BOOK REVIEWS

WILLIAM LYONS. ***The Disappearance of Introspection.*** Cambridge, Mass.: A Bradford Book/ The MIT Press (1986), xv + 193 pp. $22.50 (cloth).

What, if anything, is introspection? "'Introspection'" Ryle observed, "is a term of art and one for which little use is found in the self-description of untheoretical people." Lyons quotes this passage approvingly, but underestimates the extent to which this fact might confound his efforts. Since introspection is not only a term of art, but an ill-defined one, it has meant different things to different theorists. According to Lyons, theorists have thought, confusedly, that introspection is a "special and privileged executive monitoring process" (p.113), and their efforts to accommodate this special phenomenon in their theories of the mind/brain have led them to "bizarre" solutions, "fantastic and incredible suggestions brought on by assuming that 'introspection' is a meta-process that monitors first-level occurrences of perception, memory, imagination, thinking, and so on" (p. 123). There is no such meta-process, according to Lyons; what has been confusedly called "introspection" is just a certain variety of "operations by perceptual memory and imagination when their operations are at the level of conscious attention" (p. 124).

The first two parts of the book are devoted to criticism of earlier theories, and there can be little doubt that Lyons' quarry, as he portrays it, is a conception we should all wish to disappear. First, Lyons patiently presents the history of the introspectionist psychologists—Wundt, Titchener and their rivals—and their downfall; this is followed by an account of the behaviorist successors—Watson, Skinner and their rivals—and their downfall. Much of this material is unfamiliar to philosophers, and it adds a refreshing perspective to stale debates. Then he turns his critical attention to the philosophers of mind: Armstrong's idea of introspection as brain-scanning, and my own elaboration of Putnamian functionalistic ideas of introspective reports as the expression in "printouts" of logical states of the brain's "program". His criticisms of these proposals are then taken to clear the decks for his own positive account, which occupies the third part of the book.

Lyons makes many good observations about the phenomena in question, and the shortcomings of various theories of them, but he is confounded by his insufficient appreciation of a problem with *all* the literature on the topic (his own and my own included): too much reliance on metaphor and handwaving. In order to compose the opposition into targets suitable for criticism, Lyons must time and again make assumptions about just how literally these theorists meant to be taken, and about just how they would extend their accounts to other cases not explicitly discussed. And time and again, in my estimation, he ends up imputing more folly to his predecessors than they actually exhibited, though their own reliance on handwaving and metaphor invites just such reconstruction. In the process of rendering his rivals explicit and literal on the issues that concern him, he pushes them into extreme ("true blue") positions on such issues as reduction versus elimination, the properties of mental images, privileged access, and the way subjects' reports might be informative about their brain states. The bizarre, fantastic, incredible theories he then attacks are thus to some extent his own creations, descendants of theories that sometimes tackle issues with which his own positive account does not come to grips. This leaves him, at best, replacing a seriously confused idea by a vague and noncommittal idea.

Lyons' positive theory is that what is called introspection is actually "replay" (always in scare-quotes) of processes of perceptual memory and imagination. We are not given a theory of the operation of perceptual memory and imagination, nor any account of what their being "at the level of conscious attention" might come to. Suspicions arise about how any theory of these residual phenomena would avoid the pitfalls of the bad old visions.

Lyons asserts: "We do not need to 'introspect' in order to say what we are looking at, hearing, tasting, smelling, or feeling" (p. 150). But what do we do, in order to say these things? Surely *any* theory of how we say what we are looking at, etc., must posit *some* "meta-level" activity to accommodate the obvious facts about the phenomena. For in-

stance, on such occasions we offer (as Lyons himself often puts it), an "edited" or "interpreted" version of what we are looking at (etc.)—and can, for instance, refrain from issuing a report on details of our experience that are not deemed worthy of report—but in order to achieve this circumspection, this capacity to "edit" or "interpret" or "censor", something "meta" must be going on. If Lyons thinks a theory can be given that denies this, he offers us no reasons to agree with him.

Again, privileged access takes some heavy blows in the early chapters, but in the end Lyons ends up endorsing a variety of privileged access—incorrigibility about content, but not about veracity of content (pp. 130–132)—that is not so "thin" as he makes out; it is (apparently) the very notion of privileged access that he applauds me for abandoning (pp. 77–78). And Lyons' own way with imagery has at least as much handwaving in it as any other current version, for example: "What is not stored are 'pictures'. . . . What is stored is an abstraction from the myriad perceptual data " (p. 135).

Lyons' book thus makes contributions to philosophical work on introspection at two levels. At the first, intended level, he makes many acute observations on the pitfalls and false temptations of theory, and on the nature of the phenomena. On another, unintended level, his book is an eloquent witness to the futilities of investigations carried out at such a level of handwaving that the participants have difficulty recognizing when they are agreeing with each other. *Daniel C. Dennett, Tufts University.*

STEPHEN E. BOËR AND WILLIAM G. LYCAN. ***Knowing Who.*** Cambridge, Mass.: A Bradford Book/The MIT Press (1986), xii + 212 pp. $22.50 (cloth).

Readers of the previous work of these authors (especially those familiar with Bill Lycan's work) will expect detail and clarity combined with an alarmingly omnivorous appetite for the literature. *Knowing Who* is no exception, though potential readers should realize that the technical development of the semantics is forbiddingly dense in places. Hence this books is no easy read. But aficionados of the propositional attitude semantics industry will not want to miss it.

The book focuses on a particular species of propositional attitude ascriptions: locutions of the form "*X* knows who *N* is" where *N* is the proxy for a name, a description, or a demonstrative. But the book is of wider interest than this would suggest, for three central chapters develop a general theory of the semantics of propositional attitude sentences. The *primary* interest of the work is semantic rather than psychological, that is, on the truth conditions of knowing-who locutions rather than the nature of the psychological states those locutions report (though the latter is not entirely neglected). I will start by sketching the theoretical background, and then outline the intuitive idea of the authors' treatment of knowing-who. (I have neither the space nor the competence to assess its technical development.)

The central innovation of the book is to combine and develop two familiar ideas: the "two factor" theory of attitude individuation and the "paratactic" analysis of attitude sentences. The two factor theory of attitude ascription holds that we group (for example,) belief tokens into types according to two different but co-equal criteria, depending on our theoretical purposes. If we are interested in the explanation of behavior, we group tokens into types by virtue of similarity of conceptual role. If our purposes are the explanation of success or failure of behavioral plans, we group tokens into types via identity of the belief tokens' truth conditions. Lycan has for some years defended the importance of this approach by showing how it provides a natural solution to Kripke's puzzle about belief (4.2), the nature of beliefs about ourselves (6.1), and how it helps unravel the tangles about "de re" belief (5.3). Indeed, one virtue of the book is its bringing these discussions together in an accessible place. The opacity of propositional attitude inscriptions in general, and knowing-who constructions in particular, is a reflex of our two different schemes of attribution. Attributions guided by the criterion of conceptual role are necessarily opaque, for two terms being co-referential is notoriously insufficient for their having the same conceptual role.

The second idea is to take over, develop, and formalize the Davidsonian analysis of