Book review

Relational frame theory: A post-Skinnerian account of human language and cognition
Edited by Steven C. Hayes, Dermot Barnes-Holmes & Bryan Roche

Ever since Chomsky’s (1959) review of Skinner’s Verbal Behavior (1957), which played a seminal role in promoting a cognitive approach to human behaviour, critics of behaviourism have often assumed that their case is closed and won. Behaviourism is an evolving research program, however, and may well develop resources unanticipated by its critics. Although most recent advances in behavioural theory have been concerned with operant processes in non-human animals (e.g. Davison & McCarthy, 1988), Relational frame theory: A post-Skinnerian account of human language and cognition is an attempt to address verbal behaviour and human cognitive phenomena within a renewed behavioural paradigm.

Over 13 chapters, the book presents and defends an approach to cognition known as relational frame theory (RFT). The first eight chapters emphasize the origins of RFT (Chapter 1), its principal concepts (Chapters 2 and 8), its account of derived stimulus functions (Chapter 3), and its relevance to linguistic behaviour and thinking (Chapters 4 to 7). The remaining chapters, which the authors describe as ‘more speculative’ (p. 155), discuss possible applications of the theory to development, education, social processes, psychotherapy, and religion. Although chapter authorship varies, the contributors (whom I refer to as ‘the authors’ for brevity) always include one or more of the editors, and this book presents a unified viewpoint rather than a collection of independent essays.

Some readers will not be familiar with the experimental research reported here, which has been published mainly in behaviour-analytic journals. RFT research originates in studies of matching to sample, in which participants are trained to select combinations of stimuli among various available ones. Once a few particular selections have been trained, choices among novel stimulus combinations often emerge in a predictable way, a finding that is difficult to explain on the basis of traditional reinforcement concepts. The stimuli related by such emergent choices may also display new, untrained behavioural functions. Thus, a stimulus matched, perhaps quite indirectly, to an aversive stimulus may itself become aversive (as when being told that ‘smoking causes cancer’, for example).

Some behaviour analysts have seen such findings, reminiscent of productivity and systemativity, as promising ones for addressing cognitive aspects of human performance. Over the last 20 years, basic human research in behaviour analysis has branched into a variety of directions (e.g. Critchfield, Galizio, & Zentall, 2002), a conspicuous one being RFT. Relational frame researchers have been quite clever at developing novel laboratory preparations that allow emergent stimulus functions to be studied; RFT work in progress may thus involve relations of comparison among stimuli, identity, temporal sequencing, part-whole relations, etc. (Chapter 3), as well as their effects on emergent behaviour.
One of the distinctive features of RFT as a theoretical approach is that it tries to explain emergent performance by expanding, rather than sidetracking or discarding, the concept of operant reinforcement. In RFT, emergent phenomena are conceived as patterns of generalized operant responding or ‘frames’ (e.g. pp. 25 - 29). Because the functions of most of the stimuli that participate in these frames have never been trained explicitly, RFT goes beyond the crude reinforcement-based accounts that have been criticized (e.g. Mahoney, 1974) for their reliance on direct training as the primary means of behaviour change. From the perspective of RFT, most human behaviour is governed by networks of relational frames. The authors correctly criticize Skinner’s approach (pp. 16 – 17, p. 104) as providing deficient explanations (or labels instead of explanations when the latter were found to fail). Relational frame theorists try to develop process-based accounts, in which distinctively human abilities are explained and not just labelled. Another merit of the book is that—with one exception which I’ll discuss later—the authors discuss relevant research from the cognitive literature and aim at a broader scholarship than has sometimes been the case in behaviour analysis. The dedication of the authors is evident, and their methodological contributions interesting. More troubling is the philosophical style in which their theory is described. RFT is supposed to be guided by the data themselves (‘the data will guide us’: p. 255) and supported by numerous experiments, ‘all with confirmatory results’ (p. 43). As to whether RFT actually provides the analysis they are aiming at, however, the authors admit that the empirical support ‘is still too limited to know for sure’ (p. 255). They insist that the theory is naturalistic (p. 251), coherent (p. 254), and precise (p. 145). Its terminology is careful (p. 89). The authors emphasize that ‘RFT explains what cognition is, rather than relying on common sense terms or examples’ (p. 228); indeed, part of the RFT material may be ‘arcane and esoteric’ (p. 1). Above all, RFT embraces ‘contextualistic thinking’ (p. 34).

The authors promise an enticing program, but does it deliver the goods? Closer inspection reveals logical difficulties in the concepts used in RFT and the explanations it proposes. Remember one main technical problem that RFT attempts to address: How can we behave in the presence of objects or events (A) as if they had a relational property that, in fact, they do not have? Behavioural theorists of any persuasion will appeal to our previous experience with other objects or events (B) that did have the relational property in question, and somehow affect our performance with respect to the novel objects (A). The question is how. Here, RFT assumes that what seems to be a relational property transferring from B to A (e.g. same-as or more-than) is actually a behaviour or behaviour class (p. 32, p. 33, p.88, p. 145, p. 255). The authors thus state that a stimulus relation (same-as, bigger-than, etc.) involves ‘responding to one event in terms of the other’ (p. 29), and that this type of ‘relational response’ can be ‘brought to bear’ on other relata (p. 32) to generate novel performance.

But there usually are no such relational responses. When a child is trained to point at the object identical to a particular sample, for instance, the child's response is pointing and not ‘same-as,’ and qualifies as ‘relational’ (p. 25) only by courtesy. What is relational is the stimulus that governs pointing (namely, the relation of identity holding between the two objects), and not the child’s response to it. The dilemma for RFT is that whereas the same-as stimulus is relational, not being a response it can’t be ‘brought to bear’ on novel tasks (cf. p. 25), still less be reinforced or punished. In contrast, the response of pointing could certainly be reinforced, punished, or emitted in novel tasks; but this response isn’t relational and can’t be a relation of same-as. Either way, the RFT account fails.

Perhaps the authors do not literally mean that the stimulus relation of same-as is a behaviour, however. Rather, they might mean that after a response is emitted conditionally on a stimulus array A, emitting the same response (whatever it is) can be reinforced in the presence of another stimulus array B; so that the reinforced ‘response’ becomes emitting-in-the-presence-of-B-whatever-was-emitted-in-the-presence-of-A. Similarly, the behaviour of more-than might consist in emitting a higher number of elementary acts in the presence of a stimulus A than in the presence of another stimulus B—and the authors indeed appear to defend such an interpretation (p. 25).
However, neither that interpretation nor the previous one can deal with some of the crucial findings that RFT is supposed to explain. For example, in a study by Dymond and Barnes (1995) illustrating the notion of relational frame (p. 60), the emergent behaviours that the present authors attribute to ‘relations’ of more-than or less-than (p. 61) did not consist of emitting more or less responses than in previous circumstances. The behaviours emitted in testing consisted instead of selecting arbitrary symbols as if some antecedent stimuli were ‘more’ or ‘less’ than others, none of these stimuli being actually ‘more’ or ‘less’ than any other stimulus. What the authors could mean here by responses of more-than and less-than is a puzzle.

Of course, if the authors abandoned all claims to develop a behavioural theory, then they could easily make sense of their own approach by admitting that the ‘responses’ they postulate are not responses at all, but internal, symbolic processes. From the viewpoint of classical cognitive science, for example, frames of more-than, less-than, after, before, etc., can be brought to bear on novel relata to generate adaptive performance (Minsky, 1975). These frames, however, are not behaviours or even classes thereof, but cognitive structures; and these are ‘brought to bear’ not on external stimuli, but on internal representations of external stimuli (by variable binding). In this context, it is unfortunate that the authors nowhere mention, still less discuss, the cognitive concept of frame.

As the chapters unfold and the topics broaden, the RFT treatment becomes less precise and, contrary to what the authors claim (p. 228), their analyses seem to proceed more by common sense and example than by documenting how the frames they invoke could be behavioural. That the theory fails in well-controlled experimental cases (such as Dymond & Barnes, 1995) does not bode well for its interpretations of development, social issues and psychopathology – although it is difficult to know for sure, since the behavioural nature of the relevant ‘frames’ is largely assumed instead of being demonstrated. Thus, the RFT approach to cognition freely appeals to responses of ‘part-whole’ and ‘attribute of’ (p. 37), behaviours of ‘I and You, Here and There, and Now and Then’ (p. 122), and relational acts of ‘IS, TIME, BECAUSE, BETTER, ATTRIBUTE, and PERSPECTIVE’ (p. 220). Later, however, the authors assimilate ‘relating events’ to ‘thinking’ and add that the former ‘is not enough to produce action’ (p. 229).

In the Preface, the authors claim that the phenomena they address have not been ‘prematurely forced into a behavioral mold’ (p. xiii), but the evidence suggests otherwise. Logical difficulties should, after all, be expected when operant reinforcement concepts are pushed beyond their natural limits. The RFT attempt at explaining novel behaviour in terms of operant reinforcement is suspect on logical grounds, since reinforcement, by definition, explains the maintenance of performance but not its origins. However important the issue of maintenance is, it cannot substitute for understanding the ways in which actual stimulus relations affect behaviour across time (e.g. Rescorla, 1988) and induce the performance on which reinforcement eventually works. Whether a technical, coherent treatment of human cognition in behavioural terms is possible remains to be seen, but it is doubtful that operant reinforcement will occupy in it the place it occupies in RFT.

FRANCOIS TONNEAU
(Centro de Estudios e Investigaciones en Comportamiento, Universidad de Guadalajara, México)

References
