BEHAVIORAL AND BRAIN SCIENCES (1990) 13, 585–642
Printed in the United States of America

Unintended thought and nonconscious inferences exist

James S. Uleman[a] and Jennifer K. Uleman[b]
[a]*Department of Psychology, New York University, New York, NY 10003* and [b]*Department of Philosophy, University of Pennsylvania, Philadelphia, PA 19104*
Electronic mail: *uleman@nyuacf.bitnet*

What distinguishes "mental" events from other phenomena, and what does this tell us about cognitive science? These are important questions. Useful answers will have to include a wider range of putatively cognitive phenomena, however, and use a more differentiated set of analytic categories than that provided by Searle.

We agree with one of Searle's conclusions: that the idea of intentional mental states which are "in principle" inaccessible to consciousness is incoherent. The distinction between "intrinsic" and "as-if" intentionality is important. One of us has written as much (Uleman 1989, especially pp. 430–34). But we are dissatisfied with the way he reaches this conclusion, the implications he draws from it, and the limited categories he provides.

Searle argues that neither computational machinery nor "brute, blind" neurophysiology can be intrinsically intentional, precisely because they are not *conscious*. The claim that plant leaves and the VOR (vestibular ocular reflex) lack intrinsic intentionality is the same, for him, as the claim that when they "act" they have nothing in mind, they follow no rule, they are not conscious.

To develop this argument, Searle introduces aspectual shapes. These are (a) uniquely characteristic of consciousness, (b) impossible to account for exhaustively with "objective" evidence ("third person, behavioral, or even neurophysiological"), and (c) not even detectible from such evidence (cf. sect. 6,

para. 5 on Quine and zombies). This prevents equating intrinsic and as-if intentionality. No amount of looking or acting intentional warrants our calling a system intentional. We must (somehow) know that it performs consciously. (One might wonder whether aspectual shapes aren't just the sort of ethereal apparition that commit Searle to a dualism he wants to avoid; see Note 4.) This tack has the unfortunate effect of obscuring how or why we attribute intentionality to each other, because surely we have only "objective" information about each other.

Searle identifies unconscious intentional states with neurophysiology, given a puzzle – the sleeping man who knows the capital of Colorado – that might equally well be solved by dropping the assumption that intentional content (aspectual shapes) must be located in a head, and cannot reside in shared linguistic practice or other recurrent features of the environment. Dropping this assumption is tantamount to abandoning the view that intentionality must refer to some incorrigibly subjective feature of the agent. But the "subjectivist" tendencies of Searle's analysis cannot be pursued further here.

Is it correct, however, to insist that a given process, to be *mental* at all, need be in principle accessible to consciousness? Are there phenomena (of social cognition and cognitive psychology) to which Searle's categories cannot do justice?

First, there is Searle's assumption that "it is reasonably clear how [an intention] works for conscious thoughts and experiences." But Nisbett and Ross (1980) have shown that it is not at all clear how intentions work. People's best attempts at describing how they make intentional judgments (what rules they followed, or which factors they considered) are often wrong. They may offer plausible but demonstrably inaccurate theories, as when they "intentionally" judge how influential someone is in a two-person discussion, but are in fact responding only to that person's salience in their own perceptual field (Taylor & Fiske 1975). Or they may intend to perform some cognitive operation, being incapable of doing so but not knowing this, and so believing in the end that they have actually succeeded "intentionally." For example, beliefs may persist even after the basis for them has been discredited and the believer thinks he has taken the discrediting into account (Ross et al. 1975).

Are these cases merely neurophysiological, or are they accessible "in principle" to consciousness? They are neither, but are clearly mental events. Searle offers only a dichotomy, however: "There are brute, blind neurophysiological processes and there is consciousness; but there is nothing else" (sect. 3, para. 1).

Consider a second set of cases, illustrated by the Stroop effect in which people must quickly name the color of the ink while ignoring the printed words' meanings, and are unavoidably slowed down by the unintended activation of those meanings. Or consider priming effects, in which supra- or subliminal exposure to relevant concepts biases subsequent intentional interpretations of ambiguous behavior. (See Bargh, 1989, for a review and conceptual suggestions.) Is concept activation merely a blind neurophysiological process that influences conscious intentional processes? And does such a classification enrich or diminish our understanding of the phenomenon? Again, what can we do with skilled typists' or musicians' performances, in which behavior that is initially guided intentionally is replaced by relatively automatic routines that are *disrupted* when the performer tries to make them conscious?

Third, this scheme denies the possibility of "unconscious inferences." Yet when people read about actions that imply personality traits, they infer those traits without any intention or awareness of doing so; they do so although these inferences are not called for, and may even interfere with another primary concurrent task. (See Uleman, 1987, for a review of cued-recall results; we also have unpublished evidence from increased recognition reaction times, using McKoon & Ratcliff's [1986] procedure.) Are these somehow not inferences, but mere neurophysiological processes? Or will the gray area to which Searle assigns our attributing intentions to grasshoppers also include such cases? How much expansion can this marginal zone withstand before it alters the distinction between intrinsic and as-if intentionality?

Fourth, Searle apparently assumes that intentional processes operate in the same way whether or not they are actually conscious. But we (Newman & Uleman, in press) and others have shown that primed concepts have different effects on subsequent judgments, depending on whether they are conscious when the judgment occurs. Furthermore, the extensive literature on implicit memory (e.g., Schacter 1987) contradicts this assumption.

Only space limitations preclude extending this list of phenomena that are mental events but are not intrinsically intentional. In throwing out the bath water of as-if intentionality, Searle has also dumped a tubful of babies that cry out for recognition and coherent organization.