Lecture 02 : Philosophical Issues in Behavioural Science

Stephen A. Butterfill
< s.butterfill@warwick.ac.uk >

Monday, 7th October 2024

Contents

1. Introduction

This week we will deepen our understanding of the dual-process theory of instrumental action, examine some evidence supporting it (which will involve becoming familiar with some experimental paradigms), and consider how this theory might complicate attempts to solve The Problem of Action.

By the end of this lecture you should understand how to disentangle the contributions of habitual and goal-directed processes in producing an action. You should also be familiar with some evidence for the existence of habitual processes in humans. And you should be able to say how their existence leads to an objection to standard philosophical attempts to solve The Problem of Action.

This lecture is linked to one of the questions set for your first assignment, the short essay. The question is how, if at all, discoveries in the behavioural sciences should inform attempts to solve The Problem of Action.

This lecture depends on you having studied some sections from a previous lecture:

- *Goal-Directed and Habitual Processes* in Lecture 01
- *Philosophical Theories of Action* in Lecture 01

For the minimum course of study, consider only these sections:

- *The Minor Puzzle about Habitual Processes* (section §3)
- *The Problem of Action meets Habitual Processes* (section §5)

2. There Are No Habitual Actions

Only processes are habitual.

The term 'habitual' is used for at least two different things.

One use is for 'things someone usually does'. This use is common enough in everyday situations to be in a dictionary:

> 'A habitual action, state, or way of behaving is one that someone usually does or has, especially one that is considered to be typical or characteristic of them.' (collinsdictionary.com)

In this course, we never use 'habitual' in this first way.

The other use of 'habitual' applies to processes, not actions. It comes from comes from dual-process theory of instrumental action. This is the notion of habitual process in the glossary.

In this course, we always use 'habitual' in this second way.

It is hard to related the notion of a habitual process to the idea that an action is 'habitual' because all actions likely involve the influence of multiple types of process (see *The Minor Puzzle about Habitual Processes* (section §3)). The best sense we can make of the phrase 'habitual action' would be to say that it refers to particular actions which were dominated by habitual processes.

Philosophers and scientists use the term 'habitual' in several different ways (Du et al. 2022, p. 374). But as several researchers have argued at length (Du et al. 2022; Gardner 2015), if our focus is scientific discoveries about action we should think of processes responsible for actions as habitual or not rather than the actions themselves.[1] One reason for this is simply that even a paradigm small-scale case of so-called 'habitual action' likely involve multiple processes some but not all of which are habitual.

3. The Minor Puzzle about Habitual Processes

A rat has been given food contingent on its pressing a level. When it presses the lever, is its action habitual or instrumental? By the end of this section you should understand why this question is puzzling and also how to resolve the puzzle. You should also understand devaluation, and be able to understand an experiment that provides some of the foundational evidence for the dual-process theory of instrumental action.

You see a rat and a lever. The rat presses the lever occasionally. Now you start rewarding the rat: when it presses the lever it is rewarded with a particular kind of food. As a consequence, the rat presses the lever more often. This indicates that the rat's lever pressing is an instrumental action, for manipulating the outcome of the action has changed its frequency. But is this lever pressing primarily a consequence of habitual processes?

In thinking about this question, consider how we characterised habitual and goal-directed processes (in *Goal-Directed and Habitual Processes* in Lecture 01). What does the hypothesis that the rat's lever pressing is dominated by habitual processs predict? And what does the alternative hypothesis that the rat's lever pressing is dominated by goal-directed processs predict?

Because the aim of this section is to get you thinking about the questions, these notes do not answer them. The recording will take you through some considerations.

[1] 'We conclude that it is more fruitful to think about habits as a property of the intermediate computations that precede response generation (Figure 5B,C) rather than as a property of the response itself (Figure 5A).' (Du et al. 2022, p. 380)

3.1. The Minor Puzzle

Dickinson (1985) found that when a rat has learned to perform an instrumental action to obtain a food and when the food is devalued, the frequency with which the rat performs the action is reduced but the rat does nevertheless continue to perform the action

1. If the action is habitual, why is it influenced by devaulation at all?
2. If the action not habitual but controlled by goal-directed processes, why does it still occur (albeit less frequently) after devaluation?

3.2. The Dual-Process Theory of Instrumental Action Revisited

As the term 'habitual' is used on this course, *actions* are the wrong kind of thing to be described as habitual. It is the *processes* that trigger and guide actions that can be habitual, not the actions themselves.

This matters because on the dual-process theory of instrumental action, one action may be simultaneously guided by two or more distinct kinds of process, one goal-directed and the other habitual.

The Minor Puzzle is telling us that, in the case of the rat's action, both kinds of processes are influential. The frequency with which the rat performs the action is reduced, indicating that it is influenced by goal-directed processs but the rat does nevertheless continue to perform the action, indicating that it is influenced by habitual processs.

3.3. Conclusion

Actions are controlled by two or more distinct kinds of process, one goal-directed and the other habitual. If an action were very strongly dominated by habitual processes, we might informally label the action 'habitual'. But, as we will further explore in *Goal-Directed and Habitual: Some Evidence* (section §4), the actions of humans, like rats, are often significantly influenced by both kinds of process.

3.4. Appendix: Reflexes

Reflexes are an example of instrumental actions whose occurrence is fully explained neither by the habitual nor by the goal-directed process:

> 'A light puff of air directed at the cornea makes the eye blink. A tap just below the knee causes the leg to kick. A loud noise causes a startle reaction. These are all examples of reflexes. A reflex involves two closely related events: an eliciting stimulus and a corresponding response. Furthermore, the stimulus and response are linked. Presentation of the stimulus is followed by the response, and the response rarely occurs in the absence of the stimulus.' (Domjan 2010, p. 30)

4. Goal-Directed and Habitual: Some Evidence

According to the dual-process theory, instrumental actions can be a consequence of both goal-directed processs and habitual processs. So far we have mainly relied on testimony for this key premise. It's now time to consider evidence for it.

This is an optional section. We did not cover it in lectures. If you completed the reading for Seminar 1 *in Seminar Tasks, you have already encountered the evidence here (although perhaps not all of the details).*

Until *The Minor Puzzle about Habitual Processes* (section §3) we had not encountered any evidence at all for the dual-process theory of instrumental action. What evidence supports this theory?

The section introduces three sources of evidence:

1. cognitive load (via stress) - Schwabe & Wolf (2010)[2]
2. representation of contingency - Klossek et al. (2011)
3. neurophysiology - Dickinson (2016)

If you have difficulty with this (perhaps you are new to psychology, or perhaps you just struggle to follow the lecturer), please consider just the first of these.

It would be much better to have a firm understanding of Schwabe & Wolf (2010) than to have a sense of what each of the three sources of evidence involves.

[2] Note that Buabang et al. (2023) report a failed replication of this finding. If you rely on Schwabe & Wolf (2010), it would be good to consider whether this failed replication should undermine confidence in the original result. My own view is that it should not. This is because whereas Schwabe & Wolf (2010) used sateity to devalue, Buabang et al. (2023) used 'Tween 20 (Polysorbate 20), a colorless and odorless substance that creates a bad taste'. As the authors note, this creates an aversion to the food. But there is a distinction between a change in desire for a food and an aversion to it. We would expect habitual behaviours to be influenced by change in aversion but not by a change in desire (see *Preference vs Aversion: A Dissociation* in Lecture 08).

4.1. Speed vs flexibility

In the lecture I justify some of the predictions tested with the consideration that any broadly cognitive process must make a trade-off between speed and flexibility. This idea is further developed by Daw et al. (2005, p. 1705) who contrast the use of cached values (which is fast but insensitive to rapid changes in the environment) with values computed on the fly (which may demand time and effort but allows more flexibility).

In essence, the idea is that the goal-directed process involves searching through potential actions, predicting their likely consequences and anticipating how valuable (or not) those consequences would be. This 'poses severe demands on computation and memory and rapidly becomes intractable with growing complexity.' (Wunderlich et al. 2012, p. 786). By contrast, the habitual process is much less demanding as it does not even require memory of the consequences of actions. But there is a trade-off: in return for being less demanding, the habitual process is unreliable in a rapidly changing environment or where there is insufficient learning.

5. The Problem of Action meets Habitual Processes

Does the fact that habitual processs and not only goal-directed processs influence instrumental actions pose a challenge to the Standard Solution to The Problem of Action? Might this fact even assist us, eventually, in developing a challenge to the Causal Theory of Action?

5.1. Why Focus on The Problem of Action?

From philosophy we want a framework that supports theorising about action in the behavioural and social sciences. Minimally, the framework should

- allow us to make all the important distinctions;
- enable us to formulate questions about how and why agents act; and
- support deriving predictions from hypotheses about the answers to these questions.

That, at least, is the framework we (well, mainly you[3]) are attempting to construct in thinking through philosophical issues in behavioural sciences.

It seems reasonable to expect that any such framework must solve The Problem of Action. After all, the distinction between an action and event that

[3] Your lecturer enjoys the luxury, in teaching, of being able to point to multiple conflicting sources, leaving to you the hard work of arriving at the truth and discerning the limits of what we know. Their role is to introduce and motivate questions, yours to answer them.

merely happens to you looks fundamental. So while solving this problem is not sufficient for our aims, doing so does seem to be necessary.

5.2. Objection to the Standard Solution

The Problem of Action is to say what distinguishes your actions from things that merely happen to you (see ** **ERRoR! MISSING xref FOR** $unit : philosophical_theories_habits$ **).

According to the Standard Solution to this Problem, actions are those events which stand in an appropriate causal relation to an intention (see ** **ERRoR! MISSING xref FOR** $unit : philosophical_theories_habits$ **).

What counts as 'appropriate' here? This turns out to be a hard problem to answer. Davidson (1980, p.79) noticed, in effect, that intentions can cause events which would not thereby count as intentional actions. We therefore cannot say simply that actions are events caused by intentions; they have to be caused 'in the appropriate way', whatever that is.

For our purposes (considering an objection to the Standard Solution), we need not fully specify what counts as 'appropriate'.[4] It is enough to notice that, for the causal relation to be appropriate, minimally:

- the action should not manifestly run counter to the agent's intentions; and
- neither should whether the action occurs be independent of what the agent intends.

Objection to the Standard Solution: some actions are dominated by habitual processes and may therefore manifestly run counter to your intentions. For example, it is possible to continue seeking out a sweet chocolate drink instead of peppermint tea despite being sated on the drink and therefore currently preferring the peppermint tea (compare Schwabe & Wolf 2010 discussed in *Goal-Directed and Habitual: Some Evidence* (section §4)). Since it is irrational to intend to knowingly seek out a less preferred alternative at no greater cost than seeking a more preferred alterantive, it is possible for this action to occur counter to your intentions. Therefore not all actions do stand in an appropriate causal relation to an intention.

Other cases illustrating how habitual processes are insensitive to intentions and can therefore run counter to them about. (Wood & Rünger 2016, p. 293) cite two:

[4] Shepherd (2021, chapter 3) offers a recent attempt.

1. 'when students who frequently went to the sports stadium on campus were incidentally exposed to an image of the stadium, they raised their voices as they would habitually in that context, despite no change in their motivation to speak loudly (Neal et al. 2012)'
2. 'in a study conducted in a local cinema, participants with stronger habits to eat popcorn at the movies consumed more than those with weak habits, even when they disliked the popcorn because it was stale and unpalatable (Neal et al. 2011).'

5.3. First Response to the Objection

In response to the above Objection, consider the possibility of insisting that in every case the agent really does have a guiding intention after all. Could there be a good reply based on this response?

Note that insisting on something contrary to what has been argued is not properly a *reply* to the Objection but merely a *response*.[5] Your challenge is to evaluate whether or not this line of response might be turned into a way of overcoming or avoiding the objection.

5.4. Refined Version of the First Response to the Objection

A more refined version of this response might focus on the various causal roles intentions can have. As well as guiding actions directly, intentions may guide actions indirectly via habitual processes. To illustrate, you may act on an intention to get fitter by doing some exercises right after your alarm clock rings each day, and then giving yourself a little reward. Over time, if you are lucky, the habitual process may take over so that you no longer have to remember to exercise and just find yourself doing it. At this point, your actions are independent of your intentions in one sense. But the intention to exercise is still indirectly related to your actions through habitual processes. Perhaps, then, we should refine Davidson's idea about what distinguishes actions from things that merely happen to you so that it can accommodate the ways that habitual processes can mediate between intentions and actions.

The refined response allows that actions can happen contrary to your current intentions. Perhaps you have a new partner who gets up later and would be disturbed by your exercise. Valuing the relationship over your work out, your sole concern is not to disturb them. Yet force of habit is too strong and, despite your clear intention to the contrary, you are dismayed to find

[5] See further Argument Clinic (2021).

yourself exercising vigorously one morning. This is a complication that the refined response should address. However, it also faces a deeper objection …

Many habitual processes are established independently of intention. What is required is just that a stimulus is followed by an action which, in turn, is followed by a reward. These originating actions may not involve intention at all (as in utilisation behaviour), or else they may involve an intention but one that is not supposed to be habit forming. You may intentionally eat some chocolate after a meal relying on the idea that this will be a rare treat and yet, contrary to your intentions and policies, be landed with a habit. So while it is an important insight that habitual processes *can* mediate between intentions and actions, it also true that that habitual processes are autonomous in the sense that intentions are not necessary.

Given this objection, one might try to defend the idea that the actions habitual processes cause are actions when, and only when, they habitual processes involve intentions in an appropriate way.[6] Is this position defensible?

5.5. Second Response to the Objection

In response to the above Objection, consider restricting both The Problem of Action and the Standard Solution to intentional action.

Could there be a good reply based on this response? It may avoid the Objection, given the further assumption that actions dominated by habitual processes are not intentional actions. But this appears to be a hollow victory. After all, what was supposed to be a bold revelation about action would, if we accepted the reply, turn out to be merely the claim that intentional actions are things that stand in an appropriate causal relation to an intention.

Further, unless we think that all actions are intentional actions (which would be hard to square with the above Objection), the original Problem of Action is still a good question. We still need to know what distinguishes actions of all kinds from things that merely happen to you.

Are there better replies to the Objection? Or can this response be developed in a much better way? If neither, should we revise or reject the Standard Solution?

We might minimally revise the Standard Solution by saying that actions are those events which stand in an appropriate relation to either a goal-directed process or a habitual process. This quite minor revision allows us to retain the Causal Theory of Action.

[6] This possibility is suggested by Kalis & Ometto (2021, p. 645).

But can stimlus-action links and habitual processes really be relevant to solving The Problem of Action?

5.6. Further Replies to the Objection

5.6.1. An Anscombian Perspective

Kalis & Ometto (2021, p. 640ff) provide a critical overview of several philosophers' attempts to reply to a variant of the above Objection. These authors propose their own response, which does involve rejecting the Standard Solution in favour of an Anscombe-inspired alternative.[7]

If exploring further work by philosophers, be careful to check whether their understanding of habitual process matches yours. You can tell that this will be tricky from the fact that Kalis & Ometto (2021, p. 640ff) write about 'habitual actions', whereas, strictly speaking, no such things exist on the dual-process theory of instrumental action (as explained in *The Minor Puzzle about Habitual Processes* (section §3)).

In fact, Kalis and Ometto do appear to have a view on which actions dominated by habitual processs are not actually (intentional) actions:

> 'The concept of an intentional action, then, is essentially the concept of a kind of behavior that makes sense to the agent as her action. An intentional action is action that is, therefore, partially constituted by the agent's point of view, or her own take on what she is doing. (Kalis & Ometto 2021, p. 644)

5.6.2. Basic Actions?

A different line of response might be to appeal to so-called basic or primitive actions, that is, actions which you can perform without performing any other action (Davidson 1971).[8] In cases like popcorn eating where, supposedly, actions can run counter to any intention, consider that there is a distinction between the larger action (eating popcorn) and component actions like reaching for some popcorn, grasping it, transporting it to the mouth and eating it. Regardless of whether the larger action runs counter to any intention, might these component actions nevertheless be appropriately related to the agents intentions? If so, could we revise the Standard Solution to avoid the Objection above?

---

[7] Note that these authors' are presenting a slightly different objection from the one above, as you can see from their diagnosis of how the objection arises (Kalis & Ometto 2021, p. 642).

[8] As Schlosser (2019, footnote 17) notes, there is no agreement about how to characterise basic actions. This notion should be invoked with caution and avoided where possible.

Discoveries about motor representation (see *Motor Representation* in Lecture 07) complicate this line of response in two ways—they make it harder to characterise actions like reaching and grasping as basic actions, and they indicate that may be no need to postulate intentions concerning these actions specifically (as they are already well taken care of by motor representations).

5.7. Bonus Dangling Question: Alternative to the Causal Theory?

According to the Causal Theory of Action, an event is action 'just in case it has a certain sort of psychological cause' (Bach 1978, p. 361). If we retain the Causal Theory and if we also accept that some actions are dominated by habitual processes and may therefore run counter to your intentions, then we will have to invoke not only beliefs, desires and intentions but also stimlus-action links in distinguishing actions from events that merely happen to you.

This may motivate considering alternatives to the Causal Theory.

Consider two questions:

1. What distinguishes instrumental actions from things which merely happen to an agent (and from noninstrumental actions, if there are any)? [This is 'The Problem of Action']

2. Which states cause instrumental actions?

Fully understanding action requires answering both questions (and more).[9] But the Causal Theory of Action insists on answering the first question in a way that also involves answering, partially or wholly, the second. The idea is not simply that better understanding answers to the second question might guide us in working out the answer to the first question. On the Causal Theory of Action, any answer to the first question must already involve answering the second. There is no possibility, not even in principle, of answering the first question correctly but then discovering that everything we thought we knew about the second question is wrong.

Let us say that any answer to the first question which does not involve making commitments concerning which states, or structures of states, cause instrumental actions is *mechanistically neutral* (as opposed to a *mechanistically committed* answer, which the Causal Theory of Action requires).[10]

---

[9] Of course there are philosophers who might deny that the second question bears on any philosophical questions about action (Ginet (1990), for example).

[10] Note that the possibility of characterising A in terms which do not mention B does not in general imply that it is possible for there to be As without corresponding Bs. Proponents

If we reject the Causal Theory of Action, we will need a mechanistically neutral solution to The Problem of Action. What might that be?

6. Conclusion

The previous lecture established that there are two theoretically coherent models of instrumental action, one involving goal-directed processes and the other habitual processes.

In this lecture, we considered some evidence indicating that each of the models does actually explain some actions.

This creates an obstacle for understanding what actions are. In philosophy there is a widely-held, standard view about this. But that view seems to be inconsistent with the dual-process theory of instrumental action. If so, we need an alternative philosophical framework to support theorising about action in the behavioural and social sciences.

Glossary

Causal Theory of Action According to this view, an event is action 'just in case it has a certain sort of psychological cause' (Bach 1978, p. 361). 6, 9, 11, 12

devaluation To *devalue* some food (or video clip, or any other thing) is to reduce its value, for example by allowing the agent to satiete themselves on it or by causing them to associate it with an uncomfortable event such as an electric shock or mild illness. 3, 4

dual-process theory of instrumental action Instrumental action 'is controlled by two dissociable processes: a goal-directed and an habitual process' (Dickinson 2016, p. 177). (See instrumental action.) 2–5, 10, 12

goal-directed process A process which involves 'a representation of the causal relationship between the action and outcome and a representation of the current incentive value, or utility, of the outcome' and which influences an action 'in a way that rationalizes the action as instrumental for attaining the goal' (Dickinson 2016, p. 177). 3–6, 9

of a mechanistically neutral approach may therefore accept that instrumental actions are caused by intentions and could not be caused in some other way .

habitual process A process underpinning some instrumental actions which obeys *Thorndyke's Law of Effect*: 'The presentation of an effective [=rewarding] outcome following an action [...] reinforces a connection between the stimuli present when the action is performed and the action itself so that subsequent presentations of these stimuli elicit the [...] action as a response' (Dickinson 1994, p.48). (Interesting complication which you can safely ignore: there is probably much more to say about under what conditions the stimulus–action connection is strengthened; e.g. Thrailkill et al. 2018.) 2–6, 9, 10

instrumental action An action is *instrumental* if it happens in order to bring about an outcome, as when you press a lever in order to obtain food. (In this case, obtaining food is the outcome, lever pressing is the action, and the action is instrumental because it occurs in order to bring it about that you obtain food.) You may encounter variations on this definition of *instrumental* in the literature. For instance, Dickinson (2016, p. 177) characterises instrumental actions differently: in place of the teleological 'in order to bring about an outcome', he stipulates that an instrumental action is one that is 'controlled by the contingency between' the action and an outcome. And de Wit & Dickinson (2009, p. 464) stipulate that 'instrumental actions are *learned*'. 3, 5, 12

mechanistically neutral A characterisation of instrumental action (or of joint action) is *mechanistically neutral* just if it does not involve making commitments concerning which states, or structures of states, cause instrumental actions (or cause joint actions). 12

motor representation The kind of representation characteristically involved in preparing, performing and monitoring sequences of small-scale actions such as grasping, transporting and placing an object. They represent actual, possible, imagined or observed actions and their effects. 11

Standard Solution (to The Problem of Action). Actions are those events which stand in an appropriate causal relation to an intention. 9, 10

The Problem of Action What distinguishes your actions from things that merely happen to you? (According to Frankfurt (1978, p. 157), 'The problem of action is to explicate the contrast between what an agent does and what merely happens to him.') 1, 2, 6, 9, 10, 12

References

Argument Clinic (2021). Argument Clinic — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Argument_Clinicoldid=1061477940. Page Version ID: 1061477940.

Bach, K. (1978). A representational theory of action. *Philosophical Studies*, *34*(4), 361–379.

Buabang, E. K., Boddez, Y., Wolf, O. T., & Moors, A. (2023). The role of goal-directed and habitual processes in food consumption under stress after outcome devaluation with taste aversion. *Behavioral Neuroscience*, *137*(1), 1–14.

Davidson, D. (1971). Agency. In R. Binkley, R. Bronaugh, & A. Marras (Eds.), *Agent, Action, and Reason,* (pp. 3–25). Toronto: University of Toronto Press. Reprinted in Davidson, D. (1980) *Essays on Actions and Events*. Oxford: Oxford University Press.

Davidson, D. (1980). *Essays on actions and events*. Oxford: Oxford University Press.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.

de Wit, S. & Dickinson, A. (2009). Associative theories of goal-directed behaviour: A case for animal–human translational models. *Psychological Research PRPF*, *73*(4), 463–476.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. In L. Weiskrantz (Ed.), *Animal Intelligence*. Oxford: Clarendon Press.

Dickinson, A. (1994). Instrumental conditioning. In N. Mackintosh (Ed.), *Animal Learning and Cognition*. London: Academic Press.

Dickinson, A. (2016). Instrumental conditioning revisited: Updating dual-process theory. In J. B. Trobalon & V. D. Chamizo (Eds.), *Associative learning and cognition*, volume 51 (pp. 177–195). Edicions Universitat Barcelona.

Domjan, M. (2010). *The Principles of Learning and Behavior*. Wadsworth Cenage Learning.

Du, Y., Krakauer, J. W., & Haith, A. M. (2022). The relationship between habits and motor skills in humans. *Trends in Cognitive Sciences*, *26*(5), 371–387.

Frankfurt, H. (1978). The problem of action. *American Philosophical Quarterly*, *15*(2), 157–162.

Gardner, B. (2015). A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. *Health Psychology Review*, *9*(3), 277–295.

Ginet, C. (1990). *On Action*. Cambridge University Press.

Kalis, A. & Ometto, D. (2021). An Anscombean Perspective on Habitual Action. *Topoi*, *40*(3), 637–648.

Klossek, U. M. H., Yu, S., & Dickinson, A. (2011). Choice and goal-directed behavior in preschool children. *Learning & Behavior*, *39*(4), 350–357.

Neal, D. T., Wood, W., Labrecque, J. S., & Lally, P. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *Journal of Experimental Social Psychology*, *48*(2), 492–498.

Neal, D. T., Wood, W., Wu, M., & Kurlander, D. (2011). The Pull of the Past: When Do Habits Persist Despite Conflict With Motives? *Personality and Social Psychology Bulletin*, *37*(11), 1428–1437.

Schlosser, M. (2019). Agency. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019 ed.). Metaphysics Research Lab, Stanford University.

Schwabe, L. & Wolf, O. T. (2010). Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. *Psychoneuroendocrinology*, *35*(7), 977–986.

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

Thrailkill, E. A., Trask, S., Vidal, P., Alcalá, J. A., & Bouton, M. E. (2018). Stimulus control of actions and habits: A role for reinforcer predictability and attention in the development of habitual behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*, 370–384.

Wood, W. & Rünger, D. (2016). Psychology of habit. *Annual review of psychology*, *67*, 289–314.

Wunderlich, K., Dayan, P., & Dolan, R. J. (2012). Mapping value based planning and extensively trained choice in the human brain. *Nature Neuroscience*, *15*(5), 786–791.