

## *Behaviourism*

The minds of other people are frustratingly private to us. We guess what other people are thinking, and are sometimes suspicious of their sincerity and motives. Other aspects of nature are largely open to inspection, and science has carefully delved into them with experiments, measurements, and the spotting of mathematical patterns. It thus occurred to psychologists to focus on those aspects of mental life which are observable and measurable, and that can only mean behaviour. Thus psychology became a highly empirical subject, and philosophers began to explore the implications of their approach.

Logical positivism is an empiricist quasi-scientific doctrine, asserting that (apart from fairly empty tautologies) language is meaningless if it has no grounding in experience. By connecting this proposal with the empirical methods of behaviourist psychology, the thought emerged that the only meaningful assertions that relate to the privacy of mind are those that relate to behaviour, and that therefore statements about the hidden mind can either be eliminated, or reduced to statements about behaviour. '**Logical** behaviourism' claims that all mental language reduces to behavioural language. '**Eliminative** behaviourism' (the boldest form of the doctrine) denies the existence of anything other than behaviour in our accounts of the mind. Thus the problem of other minds is dissolved. We even know our own minds entirely through awareness of how we behave. A more modest version is '**black box** behaviourism', which concedes that there is an inner mental life, but that it is unknowable, apart from the behaviour the box produces. The weakest form of the doctrine is '**methodological** behaviourism', which just says that observation of behaviour is the best (and perhaps only) scientific route to understanding the mind. If 'the mind just is behaviour' seems too strong, 'the mind can only be defined behaviourally' may seem more plausible. That fits with the modern idea that epistemology should be an empirical science, rather than a priori philosophy. The behaviourist approach is attractive because it offers the prospect of a clear and precise account of the mind, and of removing the mystery from obscure features of the mind such as 'aboutness', the 'content' of thoughts, and 'reference' in language.

However, introspection suggests that mental life is fast-moving, highly varied, vivid and unpredictable, whereas a quietly seated person exhibits very little observable behaviour. Actual behaviour is a very inadequate replacement or reduction for the mind, and behaviourism is only plausible if it invokes potential as well as actual behaviour. Thus if the seated person is hungry, or doing arithmetic, or is remembering something, this is reducible to the behaviour which would result, if the stimulus were food, or the request for an answer, or a reminder of the past event. It is true that we get to know both ourselves and other people by seeing how they might respond, perhaps in an interview. A person is brave only if they have a **disposition** to respond boldly to danger. This can all be expressed as 'if stimulus x occurs, then behaviour y will result', and a mind is a map of the resulting vast set of conditional truths.

A further problem is that there does not seem to be a straight correlation between specific types of mental events and specific behaviour. People are unpredictable, even when you know what they are thinking. There is certainly no necessary connection between the inner mind and the outer behaviour. People don't grab at food as soon as they feel hunger, or book a plane flight as soon as they think of Paris. To get round this problem, behaviourism needs to be more **holistic**. That is, if we feel hungry, we have other thoughts, beliefs and feelings at the same time, and the result is not a single disposition but a complex of drives, sometimes united and sometimes in conflict. This complexity makes prediction and necessary connections impossible, but still allows a theoretical identification of the tangle of mental dispositions with the much simpler behavioural output. If behaviourism is to be defended, it can only be in this dispositional and holistic form.

So far the theory has focused on the person to be explained, but a slightly different story arises if we attend to the observer of this person. When we encounter other people, what do we think is going on in their minds? The truth is that though we refer to their beliefs and feelings, we are merely observing nuances of their behaviour. A variant of behaviourism (the '**intentional stance**') says that this talk of beliefs and feelings is a fiction. The language of 'folk psychology' (our everyday terms referring to mental life) has no basis, other than patterns of behaviour. Even if playing chess against a computer, we can't help attributing these fictional mental states to it. Mental language is a tool for predicting behaviour, and doesn't pick out any facts (because interior facts are too complex and obscure).

In recent times behaviourism has ceased to be a popular doctrine, in either psychology or philosophy, and the strong versions have few supporters. A number of tricky cases have arisen, which don't translate easily into behaviour, such as specific images (in memory or imagination), the qualities of an experience (such as pain or colour), mathematical thought, and our sense of values. If someone had 'inverted qualia' (experiencing red and green traffic lights with the experiences swapped), that might be undetectable in the behaviour. Physical pain may produce predictable behaviour in humans, such as facial expressions, but what is presumably a very similar feeling in reptiles results in a different output, so pain can't be the same as our wincing and grimaces.

The main challenge, though, is when thought and behaviour can be pulled apart. Good actors can produce behaviour (such as terror or love) when the actual feeling is absent. A strong-willed person can totally suppress the behaviour provoked by a feeling such as pain or greed, and a totally paralysed person might have a mental life, but virtually no behaviour. If mental events can have no behaviour, or behaviour have no mental events, it seems unlikely that they are the same thing.

The most general and perhaps most important objection concerns what the theory misses out. It is intended to give a complete account of mental life, but it confines itself to observable patterns of behaviour, and refuses to speculate about inner mental life (inside the black box). It thus not only offers no explanation of behaviour (which may be wise caution), but also asserts that there is no explanation. We take the contents of our thoughts (whether or not they are physical) to have important causal powers, but behaviourism ignores this. The scientific commitments of behaviourism remain very influential, but the details of the theory face very large obstacles.