not very clear here about whether he meant apperception1, 2, or 3, but we need not take him to be identifying consciousness in general with apperception.

Consider also the well-known passage from the Monadology:

The passing state which enfolds and represents a multitude in unity or in the simple substance is merely what is called perception. This must be distinguished from apperception or from consciousness, as what follows will make clear.(Mon.
This is probably the most problematic passage for our interpretation, but perhaps we can view the last 'or' as meaning something more like 'and.' This rings true because of the intended contrast: perception...must be distinguished from apperception and (even) from consciousness. Perception must be contrasted with both apperception and consciousness because there can be a 'multitude' of perceptions that never reach the level of consciousness. So Leibniz wanted to distinguish bare perception from both (first-order) consciousness and (higher-order) apperception, and did not merely wish to contrast perception with apperception which just is consciousness. Once again, the French is not very helpful because of the ambiguity of the word 'ou'. Another possibility is that Leibniz is treating apperception as higher-order states which involve consciousness. This would be to say that apperception entails consciousness (which is true), even though they are not identical. Granted that Leibniz was not very careful here, but I do not believe that this one passage should cause us to change our position.

It is also worth noting that Ariew and Garber are so biased in the other direction that it affected their translation. I quote Monadology 14 from Loemker because it remains more faithful to the original Leibnizian text: "...perception...must be distinguished from apperception or from consciousness..." whereas Ariew and Garber insert a comma after 'apperception': "perception...should be distinguished from apperception, or consciousness, ..." (AG p. 214) This makes it seem more likely that Leibniz was identifying consciousness and apperception. The comma, however, is not Leibniz's and they should not have forced this interpretation into the text.19

6. Do Beasts Apperceive?

I have argued at length that most animals are conscious and so are self-conscious (i.e. apperceive) in some sense,20 but more to the point and despite what some commentators have said, I believe that Leibniz also thought so. If my interpretation of Leibniz as a HOT theorist is
correct, then he must have thought that beasts (i.e. non-human animals) apperceive. This is because (a) apperception is necessary for conscious perception and sensation; and (b) Leibniz clearly believed that beasts are conscious. Thus, beasts are at least capable of some form of apperception.

We have already seen the evidence for (a) and how it fits in with the HOT theory. A HOT is necessary in order for one to have a conscious mental state. Thus, although apperception is not identical with sensation, this helps to explain the close connection in Leibniz's use of the terms 'sensation,' 'consciousness,' and 'apperception'. Even McRae (p. 30) is forced to admit that "apperception is a necessary condition of sensation," but then, since he believes that all apperception is restricted to spirits, he accuses Leibniz of inconsistently holding that sensation is present in animals. I suggest that there is no inconsistency for Leibniz, but rather a serious problem for McRae's account. Apperception is indeed a necessary condition of all conscious states and, since both humans and animals have sensations, both humans and animals have apperception.

Moreover, there is a great deal of uncontroversial textual evidence in favor of (b). Leibniz wanted to separate himself from Descartes on animal consciousness, although he did also want to maintain that there is something very special about humans. Beasts are not "simple" or "bare" monads; they are further up in Leibniz's hierarchy. Beasts have souls, i.e. monads with conscious awareness, including distinct perceptions and memory. Leibniz does not wish to attribute "reason" to beasts and so they are not what he calls "spirits" or "minds", i.e. souls that are capable of reason and knowing necessary truths. But this does not affect his view that "...beasts have souls and sensations" (RB p. 72) and he often speaks of "the souls of brutes" (AG p. 78; Letter to Arnauld). And again: "I...believe that beasts have some knowledge and that there is something in them...which can be called a soul..." (L pp. 275-6; Letter to Von Tschirnhaus)

So Leibniz distinguishes between (a) simple monads (monads with only unconscious
perceptions), (b) souls (monads with conscious perceptions and memory), and (c) spirits or minds (monads with the capacity to reason). He says that "[w]hen these beings have sensation they are called souls, and when they are capable of reason they are called minds." (AG p. 191; Letter to Queen Sophie Charlotte of Prussia) We also have the following passages:22

...since sensation is something more than a simple perception, I think that the general name of monad and entelechy is sufficient for simple substances which only have perceptions, and that we should only call those substances souls where perception is more distinct and accompanied by memory. (Mon. 19; AG p. 215)

...sensation, that is, [a] perception accompanied by memory...Such a living thing is called an animal, as its monad is called a soul. And when this soul is raised to the level of reason, it is something more sublime, and it is counted among the minds...
(PNG 4; AG p. 208)23

Furthermore, as Kulstad explains (pp. 19-28), there are passages where Leibniz more directly attributes apperception to beasts. In the New Essays, he speaks of a boar "apperceiving impressions" (A vi. VI, 173; RB p. 173), which certainly sounds like a case of a second-order mental state directed at a first-order state. We also have Leibniz explaining that animals, upon death, are "reduced to a state of confusion which [suspends apperception] but which cannot last forever." (RB p. 55)24 Kulstad rightly observes that "apperception could not be suspended in beasts if beasts never apperceived in the first place." (p. 20)

Despite my argument and the above passages, some (e.g. McRae) insist on holding what Kulstad calls the "standard view"; namely, that beasts do not apperceive. One has to wonder how McRae could possibly say that Leibniz "never once attributes consciousness, apperception, or reflection of any kind to animals." (p. 33) In any case, Kulstad (pp. 41-52) examines five arguments for the conclusion that beasts do not apperceive, which is tantamount to the claim that
only spirits apperceive. I will critically discuss three of them here, since they are particularly relevant to my concerns in this paper. Given the ammunition now at our disposal, we can disarm them without too much difficulty while raising some other issues.

**A. The Monadology Argument**

1. Apperception is consciousness. (Mon. 14)
2. Only spirits are conscious.\(^{25}\)

Therefore, (3) Only spirits have apperceptions.

Premise 1 is clearly false as it stands because apperception is not to be understood as involving outer- or world-directed consciousness. Again, it is true that apperception\(_1\) is necessary for consciousness, but that is not to say that it is identical with consciousness. Moreover, given the threefold distinction within the degrees of apperception, it is a further mistake to identify apperception *simpliciter* with consciousness (see again section 5).

**B. The Reflection-PNG 4 Argument**

We might reconstruct this argument as follows:\(^{26}\)

1. Apperception entails reflection. (PNG 4)
2. Beasts do not have the faculty of reflection.

Therefore, (3) Beasts do not apperceive.

We can now see that premise 1 is false since reflection is best understood as apperception\(_3\). Thus, reflection entails apperception, but not *vice versa*. One can apperceive (i.e. be self-conscious) without having the capacity to reason or reflect in any deliberate way. However, Leibniz clearly believes that premise 2 is true since he repeatedly denies to animals the capacity to reason and reflect (apperception\(_3\)). Leibniz links the capacity to reflect with the ability to reason and learn necessary truths through abstraction. For example: "[beasts] apparently recognize whiteness, and observe it in chalk as in snow; but this does not amount to
abstraction, which requires attention to the general apart from the particular, and consequently involves knowledge of universal truths which beasts do not possess." (RB p. 142)

Leibniz admits to Samuel Masson that he has "denied that beasts are capable of reflection" (AG p. 228) But the real key lies in the ability to reason: "...the knowledge of eternal and necessary truths is what distinguishes us from simple animals and furnishes us with reason and the sciences". (Mon. 29; AG p. 217; cf. L. p. 588) In one place, Leibniz gives some credit to the use of language which "enables man to reason to himself, both because words provide the means for remembering abstract thoughts and because of the usefulness of symbols." (RB p. 275) All of this is perhaps best summarized in the following portion of a letter to Queen Sophie Charlotte of Prussia:

For since the senses and induction can never teach us truths that are fully universal, nor what is absolutely necessary...and since, nonetheless, we know some universal and necessary truths in the sciences, a privilege we have over the beasts, it follows that we have derived these necessary truths, in part, from what is within us. (AG p. 191; cf. L p. 325)

Even if we agree with Leibniz that beasts cannot reason in some sophisticated sense, the question still arises for us: Is having apperception3 the only way for a creature to have a concept of "I" or a "self-concept"? I do not believe so, but sometimes Leibniz seems to think so:

It is also through the knowledge of necessary truths and through their abstractions that we rise to reflective acts, which enable us to think of that which is called "I" and enable us to consider that this or that is in us...these reflective acts furnish the principal objects of our reasonings. (Mon. 30; AG p. 217)

...true reasoning depends on necessary or eternal truths...those who know these necessary truths are...properly called rational animals, and their souls are called
Leibniz is closely linking some rather sophisticated reflective capacities to having a concept of "I". But if self-consciousness comes in degrees, then we should allow that beasts can have more primitive self-concepts. Not only are animals capable of apperception, but it is also reasonable to attribute apperception (i.e. momentary focused introspection) to beasts for several reasons:

(a) Leibniz agrees that beasts are capable of episodic memories of past experiences. But I suggest that if a creature C has an episodic memory of an experience, then that involves C's having a thought about itself, e.g. as the previous subject of some experience. If I remember having an experience of something that happened in my past, then I am having a thought about myself as an enduring subject of experience. (b) We should also allow that any conscious creature can at least differentiate itself from outer objects. This entails having "I-thoughts" and being able to distinguish oneself from other objects. (c) At the very least, it seems that merely having conscious pains and feelings entails having conscious second-order thoughts. If a dog has a conscious pain or emotion, it surely has the higher-order capacity to (consciously) think about the pain or emotion, if only momentarily.

So perhaps Leibniz did not account properly for apperception in suggesting that only minds (and not animal souls) employ self-concepts. We should not suppose that reasoning ability and knowledge of universal truths are necessary conditions for self-consciousness, even though they are sufficient. Perhaps this is all that Leibniz intended, but, if so, then he should have been more careful and simply left open whether or not (animal) souls are also able to have "I-thoughts". What is correct about Leibniz's line of thought is that in order to learn necessary truths we must reflect on our own minds and use reason in the process, which, in turn, involves using self-concepts. But none of this rules out animals from having more primitive forms of self-consciousness, including apperception.
C. The Rationality Argument

Kulstad (pp. 48-51) also discusses a related argument that explicitly mentions the capacity to reason. Let us put it as follows:

(1) Only spirits are capable of reasoning.
(2) Only those souls capable of reasoning are capable of apperceiving.

Therefore, (3) Only spirits can apperceive.

As we have seen in our discussion of the Reflection-PNG 4 argument, premise 1 is true for Leibniz. However, it should be clear by now that premise 2 is false because many apperceiving or self-conscious creatures cannot reason. That is, animals can have apperception\(^1\) and even apperception\(^2\) without being capable of apperception\(^3\).\(^{28}\)

Let us explore more fully why Leibniz thought that beasts do not reason. Perhaps the most important quote is the following:

...beasts are sheer empirics and are guided entirely by instances. While men are capable of demonstrative knowledge, beasts...never manage to form necessary propositions...That is what makes it so easy for men to ensnare beasts, and so easy for simple empirics to make mistakes...The sequences of beasts are only a shadow of reasoning, that is, they are nothing but a connection in the imagination -- a passage from one image to another. (RB pp. 50-51)\(^{29}\)

Thus, for Leibniz, a dog's sequence of mental states is purely guided by particulars and instances. It cannot abstract from instances to grasp general or universal propositions. The dog can connect images in imagination, but this is only a "shadow of reasoning" such that the mind passes from one image to another out of expectation based on memory. Although human minds also often function in this way, Leibniz's point seems to be that animal minds always do:

Memory provides a kind of sequence in souls, which imitates reason, but which
must be distinguished from it. We observe that when animals have the perception of something which strikes them, and when they previously had a similar perception of that thing, then, through a representation in their memory, they expect that which was attached to the thing in the preceding perception, and are led to have sensations similar to those they had before. For example, if we show dogs a stick, they remember the pain that it caused them and they flee. (Mon. 26; AG p. 216)

Beasts pass from one imaging to another by means of a link between them which they have previously experienced. This could be called 'inference' or 'reasoning' in a very broad sense. But I prefer to keep to accepted usage, reserving these words for men and restricting them to the knowledge of some reason for perceptions' being linked together. (RB p. 143; cf. RB 271)

Even if Leibniz is right about some animals, I think it is fair to say that he did underestimate the psychological capacities of many "higher" animals. For example, they often seem to be able to "figure out" what to do in an unexpected or novel situation. Leibniz's example of a dog being repeatedly hit with a stick does not address those cases. Moreover, Leibniz was clearly unaware of the degree to which animals can communicate to each other and even to us.30

Leibniz also does not give us any reason to rule out that some animals can grasp general truths (even if they are not "necessary" or "universal" in some deeper sense). Leibniz's own example could be used as evidence for the claim that the dog has grasped the general truth that "whenever someone hits me with a stick, it is going to hurt" or "whenever someone raises a stick to me, I should flee". Given that Leibniz grants to animals a fair degree of memory, it is hard to see how he rules out that some animals do at least form inductive generalizations from repeated particular instances and so are able to form more general thoughts of this kind. Thus, some
animals may indeed be able to "abstract" from particular instances, even if they fall short of demonstrative knowledge. Of course, none of this may count as genuine "reasoning" or "inference" for Leibniz, but I confess that I am not sure what the "accepted usage" was or is.

In any case, I do agree with Leibniz that "expectation based on memory" does not automatically show an ability to reason or infer. I also agree that humans generally have a capacity to reason and reflect that is absent in all animals. Interestingly, when speaking of possible worlds where there are intermediate species between us and animals, Leibniz asserts that "nature has seen fit to keep these at a distance from us so that there will be no challenge to our superiority on our own globe." (RB p. 473) But recognizing the important objection that there are defective individual humans who are clearly not psychologically superior to every individual brute, Leibniz replies that those humans "suffer from [what] is not a lack of the faculty [of reason] but an impediment to its being exercised." (RB p. 473) The idea seems to be that, say, a mentally defective human still has the faculty of reason but simply cannot use it well or at all. Reason is thus treated as an inherent capacity of human minds even if some humans cannot manifest it due to some illness or defect. I suppose the implication is that we could ideally cure such mental defectives or correct their conditions (e.g. through medical technology) so that their capacity to reason could be manifested. Animals, on the other hand, cannot be made to have the capacity in the first place. Leibniz may be right about this, but I am not sure that it warrants his rather bold claim that "the stupidest man...is incomparably more rational and teachable than the most intellectual of all the beasts." (RB p. 473).

In any case, we see how making a threefold distinction within apperception helps us to reply to the above arguments and shows how we can accept the view that beasts apperceive. We also can see how thinking of Leibniz as a HOT theorist sheds light on his theory of mind.

My interpretation can also help to explain the following potentially troubling passages: (1)...beasts have perception, [but] they don't necessarily have thought, that is, have reflection or
anything which could be the object of it. (RB p. 134) (2)...‘understanding’ in my sense is what in Latin is called intellectus, and the exercise of this faculty is called ‘intellection’, which is a distinct perception combined with a faculty of reflection, which the beasts do not have. Any perception which is combined with this faculty is a thought, and I do not allow thought to beasts any more than I do understanding. (RB p. 173)

These passages have led some to question Leibniz's belief in animal consciousness and thought. But we now know how to reply: Beasts do have higher-order thoughts which accompany conscious states, i.e. they have apperception\textsubscript{1} and even apperception\textsubscript{2}, but they do not have thoughts in the sense of apperception\textsubscript{3}. Leibniz is not denying that beasts have thoughts; rather, he is denying that they have certain kinds of thoughts, i.e. those linked with reflection. In the above passages, Leibniz is clearly linking thought to reflection and understanding.

7. Infinite Perceptions and Kant's "Intensive Magnitude"

In this final section, let us return to the Leibnizian emphasis on degrees of perception. This view stems, in part, from Leibniz's commitment to the Law of Continuity: "nature never makes leaps...any change from small to large, or vice versa, passes through something which is...in between." (RB p. 56) The idea is that nothing happens by jumps and so Leibniz held a staunch gradualism about everything in nature, including minds. He then rightly recognized that this implies an infinity of perceptions:

But since each distinct perception of the soul includes an infinity of confused perceptions which embrace the whole universe, the soul itself knows the things it perceives only so far as it has distinct and heightened perceptions...It is like walking on the seashore and hearing the great noise of the sea: I hear the particular noises of each wave, of which the whole noise is composed, but without distinguishing them. (PNG 13; AG p. 211)
These perceptions, however much they are multiplied, are different from one another, even though our attention cannot always distinguish them, and that is what makes confused perceptions, each distinct one of which contains an infinity because of its relation to everything external. (AG p. 229; Letter to Samuel Masson)

What is Leibniz up to here? First, he is again making room for unconscious mental states. The idea is that a conscious mind can be conscious of or notice only so many of its perceptions at any given time. Our minds have an infinity of perceptions of which we obviously can have higher-order awareness of a relative few (see RB p. 53). Second, Leibniz is recognizing his metaphysical commitment to an infinity of perceptions within each mind. This is due to the fact that each monad "mirrors" all others, i.e. contains all of the information about every other monad. Since there are an infinity of monads, there would need to be an infinity of perceptions. In order to "reflect" all of the information in the universe, each monad must "contain" an infinity of perceptions.

I wish to focus on how this might be related to Kant's notion of intensive magnitude. It is always dangerous to try to summarize briefly anything from Kant's Critique of Pure Reason, but here it goes: Kant was often concerned with the necessary conditions of conscious experience. This led him to argue for the view that some concepts or "categories" are presupposed in consciousness. Two of the categories are Quantity (including the concepts of unity, plurality, and totality) and Quality (including reality, negation, and limitation). Kant was then led to derive a priori principles which can be known antecedently to any particular experience. The principle corresponding to the category of Quantity is the Axiom of Intuition which says that "[a]ll intuitions are extensive magnitudes" (B202), whereas the principle derived from the category of Quality is the Anticipation of Perception which says that "[i]n all appearances, the real that is an object of sensation has intensive magnitude, that is, a degree." (B207) Bennett puts the contrast
as follows:

We are concerned with extensive magnitude when we ask of an object how large
it is, or of a process how long it lasted. We are concerned with intensive
magnitude when we ask how acute a pain is, how loud a noise, or how sour a
drink; the crucial idea...is that of the degree of intensity of some sensation.\textsuperscript{34}

Thus, extensive magnitude has to do with parts (temporal or spatial) and is reflected in Kant's
claim that "[a]ll appearances are...intuited as aggregates, as complexes of previously given parts."
(A163=B204) On the other hand, a sensation "occupies only an instant" (A167=B209) and so an
intensive magnitude is concerned with degrees of intensity and not with extent or parts and
wholes. It is no secret that Kant was very familiar with Leibniz's views and I suggest that this is
one place where he incorporated a Leibnizian doctrine into his theory of mind. Kant repeatedly
speaks of the continuity of degrees of intensive magnitude; indeed, the word 'degree' is
mentioned in the principle itself. He also remarks that "[c]orresponding to this intensity of
sensation, an intensive magnitude, that is, a degree of influence on the sense...must be ascribed to
all objects of perception, in so far as the perception contains sensation." (A166=B208) The key
Kantian idea is that we cannot know in advance what our next sensation will be (for that is an
empirical matter), but we can know a priori that it will come in some degree on a continuum
"between reality and negation".

Kant then goes on to adopt his own mental version of the Law of Continuity. Given any
two degrees of sensation there is always an infinite range of degrees between them, and given any
degree of sensation there is an infinite range of degrees down to nothing:

Every sensation...is capable of diminution, so that it can decrease and gradually
vanish....there is therefore a continuity of many possible intermediate sensations,
the difference between any two of which is always smaller that the difference
between the given sensation and zero or complete negation. (A168=B210) Every
sensation...has...an intensive magnitude which can always be diminished. Between reality and negation there is a continuity of possible realities and of possible smaller perceptions. Every color, as for instance red, has a degree which, however small it may be, is never the smallest; and so with heat...(A169=B211; cf. A172=B214)

This echoes Leibniz's belief that there are an actual infinity of perceptions within each mind and that between any two degrees of perception there must be an infinity. Once again, what is interesting is that Leibniz saw how the Law of Continuity was inseparable from his belief in unconscious perceptions. He speaks of it supporting "the judgment that noticeable perceptions arise by degrees from ones which are too minute to be noticed. To think otherwise is to be ignorant of the immeasurable fineness of things, which always and everywhere involves an actual infinity." (RB p. 57)

Kant also seems committed to such a view if we are to make sense of his position. Bennett (op. cit. p. 176) explains that Kant's view would be false "if we take it as saying that between any two degrees of intensity there is an infinite number of noticeably distinct intermediate degrees; for even if there is an infinite number of pain-levels between that of yesterday's toothache and that of today's, it would be absurd to claim that we can tell every pair of them apart." Kant, like Leibniz, must allow for unconscious sensations or perceptions, which are of course nothing but perceptions of which one is not aware. That is, they are perceptions that minds have but do not notice through any form of self-consciousness. Perhaps Kant and Leibniz's views are not identical. For example, perhaps Kant is committing himself only to the existence of a possible infinite number of perceptions whereas Leibniz is more clearly committed to the presence of an actual infinity of perceptions. However, Kant's use of language is somewhat ambiguous on this point, and it would nonetheless be a contrast between Kant and Leibniz that is worth noting.
This adventure into Kant's theory of intensive magnitudes can perhaps help to explain Leibniz's otherwise often puzzling references (cited above) to perceptions "containing" or "including" an infinity of perceptions. We might interpret Leibniz as claiming that (a) between any two degrees of perception there is an infinity of perceptions, and (b) for any perception or sensation, there is an actual infinite between its degree and its elimination; or, to use Kant's terminology, there is an actual infinite between the "reality" of any sensation and its reduction to "negation". 35

Perhaps this is an odd way to speak of distinct perceptions "containing" or "including" anything, but if this is not what Leibniz meant, then I have no idea what he did mean. One might understandably object that Leibniz often speaks of each distinct perception at a single moment containing an infinity of perceptions (cf. RB p. 53 and, again, the passages cited above). Leibniz's manner of speaking may cause trouble for my Kantian interpretation, but again I would then simply confess that I do not understand what he meant. However, it is tempting to adopt McRae's analysis that at least some sensible perceptions (e.g. color) "are not aggregates of insensible perceptions. Rather, they are novel emergents from a mass of insensible perceptions, emergents for an apperceptive mind which is incapable of distinguishing components." (MCR p. 38)

I suppose that more complex visual perceptions could be understood as composed of an infinity of unconscious perceptions, while one's conscious mind can only focus on or apperceive a few at a time. For example, one's visual field could be broken down into an infinite number of regions, but at any moment one can only attend to a small portion of it. Perhaps one could say that my current visual perception of my office "contains" or "includes" an infinity in that sense. In any case, all of this assumes that we can make sense of an actual infinite in the first place, but I will not open that troublesome can of worms here. 36
NOTES


3. It should be noted that I am mainly concerned here with an analysis of the locution 'x (mental state) is conscious.' This ought to be distinguished from 'x (an organism) is conscious' as well as the two-place predicate 'x is conscious of y.' For some discussion of this terminological matter, see CSC pp. 3-5.

4. In CSC.

5. See CSC pp. 17-18 for several additional reasons.


8. One might reasonably ask: "Just what makes a higher-order thought 'suitable'?" A full answer to this question would lead to a lengthy digression which I cannot pursue here. One condition, for example, would be that the HOT must be a 'momentary' and 'occurrent' state as opposed to a 'dispositional' state of some kind. See CSC, chapters three and four, for my attempt at answering the above question.

9. The terminology here can be a bit confusing. Sometimes thought is contrasted with perception whereas sometimes the term 'thought' is also used as a generic term covering virtually all kind of mental states. For our purposes, we can think of the higher-order state as some kind of higher-order awareness. For some discussion of the alleged differences between the 'perceptual' and 'thought' models, see CSC pp. 95-101. Leibniz was also aware of some of the terminological difficulties in RB pp. 171, 210. See also CSC (pp. 36-43) for at least one attempt to address them in the context of the HOT theory.
10. cf. RB pp. 53-57, 77-78, 113-119, 161-162 and 188. I will also abbreviate the frequently cited Monadology as 'Mon.' and the Principles of Nature and Grace as 'PNG' followed by the section number.

11. For a particularly inexcusable example of this type of error, see Peter Carruthers (1989) "Brute experience", Journal of Philosophy, 86, pp. 258-269. See my 1993 reply to Carruthers in "Brute experience and the higher-order thought theory of consciousness", Philosophical Papers, 22, pp. 51-69.

12. Nicholas Rescher's well-known 1979 book is Leibniz: An Introduction to his Philosophy (Rowman and Littlefield). Hereafter I will refer to it as 'RES' and all page references to Rescher will come from this book.


14. For some discussion of memory and self-consciousness, see CSC chapter 9 or my 1992 "Consciousness, self-consciousness, and episodic memory", Philosophical Psychology, 5, pp. 333-347.

15. Given his agreement with Rescher, it is unclear to me what to make of Kulstad's distinction between "apperception of external objects" and "apperception of what is in us" (KUL pp. 133-143).


17. I thank Stephen Voss for some help and advice on these linguistic matters.

18. See Robert McRae (1976) Leibniz: perception, apperception, and thought (University of Toronto Press). I will hereafter refer to this book as 'MCR' and all page references to McRae will come from it.

19. I thank Jonathan Bennett for some helpful comments on material in especially sections 3 and 5.
20. See CSC, especially chapters one through four, eight and nine.

21. For a critical discussion of Descartes' views, see Margaret Wilson's (1995) "Animal Ideas," Proceedings and Addresses of the American Philosophical Association, 69, pp. 7-25. For a critical discussion of Spinoza on animal mentality, see "'For they do not agree in nature with us': Spinoza on the lower animals," this volume.

22. I thank Jan Cover for some helpful comments on the differences between simple monads, souls, and minds or spirits.

23. cf. PNG 14; AG p. 211 and Mon. 82; AG p. 223.

24. The reason for the square brackets is (recall from section 3) that Remnant and Bennett do not use the term 'apperception'. But see G, V, 48 as an alternative source.

25. See KUL pp. 43-44 for a discussion of this premise. I have nothing to add except that it clearly seems false given Leibniz's frequent attribution of consciousness to souls as well as spirits.

26. See KUL pp. 18, 26-27, 41-42.

27. I discuss each of these at much greater length in CSC, especially chapters four, eight and nine.

28. Mark Kulstad ultimately replies to these arguments in a similar fashion (see e.g. KUL p. 171). However, as we saw in section 4, it does remain a bit unclear how his twofold distinction between simple (or mere) and focused reflection fits in with my three degrees of self-consciousness. This is important because it has an effect on his list of possible responses to the standard view (KUL pp. 19, 27). We can adopt something close to one of his possibilities; namely, that "there are [at least] two senses of 'apperception' in the Leibnizian text, in [at least] one of which it is correct to say that beasts apperceive, and in one of which it is not." (KUL p. 19)

29. cf. RB pp. 73, 81, 173, 180, 475.
30. For more on animal mentality, see CSC and Donald Griffin (1992) *Animal Minds* (University of Chicago Press).

31. This could be used as a way to handle the familiar charge from animal rights advocates that we cannot, based purely on psychological capacities, grant all humans a right to life and deny such a right to all (or most) animals. I cannot delve into this ethical issue here.

32. The other two arguments for "the standard view" have to do with "personality" and the theological problem of why God would create innocent beings capable of suffering (KUL pp. 44-48). I believe that we can respond to them in a similar fashion, i.e. by properly distinguishing between degrees of apperception, but I cannot pursue a full discussion of them here.


35. If the reader is interested in Kant's theory of mind, please see CSC where I extensively relate his views to the HOT theory and discuss them in light of contemporary theories of consciousness.

36. Some of the work on this paper was supported by the National Endowment for the Humanities through Jonathan Bennett's 1995 Summer Seminar entitled "Descartes, Spinoza, and Leibniz." I would also like to thank Jonathan Bennett, Susanna Goodin, Eric Sotnak, and Stephen Voss for comments on an earlier draft of this manuscript.