PREPRINT: for *Australasian Journal of Philosophy*

Joshua Shepherd

ICREA (Barcelona, Spain), Universitat Autònoma de Barcelona (Belleterra, Spain)

**Review: Wu, Wayne (2023). *Movements of the Mind: A Theory of Attention, Intention and Action*. Oxford University Press.**

Wayne Wu’s book is about mental action, and how various psychological elements fit together to explain mental agency. The book is full of interesting claims and arguments; it contains incisive treatments of a wide range of empirical work; it should be read by philosophers of action (who often give psychology minimal attention) and philosophers of cognitive science (who often give action minimal attention).

Part one of the book (chapters one and two) explicates and updates Wu’s views on attention and action, attention’s necessary role for action, and relationships between control, guidance, automaticity, and intention. Part two (chapters three and four) offers an unusual and interesting treatment of intention as both agentive activity and as practical memory (discussed below). Part three discusses the nature of bias and the role of biased attention in skill (chapter five), the role of attention in deductive reasoning (chapter six), and the nature of introspection and attention’s role in introspection (chapter seven). The various parts of Wu’s views come together in complex ways – I do not have space to articulate the full edifice here. But what drives much of Wu’s picture is a view on intention (discussed more below), and its relation to attention. Intention informs the way that an agent couples inputs to action with possible behavioral outputs, and it does so by biasing attention, which is (roughly) the selection of targets to guide action. Attention, for Wu, is ubiquitous across bodily and mental action, often operating automatically.

Wu has written a book that brings the mind sciences to bear on the philosophy of action. The book illustrates how views in the philosophy of action can benefit from, and might require, sustained engagement with science. Indeed, if one wants to engage with many of Wu’s interesting views and arguments, one has to develop opinions on how to interpret various empirical paradigms, and on the state of scientific knowledge regarding a range of issues. This is difficult work – for examples of philosophers who do just this to criticize Wu’s view of attention, see Buehler (2022), Lange (2023) – but it is, in my view, the way forward.

Let me illustrate this thought by discussing two aspects of Wu’s book. The first has to do with Wu’s ideas about control and automaticity. Wu often works by offering explications, or technical analyses, of central terms, and then building upon these. He is explicit about this, writing that ‘We should stop using central notions non-technically,’ (32) and urging readers ‘to take the analysis [of technical terms like automaticity and control] on board or do better’ (33). By ‘doing better,’ it seems Wu has in mind answering the issues that he takes to be central to accounts of action and agency – issues that draw on (necessarily selective) interpretations of fragments of scientific work. So, for example, Wu’s analyses of control and automaticity are motivated by the thought that, while intentional action often exemplifies control and automaticity, cognitive scientists often write as if ‘control implies the absence of automaticity’ (28). How to square the joint presence of control and automaticity in action? Wu’s answer is that control and automaticity are not kinds of processes, but features of processes. An action A exemplifies control when the agent A-s because she intends to A; the unintended features of action are automatic.

This is one way to dissolve the tension, but it raises questions about whether this is the only way to understand control and automaticity, and about whether those whose discussions of these notions are motivated by alternative problems are mistaken, or simply discussing a different topic. One might, for example, reject the crude dichotomy of control and automaticity as setting a guide for understanding these notions. In my view, rich literatures on control, automaticity, intention, skill, knowledge, and action identify dimensions of these notions that might not be best addressed by only one technical usage of a given term, but instead by exploring multiple relationships between these notions (Pacherie and Mylopoulos 2021, cf. Christensen 2021), and the ways that action is supported by a range of process-types and psychological capacities. To better see how Wu’s explications fit into this broader space, I wish Wu had engaged more with these literatures: the absence of a more substantive conversation between Wu and scholars working on these issues seems a missed opportunity.

The second aspect of Wu’s book I wish to discuss is arguably the most interesting and innovative: his treatment of intention in part two (chapters three and four). Wu has it that intention should not be understood in the traditional way as a state or event, but rather as a form of activity, what he calls *practical memory for work*. This activity consists in a few things: [a] in the maintenance of information over the course of an action, [b] in the setting of attention in a way that guides agents’ response to the selection problem, that is, the problem of how to appropriately couple inputs, or ways of taking things, to behavioral outputs, [c] in modulating what Wu calls vigilance, which is linked to the strength with which attention is sustained over time, or directed to alternative dangers or behavioral targets, [d] in modulating what Wu calls steadfastness, which helps the agent balance single-mindedness with flexibility, or openness to task-irrelevant stimuli.

How can intention do this work, and why think of it as a form of memory? Wu closely links intention to the science of working memory. This is a construct that, while defined in many different ways, is posited to explain the capacity to maintain and manipulate action-relevant information in real time, supporting sophisticated thought and action (Baddeley 1992). For Wu, the science of working memory covers sub-personal cognitive operations directed towards intention implementation and the fine-tuning of intention, and these sub-personal operations can be mapped to the personal-level activity of intending. He writes that ‘the dynamics of working memory underwrite the dynamics of intention’ (109), that intention should be understood ‘as a type of memory via its algorithmic underpinnings in the working memory construct’ (109), that there is a ‘functional convergence between intention and the central executive’ (106), which is a posited part of the working memory system, and that ‘the central executive realizes the structuring role of the agent’s intention in setting appropriate (cognitive) attention’ (106).

As mentioned above, additional work for intention (and thus working memory) consists in mediating levels of vigilance and steadfastness. Wu argues that in addition to attitude and content, intention contains an added dimension which he calls ‘activeness’ (118). The activeness of intention can be seen at work, he argues, in research on inattentional unawareness. Consider a study by Conway et al. (2001), in which participants receive distinct verbal streams in each ear, and are told to ‘verbally shadow’ (or repeat) words in one stream, while their name is uttered in the other stream. Participants with high working memory capacity (WMC) heard their name much less than those with low WMC – they were much better at remaining focused on the right stream. By contrast, if told to listen for their name in the other channel (Colflesh and Conway 2007), while remaining focused on the original channel, high WMC participants were much improved at hearing their name, and did much better than low WMC participants. Wu writes: ‘The idea is that differences in the agents’ being active is reflected in greater vigilance between groups even as the content of the intentions of the two groups remains the same, fixed by the same task instructions’ (117).

I am not convinced that activeness is a dimension of intention. Those with high and low WMC may not have the exact same intentions – experts often display subtle differences in the content of their intentions, due in part to the refined nature of their action concepts and the efficiency of their strategies (a point I make elsewhere, Shepherd 2021). Other discussions highlight levels of motivation as well as skill at various components of mental activity, like attention (Shepherd 2017), or general capacities for information management (Murray and Vargas 2020).

Whether the differences revealed in inattentional awareness studies should be tied to differences in intention depends, in part, upon whether one accepts the claim that working memory is the sub-personal implementation of intention. I have some doubts. I agree that the science of working memory can helpfully illuminate aspects of intention, but I also think that tying the nature of intention closely to the science of working memory over-estimates the unity and cleanliness of work that appeals to this construct (Gomez-Lavin 2021). When executing an intention, it seems agents can intentionally manipulate what Wu’s calls ‘activeness’ by putting in more effort, or by planning to perform certain cognitive control operations as a part of their intention. When I say ‘it seems,’ I am not appealing to intuition. The sciences of working memory and cognitive control are full of behavioral paradigms that have agents performing intentional actions operationalized as cognitive control operations of task switching, inhibition, maintenance, and so on. Certainly some operations of working memory can be called sub-personal – though I think deploying the personal/sub-personal distinction in philosophical argumentation requires care (Westfall 2022), and is indeed typically best avoided – but the sciences of cognitive control and working memory describe operations that cross the personal/sub-personal divide.

There is, of course, much more to say, and it is a virtue of Wu’s book that this is so. The book lays out an array of arguments regarding core issues at the intersection of the sciences and philosophy of agency. It is sure to move these areas of inquiry forward.

**Acknowledgements**

Thanks to Myrto Mylopoulos and Eric Mandelbaum for discussion.

**References**

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Buehler, D. (2022). Seeing circles: Inattentive response-coupling. *Ergo, an Open Access Journal of Philosophy*.

Christensen, W. (2021). The skill of translating thought into action: framing the problem. *Review of Philosophy and Psychology*, *12*(3), 547-573.

Colflesh, G. J., & Conway, A. R. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic bulletin & review*, *14*, 699-703.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, *8*, 331-335.

Gomez-Lavin, J. (2021). Working memory is not a natural kind and cannot explain central cognition. *Review of Philosophy and Psychology*, *12*(2), 199-225.

Lange, V. (2023). Decentering and attention. *Philosophical Psychology*, 1-28.

Murray, S., & Vargas, M. (2020). Vigilance and control. *Philosophical Studies*, *177*(3), 825-843.

Pacherie, E., & Mylopoulos, M. (2021). Beyond automaticity: The psychological complexity of skill. *Topoi*, *40*(3), 649-662.

Shepherd, J. (2017). Halfhearted action and control. *Ergo* 4(9), 259-276.

Shepherd, J. (2021). *The Shape of Agency: Control, Action, Skill, Knowledge*. Oxford University Press.

Westfall, M. (2022). Constructing persons: On the personal-subpersonal distinction. *Philosophical Psychology*, 1-30.