

To appear in *Journal of Consciousness Studies*

Please cite and quote from published manuscript

Précis of Movements of the Mind

Wayne Wu

In *Movements of the Mind (MoM, (Wu 2023a))*, I give a theory of agency that uncovers its internal psychological structure, revealing how creatures with minds do things. While my focus is on things we do “in our heads”, mental actions, the theory concerns all forms of agency. The book also provides a theory of attention and its essential connection to action. It characterizes intention in action as a type of memory for work, drawing on empirical theories of working memory. Further, it reveals the essential role of bias in action, providing an account of various psychological biases. The theory is then applied to three phenomena prominent in philosophical practice: implicit (better automatic) bias, reasoning deductively, and introspection of conscious perceptual experience.¹

In this overview, I briefly highlight a subset of the themes in each chapter with the hopes of enticing philosophers and scientists that there is something for them in the book that would repay engagement.

The book is divided into three parts. Part 1 (Chapters 1 – 2) gives a theory of action and attention in action; Part 2 (Chapters 3 – 4) presents the theory of intention as practical memory or memory for work, drawing on the theory of working memory to reveal intention’s dynamics; Part 3 (Chapters 5 – 7) apply the theory to implicit bias, deductive reasoning, and introspecting perceptual experience. This application aims to show that the theory can be productively applied across many types of action and, I hope, will serve as a roadmap and inspiration for philosophers to apply the theory to their domains of interest. While I argue for some strong claims in the book, in this précis, I present weakened versions of certain theses that should be broadly acceptable, embody the perspective of my approach and do substantive philosophical work. This allows readers to (hopefully) profit from the book even if they reject my stronger formulations.

In Chapter 1, I present what I take to be a metaphysically necessary feature of action, namely that it solves a *Selection Problem*. This Problem is constituted by action-

¹ Many topics are discussed along the way. Here’s a list, hopefully evocative: the necessity of automaticity, attending as an action, vigilance and attunement as readiness to attend, the agent’s steadfastness in action as expressing intention, the activity of control, the necessity of memory in intentional action, cognitive integration of intention with action capacities, how intention keeps time with action through practical reasoning, the activity of intention as active memory and as an action of remembering, attention as biased in skill and in implicit bias, cognitive attention in non-ampliative reasoning as the focusing of attention, an action-driven solution to Carrol’s paradox of inferential rules, the psychology of introspection, types of introspection and reliability conditions. This is done with sensitivity to relevant and well-established empirical work that I try to present in an accessible fashion.

To appear in Journal of Consciousness Studies

Please cite and quote from published manuscript

possibilities, with each action coupling a mental “taking” (e.g. an attitude or experience) as input and a response as output. The response could be a movement of the body or it could be a movement of the mind such as an encoding in memory. The standard Many-Many Selection Problem can be illustrated by beginning with n mental inputs ($n > 1$) and m response outputs ($m > 1$) where each n can be mapped to multiple m 's and vice versa. The Selection Problem can then be posed as a question: what is to be done? The answer is the execution of a possible action in the Problem space.

Figure 1: The Many-Many Problem as Selection Problem

Caption: This is a 2 x 2 Many-Many Problem where the subject mentally takes in targets A and B (e.g. the subject perceives or remembers them). Each can be mapped to two possible responses. To act, at least one of the action options, indicated by an input-output mapping given as a possible action due to the dotted lines, must be taken.

It is in this context that I introduce *bias*.

Figure 2: The Selection Problem and Bias

Caption: A bias is an internal factor that prioritizes one action possibility over others and, if that action is taken, explains why the action occurred. There can be multiple biases. In this case, the bias, in influencing A and the output Φ , prioritizes the A- Φ action.

A bias influences how the Selection Problem is solved. There are many kinds of bias, and more than one can be involved in influencing solutions to the Selection Problem, these working in concert or in conflict. Intention is one notable bias (e.g. the intention can be to do action A- Φ in figure 2), but there are others (see Chapter 5 of *MoM*). A weak version of my approach is that for actions of interest to theorists, these actions arise from solving a Selection Problem. If so, the framework I present can organize theoretical explanation of notable actions and their consequences.

I also argue for a technical notion of automaticity needed to characterize action. This is driven by an unnoted paradox of automaticity.

1. Intentional actions exemplify agentic control.

(1) identifies a core thesis in philosophy of action. It is what we are trying to explain. At the same time:

2. Intentional actions exemplify pervasive automaticity.

To appear in Journal of Consciousness Studies

Please cite and quote from published manuscript

Take reaching for a glass of gin (or is it petrol?). While I intentionally reach for it, many of the properties that individuate that reach as a distinct movement are automatic, say its kinematics. Similarly, on imagining a pink elephant, while my entertaining that image is intentional, the direction I imagine the animal facing might be set automatically. Trouble arises when we conjoin (1) and (2) with a fundamental assumption in psychology.

3. Control and automaticity are incompatible.

This is best expressed in the common division between two kinds of processes, namely automatic processes versus controlled processes. (1), (2) and (3) appear incompatible. Indeed, intentional action as characterized by (1) and (2) seem to be counterexamples to the psychological principle. This is the paradox.

Nevertheless, the three can be rendered consistent so long as we drop the standard reading of (3), and relativize automaticity and control to *features* of processes. Accordingly, control and automaticity are incompatible, but only because the same *property* of an action exemplified at a time cannot be *both* automatic and controlled. This gives a precise account of automaticity, roughly that a feature *F* of an action *A* at some time or temporal range *t* is automatic iff *A*'s having *F* at that time is not controlled, that is, is not due to the agent's intending to perform an *F* action (subject to all the caveats about action causation). The details are in the text, but the point is that automaticity is pervasive, it is philosophically important, and we need a precise analysis. The paradox motivates an intuitive account that I hope readers will consider using.

I have written much on attention as *selection for action*. A subject's attending to some target *T* is the subject's mentally selecting that target to guide response in action. Here, I will again note a weaker thesis that most theorists could accept: in cases of attention of theoretical interest, attention in that case is the subject's mental selection of a target for response.

This weaker position is entailed by James' famous expression of the folk conception of attention, what we all know: the mind's selection of a target among others in order to deal with. It is also the grounds for all the experimental paradigms for probing attention, given in an *empirical sufficient condition* of attention for each task (Wu 2024): if a subject *S* mentally selects target *T* to guide task performance, the subject mentally attends to *T* (e.g. visual selection/attention). The point is that the conditional expresses an intersection between folk and empirical psychology. Selection for action in this limited sense should be everyone's starting point for thinking about attention. These issues are discussed in Chapter 2.

Chapters 3 and 4 introduce a new perspective on intention: it is a type of practical memory or memory for work. In Chapter 4, I consider a thought experiment. Since memory is subject to loss (amnesia), consider having the technology or power to induce the agent to forget what she is doing as she does it, perhaps through neural disruption. If we could erase

To appear in Journal of Consciousness Studies

Please cite and quote from published manuscript

an attitude to induce such amnesia, what would we target? I suggest the agent's intention. So as the agent intentionally performs her action, obliterate her memory by abolishing her intending to so act. Perhaps for a short duration of time, the agent's agentive momentum will propel the intended action forward. Yet if we ask her as she acts why she is doing that, she might answer: I'm doing *A*? She has forgotten, after all, what she intends to do. Think of the common case of walking into a room, intentionally to do *F*, but having momentarily lost track of one's intention. One has forgotten what one was going to do.

Miller, Galanter and Pribram (1960) introduced the construct *working memory* into psychology to identify a specific mnemonic function: *the retention of an agent's plans in order to implement them*. One has to remember what one plans, after all, if action is to begin and continue. This connection was a "scales falling from the eyes" moment for me. It opened up a vast literature on working memory as a way to probe the dynamics of intention. This seems to me a fruitful and underexplored line of thought that I hope readers will be inspired to explore.

In psychology, working memory is seen as a dynamic cognitive workspace. I take it to be a substrate of an agent's intention in action, with the dynamics of working memory storage reflecting the dynamics of the agent's intention, specifically the dynamics of the targets of intention. Working memory is the basis for ongoing sensitivity in intention to relevant parts of the world, updating as the world changes. These changes, reflecting a *fine-tuning* of intention, embody a type of practical reasoning, and as the intention is refined over the course of action, the agent's access to her action through intention keeps up with the action through such reasoning.

In the last part of the book, I apply the theory. In Chapter 5, I consider biased attention, especially automatic attention. Recall that bias in my theory is the factor (often multipole) that explains the agent's action as solution to a Selection Problem. Implicitly biased actions are just one type of biased actions. I argue that how attention is biased is a salient basis of many implicitly biased behaviors of social, ethical, epistemic and political concern. For example, we know of negative biases in selecting questioners during question-and-answer (Q&A) sessions after an academic talk. This is seen in biased queues. For example, one skew favors faculty over students. Many disciplines have recognized this as an undesirable outcome. To build the queue, the moderator performs a visual search task, an attention task, so the skew begins with biased visual attention. Further, departments have addressed this skew by instituting a rule as counterbias such as the requirement, often performatively uttered at the beginning of Q&A, that students will get the first few questions. This leads to intentionally biased visual search to mitigate a latent implicit bias that has negative consequences. I also discuss how attention can be biased in positive ways, focusing on the case of radiological skill in visual search. In both cases, the bias can be taken to be implicit because the resulting biased behavior is not intended and subjects need not have access to those biases. The broader point is that attention provides a fruitful target for understanding implicit, or as I would prefer,

To appear in Journal of Consciousness Studies

Please cite and quote from published manuscript

automatic biases, precisely because attention is well-understood at multiple levels of empirical and philosophical analysis.

In Chapter 6, I turn to deductive reasoning and explicate the idea of cognitive attention in detail. I show how prominent psychological models of attention can be understood as hypothesizing how cognitive attention narrows in deduction understood in terms of the agent's selective focus on the contents of premises. Further, I discuss how we learn to do proofs as based on joint attention that helps sharpen focus on relevant logical form, and that this knowledge involves the acquisition of schemas that influence solving the Selection Problem by biasing appropriate attention.

I want to highlight, however, how understanding action can secure insight into thorny philosophical problems. Lewis Carroll posed a paradox of rules in inference. One paradox is psychological. How does logic compel us to draw the appropriate conclusion in, say, modus ponens? One answer is that we know the rule of modus ponens, but in Carroll's story, knowledge of the rule is assimilated by treating it as a premise to guide drawing the appropriate inference. This just iterates the problematic situation, inviting the same question about logical compulsion, necessitating another rule as premise, and so on. A standard response to this is to argue that rules are implicit, yet this seems to me to be an evasion. After all, *sometimes* we use rules explicitly, and these just are the problematic cases.

Here, understanding the structure of action helps for while knowledge can provide premises, inputs to the Selection Problem, these being potential targets of cognitive attention during inference, knowledge can also serve as a *bias on* attention. That is, the role of knowledge is to focus attention to enable appropriate inference. In doing so, it does not provide the rule as premise. If so, Achilles could have responded to the pesky tortoise by refusing to write down the known rule in the logical exercise book, for doing so would be to conflate a bias on attention with a target of attention. Of course, Achilles was not a philosopher of action, so the needed action perspective was not available to him.

Chapter 7 emphasizes how the structure of action can illuminate introspection of perceptual experience. What is needed is a model of introspection as a mental action. I am concerned that scientists and philosophers have for too long given introspection a pass, taking its deliverances at face value. I do not know how much of our introspective data is accurate. Much of it might be inaccurate. I emphasize that I am *not* an introspective sceptic, Introspection is no different than perception in that in some contexts, it can be reliably accurate, in others, not. What remains unknown is whether the introspective data that drives much theorizing in the theory of consciousness is gathered in unreliable, inaccurate contexts (for my concerns about introspective data as support for arguments for unconscious vision and the rubber hand illusion, see (Wu 2020; 2023b)).

I illustrate this by giving two structures for introspection of perceptual consciousness, what I call simple introspection and complex introspection. Simple introspection relies only on

To appear in Journal of Consciousness Studies

Please cite and quote from published manuscript

one attentional channel, grounded in the perceptual experience. Complex introspection, in not being simple, draws on multiple attentional channels, and that is where mischief and distortion can happen, for example the undue influence of theoretical beliefs or expectations. I illustrate the complexities by focusing on the introspection of blur, data from which has been used to argue about the metaphysics of perceptual consciousness. While philosophers have been confident and uniform in introspecting that peripheral vision is not blurry, I find that among all types of audiences I have posed the question to during talks, introspective responses are not uniform. Many subjects affirm that peripheral vision is blurry (I'd estimate up to 50%). In this case, I illustrate a number of pitfalls in getting accurate introspective data, underscoring how much needs to be settled to have any confidence that introspective data collection of peripheral experience is reliable and accurate. Effectively, one needs to run a careful, controlled introspective sampling, not just armchair reflection with an n (number of subjects) equal to 1. The upshot is that we need to think more systematically about the reliability and accuracy of introspection given all sorts of distorting influences on introspective attention. A structure for mental action can help frame possible lines of investigation. It enjoins us to identify the factors that contribute to shaping introspective judgments.

So, there the book is, at least in part. Again, as action is central in different ways to many areas of philosophical and empirical investigation, I hope readers will engage with the book. While the book aims for some strong conclusions, I also hope readers will find uncontroversial entry points to engaging with its ideas.

References

- Miller, George A, Eugene Galanter, and Karl H Pribram. 1960. *Plans and the Structure of Behavior*. New York: Holt, Rinehart and Winston.
- Wu, Wayne. 2020. "Is Vision for Action Unconscious?" *The Journal of Philosophy* 117 (8): 413–33.
- . 2023a. *Movements of the Mind*. Oxford: Oxford University Press.
- . 2023b. "Mineness and Introspective Data." In *Self-Experience: Essays on Inner Awareness*, edited by Manuel García-Carpintero and Marie Guillot, 120–41. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198805397.003.0006>.
- . 2024. "We Know What Attention Is!" *Trends in Cognitive Sciences* 28 (4): 304–18. <https://doi.org/10.1016/j.tics.2023.11.007>.