

## Précis of *Thinking Off Your Feet*

Michael Strevens

*Analysis*, 82, 303–306, 2022

*Thinking Off Your Feet* (henceforth TOYF) is a study in the psychology of philosophical inquiry. Its topic is philosophical analysis, sometimes called the method of cases or conceptual analysis, the process in which hypotheses about the nature of things like propositional knowledge or singular causation are subjected to trial by counterexample, undertaken in the hope that the last theory standing will tell us what knowledge or causation really are. At the core of the book is a theory about the empirical psychology of philosophical concepts, that is, a theory about the structure of the mental representations by means of which we think about the subject matter of philosophical inquiry—the mental representations, then, of knowledge or of causation. Armed with a plausible psychological story, I hope both to understand and, as my subtitle proclaims, to vindicate analysis.

By an understanding of philosophical analysis, I mean the provision of an explanation of the contours of analysis both on the small and the large scale. What psychological mechanisms are responsible for an individual's judgments about philosophical thought experiments—the judgment, say, that a certain Gettierized belief is not an instance of knowledge? Why do such judgments sometimes seem to be empirically incorrigible, in the sense that the judger feels confident (rightly or wrongly) that no further information could weaken or undermine their conclusion? Why does it seem, to some of us at least, that these judgments are giving us information about a philosophical subject matter, rather than about the internal workings of our minds—about

the nature of knowledge, and not just our concept of knowledge?

Then, taking a longer-term perspective, why do attempts at analysis so often seem to fall short just as they begin to close in on a successful theory, spawning a profusion of epicycles and then collapsing under their own weight? Why are so many erstwhile analysts pessimistic about the whole endeavor, sharing Timothy Williamson's (2000, 31) opinion that "the pursuit of analyses is a degenerating research programme"?

By a vindication of analysis, I mean—well, my commentators in this symposium have applied some not undeserved pressure to say exactly what I mean. At a minimum, I would like to discern psychological mechanisms and semantic structures underlying philosophical analysis that give us some hope that, when it is done right, it can provide us with substantive knowledge of its ostensible subject matter, the nature of knowledge, causation, and so on. Ideally, I would like to show that philosophical analysis in general is a respectable form of inquiry. At the same time, I'd like to make plenty of space for varieties of philosophical inquiry that depart from the armchair model—not least, TOYF itself.

My attempts at both understanding and vindication are built on what I call an "inductivist" conception of human mental representation and thought. I start out with the "language of thought" model characteristic of much cognitive psychology, in which beliefs are conceived of as something like sentences in the head, composed of mental "words" representing individual objects, properties, and so on. (As with most scientific models, it is possible to harness the explanatory potential of the language of thought model without interpreting it in an overly literal way.)

To this I add what psychologists call the "theory-theory" of concepts, according to which the concept of (say) a horse—by which psychologists mean the mental structure that guides thought about horses, and hence the deployment of the mental word "horse"—is a theory of horses, or to put it more plainly, a set of beliefs about horses, representing some of their appearances,

behaviors, their relations to other horses, and so on. The beliefs constituting such a theory are often rather shallow, and to the extent that they are deep, they represent explanatory rather than metaphysical facts—the capacities of horses, their susceptibilities, and other facts about their place in the causal economy, rather than hypotheses about what ultimately makes something a horse. Their epistemic status is equally mundane: they are vulnerable to empirical disconfirmation or indeed rethinking of any sort. They are simply, as I put it in *TOYF*, “ordinary beliefs”.

That said, they may contain a lot of information (or at any rate, opinion) about horses—enough to give the mental word “horse” a rich inferential role in the cognitive ecosystem. One aspect of that role is a set of categorical dispositions: the thinker uses their many beliefs about horses—their “theory” of horses—to determine, of particular animals, whether or not they are horses. That is how most of us recognize a horse when we see one.

What if philosophical concepts, such as the concept of propositional knowledge or the concept of singular causation, work in the same way? That is the thought that kicks off the positive project in *TOYF*. What I’m imagining, then, is that our thought about (say) knowledge, and in particular our judgments about what counts as knowledge, are guided by a set of ordinary beliefs about knowledge—not some stipulative criterion for what counts as knowledge, but just our best surmises about the role of knowledge in the appropriate explanatory network—psychological, epistemic, or whatever.

Then we might, on the one hand, have a rich basis for making sophisticated decisions about what does and does not count as knowledge in elaborate philosophical thought experiments—about whether this or that Gettierized belief qualifies as knowledge—while at the same time having no fixed view about the question of philosophical interest, namely, the question of what ultimately makes something knowledge. The only way to answer the question would be to use our category judgments, in just the way that professional philosophers do when conducting philosophical analysis. This, I propose, is

what philosophical analysis consists in.

The inductivist picture provides an appealing explanation of the fundamental character of analysis, as an inquiry into an external matter in which what is available in the armchair—our category judgments—functions as our principal source of evidence. It lays to rest the “paradox of analysis” in the most straightforward way: we are looking for a criterion that is not represented in our heads, and is therefore not the basis for the judgments we use in our search. It is no surprise if the answer to our question is surprising.

There is nothing about the inductivist picture, however, that in itself provides a vindication of analysis. Why should we suppose that there is a fact of the matter about what makes something knowledge? And even if there is, why should we suppose that our judgments about individual cases are sufficiently reliable to help us to find it?

To get some purchase on these issues, I do something that is as close as TOYF gets to armchair philosophy: I propose a certain theory of reference, on which (very roughly) a term or mental representation of a category picks out whatever category we would apply it to, if we had all the evidence we could ask for. By harnessing our beliefs involving a category term to determine the term’s reference, this dispositional approach guarantees (when circumstances are right) that the term will pick out a determinate category and that those same beliefs will be, or upon acquiring enough evidence will evolve into, a dependable guide to category membership.

Many questions remain, of which I will mention three. First, do we have reason to believe that the categories we analyze are of any objective interest? That they are not reflections of a parochial taxonomy that would better be replaced than exhaustively studied? To provide some sort of reassurance that our categories are worth the bother of analysis, I draw an analogy with natural kind concepts. Our notion of, say, gold, may have its origins in the cultural significance of the metal, which is to say, our “starter theory” of gold may be focused on properties of gold that are chiefly of local interest. But because

the theory has a certain causal structure, it has a tendency, once more facts come in, to home in on the explanatory cluster in the vicinity, a cluster that is built around gold's atomic structure and that is of a significance that quite transcends the theory's origins. I show how some philosophical concepts may home in on important aspects of the local explanatory landscape in much the same way.

Second, as remarked above, many of the most revealing case judgments in philosophical analysis—that a Gettierized belief is not knowledge, for example—have an invulnerable air. It seems to many armchair philosophers that there is no further knowledge about knowledge, or about human psychology in general, that might call on them to revise these judgments. It is not straightforward to explain this psychological phenomenon of “case certainty” using the inductivist picture. Furthermore, it points to an important difference between some philosophical concepts and natural kind concepts, undermining the analogy that I use to assuage worries about parochialism. A good part of *TOYF* sets out to tackle these problems.

Third and finally, how to account for what many philosophers take to be, along with Williamson, the long-term failure of analysis? On the one hand, the inductivist approach offers a striking account of that failure: our judgments concerning membership of any given category are dictated by a rich, complex theory; perhaps, then, there will be no succinct summary of these patterns of judgment that has the character sought by philosophical analysts. On the other hand, that leaves the project of vindicating analysis in a rather uncomfortable place. I surmise, and give what I hope is a sophisticated argument for supposing, that even in those cases where there is no such summary, the activity of analysis is a particularly effective and efficient way to map the local explanatory landscape. We may not learn what we wanted, but we learn something worthwhile. And thus we will not regret the construction and evaluation of all those clever, convoluted thought experiments that consumed so many armchair hours.

## References

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press, Oxford.