# The psychology of implicit knowledge

Zoe Drayson, UC Davis

## Implicit knowledge: the very idea

Explicit knowledge is consciously accessible to the knower: the person can introspect what it is that they know and articulate it in the form of a statement (Dummett 1991, Davies 2015, Thompson 2023). If a person possesses some knowledge which they are unable to articulate to themselves or others, this knowledge is said to be implicit rather than explicit. Standard examples of implicit knowledge include a speaker’s knowledge of language, or practical knowledge such as how to ride a bike. The concept of implicit knowledge, however, raises challenging philosophical questions. If attributions of mental states only make sense against a background assumption of rational relations between thought and action (see e.g. Davidson 1980), then it seems difficult to attribute knowledge to a person who is unable to assert what they know. And it is hard to see how there can be non-introspectable knowledge if one thinks that there is a constitutive connection between being in a mental state and having introspective knowledge about that state (see e.g. Shoemaker 1994).

Philosophers have responded to these concerns by showing how it is conceptually possible to possess knowledge which one cannot introspect or articulate. In cases of practical knowledge, for example, Stanley and Williamson (2001) have argued that the knower does in fact grasp a proposition, but the mode of presentation under which they grasp this proposition (practical rather than theoretical) accounts for the knower’s inability to introspect or articulate the proposition. And Elga and Rayo (2021) have argued that attributions of knowledge can be relativized to tasks, contexts, or purposes without compromising the rationality of the knower: certain knowledge might be available to guide the agent’s behavior in some respects without necessarily being available for verbal report. Notice that both arguments concern the conceptual possibility of implicit knowledge and the appropriateness of our attributions of implicit knowledge. Neither of them makes empirical claims about the mechanisms of cognition which causally explain cases of implicit knowledge: Stanley and Williamson (2001) intend their claims about knowledge-how to be entirely neutral on how propositional knowledge is realized in the brain, while Elga and Rayo (2021) emphasize that they are not specifying the cognitive architecture which underlies the accessibility of information to a subject relative to certain elicitation conditions. They presumably take claims about the realization of implicit knowledge in our brain’s cognitive architecture to be a task for cognitive psychology rather than philosophy.

This piece explores some of the different ways in which cognitive psychology can attempt to explain implicit knowledge and considers how these scientific explanations relate to the more traditionally philosophical questions, described above, about the concept of implicit knowledge.

## The explanatory framework of cognitive psychology

Cognitive psychology understands the human brain as an information processor: a physical computer which uses neural hardware to run cognitive software. Cognitive psychologists attempt to give a causal explanation of human cognitive capacities by appealing to the brain’s storage and transmission of physically-implemented and semantically-interpretable data. These information-processing mechanisms provide cognitive psychology with several ways to explain how a subject can have knowledge which they are unable to articulate. Before we explore these, it is worth considering how Freudian psychology and behaviorist psychology accounted for implicit knowledge prior to the emergence of cognitive psychology in the latter half of the twentieth century.

Freudian psychology has much to say about inarticulable information. Freud attempted to provide an etiological account of a certain kind of implicit knowledge, notably that which is traumatizing to the subject or socially unacceptable. He proposed that such content could become inaccessible by the subject due to a causal process of repression or dissociative amnesia. This repressed knowledge is implicit in the sense of not being available to the subject for introspection or articulation, but Freud proposed that it could be made manifest in certain behaviors (e.g. slips of the tongue, phobias, neuroses) and in dreams, and potentially recovered through psychotherapy. While Freud’s story is intended to be a causal one, he leaves the physical mechanisms of repression unspecified. At the time, there was no account of how mental states could be processed differently by the brain in virtue of their meanings or significance to the subject.

Partly due to the lack of such a causal account, behaviorist psychology resisted committing to internal mental states with semantic properties like meaning or significance. Behaviorist psychology focused instead on behavioral data: worldly stimuli and our physical response to them. Knowledge, for behaviorist psychologists, was to be understood as a repertoire of behaviors rather than in terms of internal mental states which cause the behaviors. The explicit knowledge that Paris is the capital of France, for example, would be nothing more than the behavior of asserting the claim “Paris is the capital of France”. Implicit knowledge, for behaviorists, would similarly be considered as a repertoire of behaviors, but one which does not include the assertion of a statement of the knowledge in question.

Cognitive psychology proposes that intelligent behavior can be caused by internal mental states and processes, where these mental states are understood as having both causal and semantic properties. Cognitive psychologists explain explicit knowledge in terms of the brain tokening an internal vehicle of representation (a symbol, activation pattern, or probability distribution, for example) which encodes the relevant information in some physical pattern of neural activity. The internal vehicle of representation is computationally processed in ways which makes the information accessible to the subject more generally: for the guidance of behavior, for introspection and inference, and for verbal report. (See Drayson 2018 for further discussion of why and how cognitive psychology individuates representations as physically implemented vehicles.)

So much for explicit knowledge: but how does cognitive psychology explain implicit knowledge? Several different approaches can be taken to account for our capacity to possess information which we cannot consciously access.

## The cognitive psychology of implicit knowledge

One way for cognitive psychology to explain the difference between explicit knowledge and implicit knowledge is to appeal to a distinction in the cognitive architecture between two different kinds of causal processes. This is the approach taken by ‘dual-process’ accounts of cognition (e.g. Frankish 2010). According to such approaches, the brain implements two very different kinds or systems of computational architectures. The first of these (usually known as ‘System 1’) employs parallel processing and associative operations to produce behavior quickly and automatically. The second of these (usually known as ‘System 2’) employs sequential processing and rule-based operations to produce behavior more slowly and effortfully. System 1, assumed to be evolutionarily and developmentally earlier, is thought to be responsible for intuitive and unreflective sensory-motor interactions with the world, such as facial recognition. System 2, assumed to be evolutionarily and developmentally later, is thought to be responsible for controlled and reflective cognitive capacities such as abstract thought. It is easy to see how System 2 can account for explicit knowledge, and System 1 appears to be able to account for some instances of implicit knowledge. But there are certain standard examples of implicit knowledge, such as our knowledge of language, which fail to be easily captured by dual-process accounts.

Since Chomsky (1965), our knowledge of language has been proposed as an archetypal rule-based computational process: our ability to acquire, understand and speak a language is best explained by attributing that person with knowledge of the rules and principles of the grammar which govern that language.[[1]](#endnote-1) But since we are generally unable to articulate those rules and principles, such linguistic knowledge must be understood as implicit. This would be an example of a case where rule-governed processing produces implicit knowledge in a way which cannot be straightforwardly explained by the standard dual-processing approach. Fodor (1968) proposes that we can give a similar account of practical knowledge, like knowing how to tie one’s shoe-laces, in terms of internal representations in the form of “propositions, maxims, or instructions”, even where we cannot articulate the contents of those representations.

“Now I want to say: if X is something an organism knows how to do but is unable to explain how to do, and if S is some sequence of operations, the specification of which would constitute an answer to the question "How do you X?," and if an optimal simulation of the behavior of the organism X-s by running through the sequence of operations specified by S, then the organism tacitly knows the answer to the question "How do you X?," and S is a formulation of the organism's tacit knowledge.” (Fodor 1968, 638)[[2]](#endnote-2)

Stich (1978) refers to the sorts of information posited by Chomsky and Fodor as subdoxastic: it involves belief-like informational states and processes, but it is isolated in a distinct cognitive subsystem. It is not inferentially integrated with consciously accessible cognitive processes. One way to account for this lack of inferential integration is through the framework of modularity (Fodor 1983), and information-encapsulation in particular. We can understand different cognitive subsystems as distinct computer programs, which don’t interact informationally with each other. (Notice that Fodor does not think that all implicit knowledge is explained by modularity. He proposes that our implicit knowledge about the validity of inference rules like *modus ponens*, for example, is better understood as a non-representational property of our unencapsulated cognitive processing rather than as a representational property of an encapsulated cognitive subsystem – see below.)

All the approaches outlined above appeal to the nature of the computational processes operating over an internal vehicle of representation (e.g. associative vs rule-based, encapsulated vs unencapsulated) to account for the subject’s inability to access the information encoded by the representational vehicle. Alternatively, a cognitive psychologist might focus less on the computational processes and more on the way that the vehicles store information in the first place. The formal, functional, or physical properties of representational vehicles allow them to encode content in different format: consider how a sentence and a map can carry information in distinct ways. By comparison, neural states might carry information discursively in a sentence of a computer language, or iconically by being structurally isomorphic to their contents. Such a difference in format could explain why the information is available or unavailable for linguistic articulation: perhaps iconic representations account for implicit knowledge while discursive representations account for explicit knowledge.

The vehicle approach and the processing approach described above are not mutually exclusive: implicit knowledge might the result of both the way that information is stored and the way that is available to be accessed. (In connectionist cognitive architectures, for example, it becomes difficult to draw a distinction between the format of the vehicles and the way they are processed.)

Furthermore, these approaches are not the only options. Cognitive psychologists might take an operational notion of implicitness, such that the implicitness of knowledge is a function of ways we use to measure it (see Bendaña 2023). Alternatively, they might accept that explicit knowledge involves vehicles of representation while denying that implicit knowledge does. This is the suggestion that implicit knowledge is inarticulable precisely because it is not represented by the cognitive system, but rather built into the way the system processes information. This is the sort of implicit knowledge that Dennett (1977) ascribes to the chess-playing computer program which knows to get its queen out early, and the sort of implicit knowledge that Fodor (1983) thinks we might have of inference rules (see above). Proponents of embodied approaches to cognition often take this stance, emphasizing the respects in which our practical knowledge can be explained without invoking internal vehicles of representation.

## Implications for the nature of implicit knowledge

I have been focusing thus far on the ways that cognitive psychology can attempt to give a causal explanation of the production of implicit knowledge: which mechanistic features of neural information processing can result in information which is not articulable by the subject? These are ways to address the question of *why* some information is inaccessible to the subject, while other information is consciously accessible for introspection and verbal report. Some philosophers, however, propose that the correct scientific explanation can provide insights into the metaphysical nature of implicit knowledge. Pavese (2017), for example, thinks that Stanley and Williamson’s (2001) notion of a practical mode of presentation can be given a more rigorous characterization by appealing to theories from cognitive psychology about the motor representation system. Bendaña and Mandelbaum (2021) propose that Elga and Rayo’s (2021) ‘fragmented’ approach to knowledge attributions is constrained by theories from cognitive psychology about how information is stored and accessed.[[3]](#endnote-3)

What if it turns out that the epistemological distinction with which we started, between explicit knowledge and implicit knowledge, does not map neatly on to a scientific distinction between two types of information processing or two types of representational format? It could turn out, for example, that our knowledge of language is the result of an informationally-encapsulated rule-governed process over discursive representations, while our knowledge of how to ride a bicycle is embodied rather than represented. If we take a more *a priori* or ‘top down’ approach to cognitive ontology, we could see these architectural differences as philosophically trivial: at most they tell us something about the different realizers of implicit knowledge, and not about the nature of implicit knowledge itself. But the more *a posteriori* or ‘bottom up’ our approach to cognitive ontology, the more inclined we would be to revise our categorization of implicit and explicit knowledge to be more in line with our scientific discoveries: if *a posteriori* psychology categorizes implicit knowledge differently from our *a priori* philosophical theories, then we might interpret this (as Norby 2014 does) as a problem for our initial philosophical categorization.)This is ultimately a matter of how naturalistic we are in our approach to distinguishing cognitive kinds.

## References

Bendaña, Joseph & Mandelbaum, Eric (2021). The Fragmentation of Belief. In Cristina Borgoni, Dirk Kindermann & Andrea Onofri (eds.), *The Fragmented Mind* (78-107). Oxford, UK

Chomsky, Noam. (1965) *Aspects of a Theory of Syntax.* Cambridge, M.A.: MIT Press.

Collins, John. (2023). Chomsky, cognizing and tacit knowledge. In J. Robert Thompson (ed.) *The Routledge Handbook of Philosophy and Implicit Cognition*. Routledge.

Davidson, Donald (1980). *Essays on Actions and Events*. Oxford: OUP.

Davies, M. (2015) Knowledge (explicit and implicit): philosophical aspects. *International Encyclopedia of the Social and Behavioral Sciences* (2nd edition), 74-90. Academic Press.

Dennett, Daniel C., 1977 “Critical Notice: Review of The Language of Thought by Jerry Fodor”, *Mind*, 86(342): 265–280.

Drayson, Zoe (2018). The realizers and vehicles of mental representation. *Studies in History and Philosophy of Science Part A* 68:80-87.

Drayson, Zoe (2023). The fragmented mind: personal and subpersonal approaches to implicit mental states. In J. Robert Thompson (ed.) *The Routledge Handbook of Philosophy and Implicit Cognition*. Routledge.

Dummett, Michael (1991). *The Logical Basis of Metaphysics*. Harvard University Press (Cambridge, MA).

Elga, Adam & Rayo, Agustin (2021) Fragmentation and information access. In Cristina Borgoni, Dirk Kindermann and Andrea Onofri (eds), *The Fragmented Mind* (37-53). Oxford University Press.

Fodor, Jerry A. (1968). The appeal to tacit knowledge in psychological explanation. Journal of Philosophy 65 (October):627-40.

Fodor, Jerry A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Frankish, Keith (2010). Dual-Process and Dual-System Theories of Reasoning. *Philosophy Compass* 5 (10):914-926.

Norby, Aaron (2014). Against Fragmentation. *Thought: A Journal of Philosophy* 3 (1):30-38.

Pavese, Carlotta (2017). A Theory of Practical Meaning. *Philosophical Topics* 45 (2):65-96.

Schwartz, Arieh & Drayson, Zoe (2019). Intellectualism and the argument from cognitive science. *Philosophical Psychology* 32 (5):662-692.

Schwitzgebel, Eric (2021). The pragmatic metaphysics of belief. In Cristina Borgoni, Dirk Kindermann and Andrea Onofri (eds), *The Fragmented Mind* (350-376). Oxford University Press.

Shoemaker, Sydney (1994). Self-Knowledge and ‘Inner-Sense’. *Philosophy and Phenomenological Research* 54: 249-314.

Stanley, Jason and Williamson, Timothy (2001) Knowing How. *Journal of Philosophy* 98, pp. 411-444.

Stich, Stephen P. (1978). Beliefs and subdoxastic states. *Philosophy of Science* 45 (December):499-518.

Thompson, J. Robert (2023). Introduction: in search of the implicit. In J. Robert Thompson (ed.) *The Routledge Handbook of Philosophy and Implicit Cognition*. Routledge.

**Notes**

1. Chomsky’s ideas about the epistemic status of the knowledge in question are somewhat opaque and have arguably changed over time. See Collins (2023) for discussion. [↑](#endnote-ref-1)
2. Fodor uses the term ‘tacit knowledge’ for what I am calling ‘implicit knowledge’ here. [↑](#endnote-ref-2)
3. For further discussion of the relationship between personal and subpersonal approaches to attributions of mental fragmentation, see Schwartz and Drayson (2019), Drayson (2023). [↑](#endnote-ref-3)