

Intentionality and the connection principle

Karl Pfeifer
Department of Philosophy
University of Saskatchewan

Fourth Annual Conference Of Western Canadian Psychologists
University of Saskatchewan
October 3, 1992

Intentionality and the connection principle

Ending one of his early papers on intentionality, John Searle (1979) once averred that he believed only beings capable of conscious states are capable of intentionality, but that he didn't know how to "demonstrate" this. Soon thereafter followed "Minds, Brains, and Programs" (Searle 1980), with its notorious Chinese room Gedankenexperiment, in which he maintained against certain factions of the Artificial Intelligentsia that implementing a program per se cannot ever be sufficient for having mental states. For mentation, the system in question must also have the right causal capacities. Human brains and similar biological systems have such capacities, digital computers do not. Although the issue of consciousness was not explicitly raised in that paper, one can safely presume that the causal capacities Searle had in mind would at least subsume capacities for conscious states. Now, in "Consciousness, explanatory inversion, and cognitive science" (Searle 1990), written for the tenth anniversary of "Minds, Brains, and Programs", the issue of consciousness comes to the fore. Basically Searle's current hobbyhorse is that functionalist or computationalist theories of mind, in seeking to elucidate our cognitive capacities, uncritically help themselves to the notion of unconscious mental processes, including processes by their very nature inaccessible to consciousness. Searle contends that no proper demarcation has been drawn between unconscious but nonetheless mental phenomena and nonconscious nonmental, brute neurophysiological phenomena in the brain; he contends

moreover that a proper demarcation logically requires accessibility to consciousness for unconscious mental states. If Searle's arguments succeed, they would at last constitute the elusive "demonstration" mentioned in the early paper.

However, while I am sympathetic to the idea of mental states generally having some sort of essential link to consciousness, I do not think that Searle's conception of the connection is entirely satisfactory. In what follows I will expose some of the soft spots in his account and briefly suggest an alternative.

"The ascription of an unconscious intentional phenomenon to a system implies that the phenomenon is in principle accessible to consciousness," says Searle (1990), introducing his Connection Principle. The adjective "intentional" here is not meant to be read as the ordinary-English near-synonym for "deliberate" but as a cognate of the technical term, "intentionality", for that common property of mental states of being "about" or "directed at" objects or states of affairs. If one has a belief, for example, it is always a belief that such-and-such is the case; if one has a desire, it must be a desire for something; and so on. Not all mental states are intentional in this sense, however. Tickles or pains, for example, aren't "about" anything in the way that beliefs or desires are. On Searle's view intentionality merely partitions the mental: intentionality is a sufficient but not a necessary condition for something's being a mental state or process. We should note as well that nonintentional mental states such as itches and pains are necessarily conscious states; so when we are speaking of unconscious mental states we are eo ipso speaking of intentional states.

Searle also insists on a firm distinction between intrinsic and "as if" intentionality.¹ Although I might say that my lawn is thirsty just as I

might say ~~that~~ that I am, the former is merely a figurative way of describing the lawn's capacity to absorb water, whereas the latter reports an intrinsic mental state in me. As-if intentionality, Searle adds, is not really a type of intentionality at all, but merely a type of attribution.

So why would the proper demarcation between unconscious mental and nonmental states require that the former be essentially accessible to consciousness? Searle claims that intentional states, whether conscious or unconscious, always have "aspectual shapes". Aspectual shape is part of what makes an intentional state the intentional state that it is. As Searle explains it,

Whenever we perceive anything or think about anything, it is always under some aspects and not others that we perceive or think about that thing. . . . A man may, for example, want to drink a glass of water without wanting to drink a glass of H₂O. . . . Notice also that the aspectual shape must matter to the agent. It is, for example from the agent's point of view that he can want water without wanting H₂O. In the case of conscious thoughts, the way the aspectual shape matters is that it constitutes the way the agent thinks about or experiences a subject matter: I can think about my thirst for a drink of water without thinking at all about its chemical composition. I can think of it as water without thinking of it as H₂O. (p. 587)

But Searle also maintains that the ontology of unconscious mental states, at the time they are unconscious, consists entirely in the existence of purely neurophysiological phenomena, and that aspectual shape cannot be

¹The distinction he makes in "Minds, Brains, and Programs" between original and derived intentionality can be regarded as a special case of this. The as-if intentionality of computers reflects and hence "derives" from the purposes of their designers, programmers, and users.

characterized in those terms, or more generally in any such third-person terms. If Searle is right in these claims, there seems to be a problem: How can unconscious mental states have aspectual shape and yet consist entirely in physical phenomena? How can they have aspectual shape if there is no aspectual shape at the level of neurons and synapses?

Searle's way out is to argue that the notion of an unconscious intentional state must therefore be the notion of a state that is a possible conscious thought or experience. And so, qualifying the previous ontological claim, he concludes that the ontology of the unconscious consists in objective features of the brain capable of causing subjective conscious thoughts (p. 588).

Does this work? I see two problems. The first has to do with the distinction between intrinsic and as-if intentionality.

One way of characterizing Searle's view is to say that the language of unconscious intentional mental states is largely a matter of transferred epithets. Although Searle continues to use the same idioms most of us do -- he talks of unconscious mental "states" that are "accessible" to consciousness, that are possible "contents" of consciousness, etc. -- these idioms are not to be taken literally. Such states and contents have about as much existence and accessibility as the proverbial average man. To use Searle's own metaphor, mental states cannot be pulled out of the unconscious the way fish are pulled out of the sea. So as far as mental states are concerned the "pulling out" is really a "coming to be". At bottom the notion of unconscious mental states is just a fancy way of talking about causal capacities or dispositions of the brain to generate certain conscious states. And although such capacities or dispositions are in a sense features of the brain, they themselves are not what is accessible to consciousness;

indeed, a person might not believe that a belief that just consciously occurred to him is one that he unconsciously "had" (i.e. had the disposition to consciously have) all along, even if it happens to be true.

It appears that our language of mentalistic attribution for even the unconscious states of preferred status has many of the earmarks of what Searle disparages as as-if intentionality.

To be sure, the states of the brain responsible for the relevant dispositions are real, as are the dispositions themselves. But why should the intentionality ascribed to any of this be intrinsic rather than as-if, given the undeniable figurativeness of the language of ascription? Compare:

Case 1. I wish to obtain the solution to a numerical problem, so I punch the buttons on my calculator, and get a display on its little screen which makes me think, "Aha, the answer is 42."

Case 2. Like case 1, except that the numerical problem puts me to sleep, dreamlessly, or at least as pertains to the problem, and I wake up thinking, "Aha, the answer is 42."

Handheld calculators presumably only do as-if calculation. Yet if we tell the story about the causal sequence between wish and wish-fulfillment in case 2 in the right sort of way, not much more need be going on in the brain as pertains to the so-called calculation, than is going on in the calculator. For example, imagine

Case 3, in which the calculator is attached directly to my brain with

wires and electrodes. My wish for the numerical problem's solution generates neural impulses that have the same effect as pushing the calculator's buttons and the calculator's output pulses produce in me the thought, "Aha, the answer is 42."

We can make the parallel even more forceful if we imagine

Case 4, which is like case 3 except that the calculator is miniaturized and implanted in the brain itself, perhaps as a prosthesis replacing a damaged area.

Assuming that case 2 is a case of unconscious but by Searle's lights genuinely intentional calculation, is case 4 properly assimilated to it or to case 1? Searle could respond that the intervening sequence of brain processes in case 2 is integrated with other states and processes, namely those that sustain consciousness, in a way that allows the unconscious sequence itself to be accessed by consciousness, whereas the intervening sequences in cases 3 or 4 are encapsulated. Such a response would indeed be on target since cases 3 and 4 merely suggest modules that accept inputs and outputs but whose internal intervening states are insulated from direct interaction with surrounding brain states and processes. But I submit that there is no reason there couldn't be a

Case 5, which is like case 4 except that the inner processes of the calculator are not sealed off.

Just as the calculator of case 1 need not be entirely a black box to us --

it could be in a transparent case, voltage meters could be connected to its circuits, etc. -- so too could the embedded calculator of case 4 have various wetware connections to its innards in a way that rendered its operations penetrable to brain states and processes that sustain consciousness, and thereby allowing those operations potentially to be in some way or other conscious.

The difficulty I am trying to raise is this. Given Searle's view, two exactly similar unconscious sequences of inner states, mediating between exactly similar conscious inputs and conscious outputs, might nonetheless have different ontological status as regards having or not having aspectual shape, merely because of differences in how other processes may or may not -- not even "do or do not"! -- act on them. All the deciding factors reside outside the sequences in question, for in each case its contribution is the same; there is no difference in either sequence per se that makes for a difference. Why then should one have intrinsic aspectual shape and not the other? Or, to take the flip side, why should one of these have merely as-if intentionality and not the other?

One conceivable countermove here might be to insist that unconscious mental processes have to be viewed more holistically, say, as also including the states and processes that make the appropriate consciousness possible. The problem with this, however, is that it is not likely that many of these are themselves accessible to consciousness, so the propriety of including them as a difference-making part is dubious. Moreover, going this route would seem to inject even more scope for vagueness than there already is in the dichotomies under consideration.

The second problem we should consider is how purely neurophysiological phenomena could have the causal capacity to produce conscious states

having such-and-such intentional features, in the first place. Notice for starters that what such a causal capacity is for must already have an intentional characterization, and notice also that a capacity so-characterized preexists its manifestations, if any. Presumably the difference between having a given capacity for producing certain conscious intentional states and not having it requires a difference in neurophysiological features that correlate with the various aspects (as it were) of the aspectual shape of the conscious state that is the manifestation of the capacity. So the neurophysiological phenomenon that constitutes an unconscious intentional state must already hold some analog of that intentional state. Anything else would be magic.

One reason Searle insists on a strong distinction between intrinsic and as-if intentionality is that without it everything in the universe would have intentionality. I don't think that would necessarily be so bad. At least it would seem to be no worse than sheer emergence.

I wish to propose a way of providing continuity between the merely physical and the intentional. Elsewhere (Martin & Pfeifer 1986, Pfeifer 1988) I have maintained that the usual characterizations of intentionality are incomplete because ordinary physical causal dispositions and capacities already exhibit many of the features regarded as distinctive of intentional states and processes. Let me give a couple of quick examples. Intentional states are said to be directed at objects or states of affairs which may not even exist ("intentional inexistence"), now or ever. But this is also a feature of dispositions, since dispositions are after all dispositions to affect or be affected by, or to engender, certain objects or states of affairs which may not actually exist, now or ever. Searle's (1979) broad notion of satisfaction conditions for intentional states also has its analog in physical causal

dispositions. Just as intentional states may be satisfied or not insofar as beliefs are true or false, desires are fulfilled or unfulfilled, etc. so dispositions may be satisfied or not insofar as they are manifested or not. This is not to say that mere physical dispositions or capacities are on an equal footing with fullblooded intentional states, but it does suggest that intentional states might be a special kind of dispositional state. If this is so we could have the underpinnings for the emergence of intentionality at a certain level of neurophysiological complexity -- it would be (in part) a matter of interlocking and nested dispositions, with intentionality consisting in summations of quasi-intentional features of a complexly ordered network of dispositions of neurophysiological states and processes.

So far, this account need not be inimical to Searle. At this point however he would want at least one rider, allowing as intentional only those summations that include a disposition toward conscious manifestation. This would be (inter alia) on account of the aspectual shape which makes an intentional state the state that it is and which is supposed to be essentially tied to consciousness; in order for nonconscious neurophysiological states or processes to be deemed intentional they must have at least an indirect dispositional connection with consciousness.

But again, I think aspectual shape is a feature exhibited by dispositions and capacities. Dispositions are keyed to specific kinds of objects or states of affairs; the relata of a dispositional relationship always figure in that relationship qua bearers of particular properties, namely properties in virtue of which the disposition is the disposition that it is. If this is so then aspectual shape will occur all the way down, and the subtle highlevel discriminations of aspectual shape selfconscious agents make can again be

plausibly viewed as summations of cruder aspectual shapes at various neurophysiological levels. Searle's argument that aspectual shape cannot be exhaustively characterized in terms of neurophysiological predicates does not touch my claim that the dispositional is aspectual. If, as Searle contends, neurophysiological predicates (among others) cannot characterize aspectual shape because of an "inferential gulf" between the epistemic grounds for the presence of something and that something's ontological nature (p. 587, step 3), then by the same token they (or similar physical predicates) will not be able to characterize dispositions either. So to that extent intentionality and dispositionality are on a par.

Searle's insistence that unconscious intentional states are essentially capable of being conscious can be seen to be even less motivated if one reflects on the implications for animals other than ourselves and a few higher mammals.

Animals too are conscious, to the extent that they have sensory experiences. Insofar as this sensory experience selectively projects onto objective features of their environment (i.e. has environmentally informed aspectual shape) it enables them to behave adaptively. A frog, for example, perceptually detects things in virtue of their falling under the aspect "small dark moving object" (or somesuch) and fortunately for it enough of these will turn out to be edible. All this can be described dispositionally as well, and indeed there is nothing spectacular about the aspectual shape here that would elevate it above the humdrum property-keyedness of common physical dispositions. Although aspectual shape figures in the frog's perception, the frog is not even potentially conscious of aspectual shape. Nor does the aspectual shape as such "matter" to the frog; what matters to the frog, although it has no inkling of this, is getting enough to eat.

If it is a requirement of the Connection Principle that aspectual shape itself be accessible to consciousness, then it follows that froggy perceptual consciousness is not intentional. I find this implausible (cf. Doty 1987; Ewert 1987, p. 390), as likely Searle would too. Moreover it suggests that Searle's Connection Principle is an artifact of the highlevel human examples he restricts himself to, inasmuch as in the human case consciousness of propositional attitudes such as beliefs and desires is eo ipso consciousness of their aspectual shape.

Like Searle I balk at the prospect of panpsychism, the view that mentality pervades everything; unlike Searle, I would hesitate to equate that with "panintentionalism". However this appears to be largely a matter of taxonomical preference. One can regard intentionality as coming in various grades à la Dretske (1981), instead of taking it to be criterial for the mental or some subdomain thereof. This need not be a matter of confusing as-if intentionality with the real thing. We can agree with Searle that rocks, trees, and even certain chunks of the brain don't literally have beliefs and desires, but that need not force the conclusion that these don't have intentionality either. An alternative is that whereas the intentional vocabulary of beliefs and desires is just too highgrade to apply to rocks etc., some other more basic intentional vocabulary (directedness, satisfaction conditions) could still apply. But I won't quibble over taxonomy. Whether mere physical dispositions are called intentional or quasi-intentional matters not as long as it is acknowledged that the highgrade capital-I intentionality humans take so much pride in arises naturally out of complex substrata of crude dispositions, prefiguring in cruder form some of the capital-I features.

If we want a principled demarcation of the intentional, limiting it to

where the literal use of psychological terms is appropriate, I think that Searle's requirement of the capability of consciousness draws too heavily on the peculiarities of the human case, as the frog example indicates, and overly emphasizes the effect side of causation. How then should we mark off unconscious intentional states from nonconscious dispositional states at large? My own suggestion, which mayhap extensionally approximates Searle's, is that we regard as intentional those dispositions (or summations of dispositions) involving consciousness in their etiology or in their exercise (cf. Pfeifer 1987). This characterization allows for differences in intentional sophistication while restricting intentionality to creatures with perceptual organs.

Karl Pfeifer

University of Saskatchewan

References

Doty, Robert W. (1987) Has the greedy toad lost its soul; and if so, what was it? Behavioral and Brain Sciences 10: 375.

Dretske Fred I. (1981) Knowledge & the Flow of Information. Bradford Books / MIT Press.

Ewert, Jörg-Peter. (1987) Neuroethology of releasing mechanisms: Prey-catching in toads. Behavioral and Brain Sciences 10: 337-368, 389-405.

Martin, C.B. & Pfeifer, Karl. (1986) Intentionality and the non-psychological. Philosophy and Phenomenological Research 46: 531-554.

Pfeifer, Karl. (1987) Causal dispositions + sensory experience = intentionality. Behavioral and Brain Sciences 10: 757-758.

Pfeifer, Karl. (1988) Chisholm on expressions for intentional relations. Philosophical Studies 53, 153-156.

Searle, J.R. (1979) What is an Intentional state? Mind 88: 74-92.

Searle, J.R. (1980) Minds, brains, and programs. Behavioral and Brain Sciences 3: 417-

Searle, J.R. (1990) Consciousness, explanatory inversion, and cognitive science. Behavioral and Brain Sciences 13: 585-596, 632-642.