Are Thought Experiments Done in the Head? Ule, Wittgenstein and Enactments in Thought

Nenad Miščević
University of Maribor

Abstract

The paper discusses Andrej Ule's Wittgensteinian perspective on the nature and structure of thought experiments (TEs). It briefly turns to the *Investigations* rich with fictional examples, *bona fide* TEs, and briefly suggests how the rich tradition of debating them, which has marked the entire history of analytic philosophy, can be systematized. The starting point is to understand the internal structure of the functioning of thought-experimenting. Attention to the stages of a TE offers a link between the inner structure of the respective TE and the historical vicissitudes of traditions from which it arises. Wittgensteinian tradition(s) are a fine example of this general rule. The dialogue with Ule focuses on the issue of the nature of thought-experimenting.

Keywords: thought experiments, mental models, rule following, anti-mentalism, public practice

Ali potekajo miselni poskusi v glavi? Ule, Wittgenstein in izvedbe v misli – povzetek

Prispevek obravnava wittgensteinski pogled na naravo in strukturo miselnih poskusov Andreja Uleta. Najprej se na kratko ustavi pri *Raziskavah* in njenih številnih fikcijskih primerih, ki so pristni miselni poskusi, in predlaga, kako bi bilo mogoče bogato tradicijo razpravljanja o njih, ki je zaznamovala celotno zgodovino analitične filozofije, sistematično urediti. Prvi korak predstavlja razumevanje notranje strukture delovanja miselnega poskusa. Le če upoštevamo stopnje, ki si sledijo pri izvedbi miselnega poskusa, lahko notranjo strukturo določenega miselnega poskusa povežemo z zgodovinskimi okoliščinami njegovega nastanka. Wittgensteinska tradicija oziroma tradicije so dober primer tega splošnega pravila. Dialog z Uletom se osredotoči na vprašanje narave miselnega eksperimentiranja.

Ključne besede: miselni poskusi, mentalni modeli, sledenje pravilu, anti-mentalizem, javna praksa

Introduction

Let me start by saying that I am happy and thankful for the opportunity to write a paper in honor of Professor Andrej Ule.[1] We have known each other since the early eighties when he was on the committee during the defense of my PhD thesis, and we soon became friends.

---

1   I would like to thank to Andrej Ule, Gabriele Mras, Maja Malec and Olga Markič.

He is the most prominent Slovenian Wittgensteinian; most analytic work on Wittgenstein in Slovenia has been either conducted by him, or inspired and guided by him.

In this paper I want to consider Wittgenstein's use of imaginary examples, prominent and ubiquitous in his work. Nowadays, Wittgenstein scholars (e.g. Cora Diamond, 2002; Martin Cohen, 2004) routinely call the particular instances of such use "thought experiments," and we shall be doing the same, shortening the expression to "TE." I understand a thought experiment is an investigative procedure conducted "in the armchair", which normally involves:

(1) the formulation of experimental design pointing to:
     (i) the determination of the goal(s), in particular the thesis/theory to be tested, and
     (ii) the construction of a (typically) counterfactual scenario to be considered;
(2) the presentation of the thus-constructed scenario to the experimental subject (either the author of the scenario herself, or the interlocutor),[2]
(3) the (typically imaginative) contemplation of the scenario and some piece of reasoning,
(4) the decision ("intuition") concerning the thesis/theory to be tested.

Once this result is achieved, it can be, and often is, compared with the results of other similar thought experiments.[3]

In the first part we turn briefly to Wittgenstein's *Philosophical Investigations* (1963), rich with fictional examples, some of which have become centerpieces of analytic philosophy. Unfortunately, due to space limitations, we have to limit ourselves to a few sketchy remarks (for a slightly more detailed story, see Miščević, 2013).

---

2 At a philosophy seminar, the design of the scenario is usually given as a written text. This is what is colloquially called a "TE."

3 My proposal was inspired by the extant characterizations in the literature. For example, Sophie Roux notes about TEs that they are counterfactual, involve a concrete scenario, and have a well-determined cognitive intention (2011: 19). Compare also with the definition of TE in the *Stanford Encyclopedia* (due to J. R. Brown and Y. Fehige, 2011): "Thought experiments are devices of the imagination used to investigate the nature of things. Thought experimenting often takes place when the method of variation is employed in entertaining imaginative suppositions."

   Tamar Szabó Gendler, in her proposal of characterization of a TE, also stresses imagination and cognitive purpose:
     (a) Thought-experimental reasoning involves reasoning about a particular set of circumstances (which may be specified in more or less detail), described at a greater level of specificity than that of the conclusion.
     (b) The reasoner's mode of access to the scenario is via imagination rather than via observation.
     (c) Contemplation of the scenario takes place with a specific purpose: the confirmation or disconfirmation of a certain hypothesis or theory. (2004: 1155)

   In brief: "To perform a thought experiment is to reason about a scenario /.../ which is imaginary /.../ with the aim of confirming or disconfirming some hypothesis or Theory." (*Ibid.*).

   Let me also mention James Robert Brown who in his early and pioneering work starts with a fine picture: "Thought experiments are performed in the laboratory of the mind," and then further develops it:
     Beyond that bit of metaphor it's hard to say just what they are. We recognize them when we see them: they are visualizable; they involve mental manipulations; they are not the mere consequence of a theory-based calculation; they are often (but not always) impossible to implement as real experiments either because we lack the relevant technology or because they are simply impossible in principle, as when frictionless planes or a universe devoid of all matter are involved. (2010: 1)

In the second and most important part, focused on Ule's work, I will investigate the general issue of the nature of TEs. Ule has generously discussed my view of TEs in his recent paper "Mental Models in Scientific Work" (2017), written from a clearly Wittgensteinian perspective. Here, I would like to return the favor and discuss his view in detail.

TE's in the *Philosophical Investigations* – from stages to trails and traditions

The road to the later work is paved with imaginative examples. In *Philosophical Remarks* (1975), Wittgenstein invites us to "[i]magine a gearbox whose lever can take four positions. Now of course it can only take these positions in succession, and that takes time; and suppose it happened that it only ever occupied one of these positions, since the gearbox was then destroyed." And then he asks: "Wasn't it still a gearbox with *four* positions?" His point is to compare it with language. "Similarly, I would like to say in the case of language: What's the point of all these preparations; they only have any meaning if they find a use" (*ibid*: 59).

Philosophical Investigations (1963) famously opens with a quote from St. Augustine followed by the injunction(s) to the reader to imagine a few obviously non-actual scenarios; first, a very elementary scenario of ordering and getting red apples, and then the scenario involving two builders using a very elementary "language" to talk to each other. The scenario is then followed by other similar ones, most of them involving builders; let me call the whole big "sketch" – "Builders." The other two prominent imaginary scenarios in the *Investigations* are the one focused on the (im)possibility of private language, and the one concerning rule following.[4] As I mentioned above, these imaginary examples are all seen nowadays as thought-experiments.[5] We can add to them some "quasi-TEs," scenarios that remind one of a TE but lack important elements that define full-blown TEs.[6]

In his *Investigations* (1963: viii), Wittgenstein himself famously describes the remarks as "sketches of landscapes." "This book is really only an album," he warns the reader. A lot of space within the album is occupied by sketches of real and imaginary examples, and the rest of it is filled mostly with the discussion of such examples. Some of the sketches/examples contain many smaller sketches. Indeed, it is typical of Wittgenstein to organize several small-size examples into a larger imaginative project; we can talk about micro-TEs being brought together into a meso- (middle size-) TE, or even a large-scale, macro-TE. The "Builders TE" seems to me to be either a quite large meso-TE or a macro-TE constructed of smaller ones. The larger "Private language

---

4   For Ule's approach to them, see Ule, 1990: Section 3.4.

5   For example, Cohen (2004) and Diamond (2002).

6   See the division "Invented language-games" in Baker and Hacker, 2005: 61 ff.

TE" contains the "Beatle micro-TE", and so on. For Wittgenstein the crucial point of such construction is variation, and his intention is to teach the reader the complexities of the everyday and of the ordinary.

	TEs	*Quasi-TEs*
MACRO-	Builders	??
MESO- *(middle sized)*	Private language Rule following	4 positions gear-box
MICRO-	Beetle (within Private language)	Trying to raise my arm with a torn sinew

I have discussed the three main TEs (Builders, Rule-Following and Private Language) in some detail from a contemporary, somewhat non-Wittgensteinian perspective, in Miščević (2013). Here, I want to point out a way of understanding the traditions or at least trails that have grown out of them. In fact, they have been growing even out of the smaller TEs within the larger ones. Here is David Stern's comment on the uses of the "Beetle TE":

> The story of "Wittgenstein's beetle," like the story of the builders in section 2 of the *Philosophical Investigations*, is one of a small group of imaginary scenarios which play a leading role in that book's central argument.
>
> It is also one of the most frequently retold, and most frequently cited, passages in the book. Like Plato's cave, it has taken on an afterlife, both as a standard item in the philosophy curriculum, and as a parable that has been read in a bewildering variety of ways. It has even given its name to a recent collection of such philosophical narratives, Martin Cohen's *Wittgenstein's Beetle and Other Classic Thought Experiments*. (Stern, 2007: 250).

Remember the stages of a typical TE, which I presented above. I mentioned understanding as stage one following stage zero – the construction and presentation of the scenario of the TE by the author/speaker. Understanding of the matter by the subject/interlocutor is usually taken as unproblematic, but in the case of Wittgenstein a critical trail might start here. The really radical reactions (positive but re-interpretive as well as negative-critical) might focus on the very understanding of the proposed train of thought, and thus reject the very construction, or at least demand a re-interpretation of the apparent proposal. Let me illustrate what I mean.

Consider the way Stern presents the Beetle TE from **§**293 of the *Investigations*, in which it appears that the author is developing his thoughts in a dialogue with a somewhat less sophisticated interlocutor. After telling the story of the beetle, Wittgenstein writes: "But suppose the word 'beetle' had a use in these people's language? – If so it would not be used as the name of a thing." The question comes from the interlocutor, the answer is presumably Wittgenstein's. No, says Stern, we just have two characters, the

narrator and the interlocutor, and there is no reason that the voice of the former should be Wittgenstein's. The literal understanding of the very text is now challenged. This changes the game completely: the upshot will be that both sides – the narrator and the interlocutor – are engaged in a somewhat pointless discussion. Stern argues:

> Wittgenstein's beetle has been used by most of his interpreters as an opportunity to articulate the arguments that underlie his theory of the nature of sensation; it would be closer to the truth to say that it is an excellent illustration of his attack on the intuitions that feed such theories. (Stern, 2007: 267)

A slightly less radical reaction, questioning the first and the second stage, is directed to the Builders and their simple language. In his now classic paper "Wittgenstein's Builders" from the late 1950s, Rush Rhees expresses doubts about the very imaginability of the proposed scenario:

> But I feel there is something wrong here. The trouble is not to imagine a people with a language of such a limited vocabulary. The trouble is to imagine that they spoke the language only to give these special orders on this job and otherwise never spoke at all. I do not think it would be speaking a language. (1959: 178)[7]

A much milder disagreement is to be found, as one might expect, in Charles Travis (2006). He does not focus on the early stages of Builders, but on the generalization and account of it. On the usual reading, Wittgenstein is stressing the plurality of uses, in the sense that a single quasi-sentence from a builder can mark very different speech acts: once, the builder is asking for a slab, next time pointing to one, and so on. Travis is minimizing this plurality: "One might be tempted to see the point here as merely about illocutionary force, but it is not the real point" (2006: 23). For him the point of the TE lies in the occasion's sensitivity to what counts as a slab: do only the whole ones count, or halves as well, or even the more fragmented items?

In general, the less dramatic proposals, both critical and re-interpretive, will focus on the later stages. The proponent might even agree with the intuition(s) suggested to the reader and with its (or their) generalization, and then find fault with the immediate consequences that Wittgenstein seems to be drawing from it. For example, micro-TEs from the Builders suggest a plurality of language-games related to the plurality of forms of life, the correlation being one to one.[8]

The critic might attack the connection made. Does to each form of life correspond a particular language-game? What if all people in reality speak the same language, but live different forms of life (as proposed by Newton Garver, 1990 and Donald Barry, 1996)?

---

7   For historical sources of the trail, see M. F. Burnyeat (1987).

8   As suggested by von Savigny (1991).

The least radical reaction to a proposed TE is to accept the suggested intuition, generalization and immediate consequence, but to question even the later stages: the proposed explanation of the generalization, and its place within the wider context of the relevant philosophical sub-discipline. Consider the Private Language TE. The two classics among the commentators, Norman Malcolm and P. F. Strawson, accept Wittgenstein's proposal of what is to be meant by "private language" and his immediate conclusion that a language naming private objects is impossible. The disagreement is about further consequences: for Strawson the alleged consequence that pain cannot be named is untenable. For Malcolm, the moral of the story is that we must reject the naming model, and understand our talk of pains differently.

A more difficult problem is to locate the most important reactions to the Rule-following TE, from C. Wright and J. McDowell to Cora Diamond, informed, of course, by a confrontation with Kripke's challenge. For example, how deep does Kripke's skeptical reading go? Is it a denial of some deep element in the construction of the TE or a re-interpretation of its immediate conclusions and their explanation, or does it belong more to the wider framework of integrating the Rule following considerations into a more general epistemological picture?

In any case, attention to stages offers a connection between the inner structure of the TE and the historical vicissitudes of traditions springing from it. This is typical of the understanding of the longer history of each given TE in the history of philosophy; for instance, the Evil demon or Social contract. These are long-term histories, each lasting around three hundred years if we only count modern philosophy. If we could bring together short-term and long-term stages, this would make the historical process more intelligible. For instance, take the relatively recent variations of the two aforementioned thought-experimental oldies. The Evil demon has been innovatively transformed into (or replaced with) the Brain-in-a vat, whereas the Social contract has generated the Original position proposal. What is the relation between such historical changes, and the usual micro-variations of a given TE? Once we have worked out the answer, we shall be able to integrate the meta-philosophy of TEs with information from history of philosophy. Wittgenstein's trails are a fine example of this general rule.

Explaining TEs: the place of the mental

Ule's & Wittgenstein's understanding of TEs: rules & skills vs. mental models

We are nowadays used to understand TEs as processes in thought, or exercises in thinking. I am personally in favor of "mental models" understanding of TEs. I view them as belonging to a wide family of imaginary enactments in thought, together with some mathematical problem-solving attempts, with the literary and artistic use of imaginary scenarios, as well as with various kinds of meditations. Call any such way to understand

TEs "mentalistic." Wittgenstein offers, at least in some important places in his work, a series of caveats against the very idea of thinking and inner thought as real, explanatory items. Let me start by quoting a famous passage from *Philosophical Investigations* which Ule endorses and then uses in his critique of the mentalistic approach (I added the original German expression in brackets):

> If there has to be anything "behind the utterance of the formula" it is *particular circumstances*, which justify me in saying I can go on – when the formula occurs to me.
>
> Try not to think of understanding (*Verstehen*) as a "mental process" (*seelichen Vorgang*) at all. For *that* is the expression which confuses you. But ask yourself: in what sort of case, in what kind of circumstances, do we say, "Now I know how to go on," when, that is, the formula *has* occurred to me? –
>
> In the sense in which there are processes (including mental processes) which are characteristic of understanding, understanding is not a mental process (*ist das Verstehen kein seelischer Vorgang*).
>
> (A pain's growing more or less; the hearing of a tune or a sentence: these are mental processes.) (Wittgenstein, 1963: par. 154)

One can develop a sketch of an account of the use of imaginary examples, i.e. of TEs in the spirit of Wittgenstein, and this is what Ule did in his recent paper (2017). I will add references to some other Wittgenstein scholars. Ule believes that it is possible to consider potential models of theories as mental models. Here is his most important formulation:

> Mental modeling in science thus seems to be an integral part of public practices of the usage of non-linguistic or para-linguistic representations of empirical structures in finding their full theoretical explanation, and is only secondarily a part of individual mental practices. (*Ibid*: 231)
>
> I agree, but would like to add that our ultimate justification of intuitions is often based on something like *blind rule-following*, as described by Wittgenstein in his *Philosophical Investigations*. Such intuitions are often only implicitly taken in regard to mental modeling, and not explicitly formulated. Such blind rule-following may be regarded as a further, not-justified, but indubitable foundation of other more explicit rules and "moves" in mental modeling. According to Wittgenstein, blind rule-following is not something purely internal or "minded," but forms a part of *common practice* of rule-following of people in a certain life-form. For Wittgenstein, the seemingly purely mental character of basic intuitions is rather a linguistic or rhetorical illusion, and not a mental fact. (*Ibid*: 230)

What then follows is the just quoted passage of par. 154 of Wittgenstein's *Philosophical Investigations*. It is certainly a Wittgensteinian proposal. Ule is rather moderate in his endorsement of certain mental processes, although he assigns a secondary role to them. Below, I shall quote a few lines from remarks by other interpreters of Wittgenstein, who are more radical than Ule.

A historical consideration in favor of this reading of Wittgenstein comes from his rejection of the traditional view of state or activity that is supposed to terminate a TE, and of other related enactments in thought, particularly mathematical imaginative exercises. In his early stage, Wittgenstein had nothing against the view; discussing matters with Russell, himself a great admirer of intuitions, he started turning against the view. The dramatic event was Brower's lecture in Vienna in 1928, which brought Wittgenstein back to philosophy as well as to reflections on intuitionism. In the decade that followed, as Kim van Gennip puts it, "Wittgenstein sees the basis of our knowledge not to be reflected in principles or propositions, but in our actions and reactions" (2011: 574).

Let me systematize the Wittgensteinian proposal a bit. I shall call its first component pan-grammaticism: this is the claim that imaginary examples (our "TEs") are "grammatical examples" and their philosophical use is grammatical; it concerns the "grammar" of the language we use, and its point is to clarify misunderstandings about the use. Here is a quote from Beth Savickey who, in her *Wittgenstein's Art of Investigation* (1999, Ch. 7), formulates the connection between Wittgenstein's ideas of "grammar" and his treatment of imaginary examples in a simple and perspicuous way: "[H]e invents particular cases and uses fictitious intermediate cases to clarify the grammar of our words" (*Ibid*: 181). Then, she refers to Wittgenstein's own words:

> What kind of investigations are we carrying out? Am I investigating the probability of the cases that I give as examples, or am I investigating their actuality? No, I am just citing what is possible and am therefore giving grammatical examples. (1993: 187)

Again, the use of imaginary examples points away from actuality and in the direction of what is possible and impossible, even in that of the limits of possibility. For Wittgenstein, this is the area of grammar: it is the rules of language that delineate possibilities. The fictional-imaginary points to possibilities, possibilities depend on grammar, therefore, the investigation of TEs is a grammatical investigation. As Edgar Alan Poe would put it, we do things "logic and common usage so commanding." But where does our use of grammar belong? Kim van Gennip mentions the standard response to this question, which I also endorse, aptly placing it in its historical context:

> /.../ Wittgenstein's criticisms of Russell's account of indubitability and perception signify a fundamental disagreement in their outlook on epistemology in general. Where Russell seeks for empirical, propositional premises of our knowledge, assigning these premises the status of indubitability, Wittgenstein sees the basis of our knowledge not to be reflected in principles or propositions, but in our actions and reactions. (2011: 574).

But what grounds our actions and reactions? Not mental processes, Wittgenstein seems to claim. "Language is self-sufficient, it needs no external grounding," claims Ule in his book *Filozofija Ludwiga Wittgensteina* (1990:102).

So here is the second Wittgensteinian component – anti-mentalism: TEs are not to be explained by appeal to human mental processes. Andrej Ule comments the above quoted par. 154 from *Investigations* in this spirit. I shall assume that Ule's interpretation of Wittgenstein is correct. Other informed readers of Wittgenstein have come to similar conclusions. For example, Beth Savickey defends roughly the same view. After providing a number of examples, she concludes: "If we imagine particular cases for each of these examples, or look at how we actually use these words, the philosophical idea of thinking (as a mental process which accompanies speaking) dissolves." (Savickey, 1999:187).[9]

As she further explains, the problem with speaking of the mind as the locality of thinking is the following:

> /.../ we tend to consider it *the* locality of thinking (and treat the connecting cases as merely metaphorical); for we do not want to claim that we can think with a pencil on paper and not, at the same time, in our minds. Our confusion is grammatical, however, and it involves the different uses of the expression "the locality of thinking."

Wittgenstein thus writes in *The Blue Book* (1969: 16):

> "Where does thinking take place?" We can answer: on paper, in our head, in the mind. None of these statements of locality gives the locality of thinking…we must not be misled by the similarity of their linguistic form into a false conception of their grammar. As e.g. when you say: "Surely the real place of thought is in our head." /.../ And it is, furthermore, extremely important to realize how, by misunderstanding the grammar of our expressions, we are led to think of one in particular of these statements as giving the real seat of the activity of thinking.

Ule now has to square this anti-mentalism with the assumption that scientists (and philosophers) use some kind of mental modeling in their practice. Thus, return to the quotations above and note the moderate character of the proposal in comparison with Savickey. Mental processes are involved in the work with examples and models, but their role is "only secondary." However, in his characterization of the final result, namely the "intuition," he also goes radical and denies that for Wittgenstein they constitute a mental fact at all: "For Wittgenstein, the seemingly purely mental character of basic intuitions is rather a linguistic or rhetorical illusion, and not a mental fact." (Ule, 2017: 230)

Further, he continues after quoting *Investigations*:

> I believe that the same holds true for many types of intuitive insight. We have to ask ourselves in what case, in what circumstances, do we say: "Now I see what you mean.", "Now I see how I can do this or that.", etc.; when *has* a formula, a graph,

---

9 She quotes some rather radical formulations, for example from Zettel (1981) where Wittgenstein addresses the philosophical temptation to equate "thinking" with a mental process that occurs in our brain, by examining the idea of "the locality of thinking": "One of the most dangerous of ideas for a philosopher is, oddly enough, that we think with our heads or in our heads." (605)

or a mental picture occurred to me? Sometimes, when we hit upon the seemingly indubitable statements about intuitive reasoning, we actually stumble across indubitable, but *public criteria* for meaningful use of statements on our basic intuitions, and not on some *a priori* true proposition. (*Ibid*. 231)

A contemporary, physicalistically-minded reader, might propose the way to avoid the appeal to purely mental processes by replacing it with an appeal to the neural processing. The Wittgensteinian tradition is radically against such an attempt.[10]

Ule aptly summarizes its spirit and targets the neurologists who try to explain thinking in terms of parallel neural processing:

/.../ their most "deleterious" mistake is their confusion of the internal computational (or parallel) processes taking place in the brain (which possibly cause mental states) with socially-based, everyday criteria of recognition and classification of, and knowledge about, the content of mental states. (Ule, 2005: 375)

And he connects this idea to Wittgenstein:

However, it has its principled limitations also pointed out by Wittgenstein. He opposed not just the idea that non-material processes taking place within an individual are inaccessible, but also the idea of the "internal mechanism" or "internal processes" in general. For him, the "external–internal process" difference is simply too small to be useful in explaining the predicaments arising from everyday discourse and, even more so, philosophical discourse on mental phenomena. (*Ibid*: 377)

Certainly, this is not the only way to read Wittgenstein; one could take a more moderate view and claim that when he rejects private mental states the stress should be on "private," and that a Wittgensteinian can talk in good consciousness about mental states and processes as long as she does not view them as mysterious private items.[11]

Similarly with understanding, Baker and Hacker discuss it in detail (see chapter XVII "Understanding and ability" of their 2005 book) without making strongly antimentalist assumptions. Their discussion of Wittgenstein's attack on the "mental state"

---

10 Here is a very clear formulation by Wittgenstein himself:
No supposition seems to me more natural than that there is no process in the brain correlated with associating or with thinking; so that it would be impossible to read off thought-processes from brain-processes. I mean this: if I talk or write there is, I assume, a system of impulses going out from my brain and correlated with my spoken or written thoughts. But why should the *system* continue further in the direction of the center? Why should this order not proceed, so to speak, out of chaos? The case would be like the following – certain kinds of plants multiply by seed, so that a seed always produces a plant of the same kind as that from which it was produced – but *nothing* in the seed corresponds to the plant which comes from it; so that it is impossible to infer the properties or structure of the plant from those of the seed that it comes out of – this can only be done from the *history* of the seed. So an organism might come into being even out of something quite amorphous, as it were causelessly; and there is no reason why this should not really hold for our thoughts, and hence for our talking and writing. (*Zettel*, note 608., p. 105)

11 Malcom Budd offers a very thought-friendly account (see, e.g. 1989/1991: 128 ff.)

classification of understanding focuses on "state" as opposed to other candidates (disposition and so on), and does not deal with mental/non-mental in any detail. Ule, on the other hand, stresses the non-importance of the mental, and seems ready to go quite far in this regard; additionally, the more radical reading is the only one that proposes a new and challenging view of imaginary examples/TEs, so I shall abide by it.

The problem of anti-mentalism

I shall not repeat the general criticism, but focus on our specific topic, namely the understanding of imaginary examples. Consider the two TEs: first, the omniscience we discussed in the section above, and second, the famous Builders from *Investigations*.

Consider first the strong version denying that the processes involved (like understanding) are mental processes at all. So, here is the question. How does the reader of *Lecture on Ethics* (1965) reconstruct the scenario? She is told to suppose that she "is an omniscient person." How does one suppose one is such-and-such a person if not in one's thought, most plausibly in one's imagination? Do not think of supposing as a mental process, the Wittgensteinian might advise the reader. However, the very verb "think" and noun "supposing" are about something that sounds mental. Behave as if you were supposing? How does one do this behaviorally? Write a letter demanding a Nobel Prize for one's future achievements (like the one Barack Obama got for peace)?

By the way, what does omniscience involve? Knowing "all the movements of all bodies in the world" and "all the states of mind of all human beings"? Why should one suppose that one knows such states of mind if states of mind do not count? Or, are perhaps states of mind not mental (and not neural)? Well, Wittgenstein changed his mind later, some years after the *Lecture*. But what would he have written then? Suppose you know behavioral dispositions? This does not sound very promising. Finally, what about knowing? It is not a mental state. How do I suppose I know without supposing that I believe, and so on? If I do this, I end up supposing I am in a mental state.

Things did not change in Wittgenstein's later work. Consider a later TE, the Builders, and the well-known opening imaginary scenario:

> Now think of (*Denke nun an*) the following use of language: I send someone shopping. I give him a slip marked "five red apples." He takes the slip to the shopkeeper, who opens the drawer marked "apples"; then he looks up the word "red" in a table and finds a color sample opposite it; then he says the series of cardinal numbers (*Grundzahlwörter*). (Wittgenstein, 1963: § 1)[12]

And then I have to think of the shopkeeper pairing "Grundzahlwörter" with apples. In short, I have to "think of" a process in time: one paired with the red apple, two with the next one, and so on. It does not sound like a routine topic of anybody's thinking (unless

---

12 There is a mistake in the translation: "Grundzahlwörter" are "basic numerals" not "numbers."

one is specialized in teaching elementary arithmetic to shopkeepers selling apples). Perhaps Wittgenstein refers to a routinely acquired manner of "thinking"? But what kind of routine, non-mentally directed training, should the person have in order to do the "thinking" related to the activity of the shopkeeper? Wittgenstein (in §5) asks the reader to look back at (*betrachtet*) the example in §1, and suggests that this will perhaps result in getting an inkling; German has a more romantic "*so ahnt man*." Where does the "*betrachten*" and "*ahnen*" take place? Not in any loud linguistic activity. How does Wittgenstein avoid downright behaviorism at this point?

Next, how does one explain what the reader is doing? The latter cannot be building a mental picture ("model") of the transaction if there are no intellectually-efficient mental entities around to appeal to them in the explanation. However, Wittgenstein only criticizes the "purely mental" (*seelisch*) as opposed to the physical, a charitable reader of *Investigations* might conjecture. Not at all. We noted his anti-neurology stance. The idea that the reader's cognitive neural mechanisms do the work is unacceptable.

Consider now a weaker version that Ule sometimes prefers and that offers more promise. Mental processes are there, but they do not do the work. It is the external, social "common usage" that accounts for TEs. However, the examples I listed do not describe common social practices, which we know by heart, and do not need any thinking to imagine them. Nobody is omniscient; there is no routine social practice associated with omniscience. We do not routinely write books in which we put down all the facts of the world. Well, the moderate Wittgensteinian might reply that the omniscient behavior is just generalization of our non-omniscient one. Indeed, but then it is the reader's individual thinking that does the generalization. I have not been routinely taught at school to generalize from my daily ignorance-marked pseudo-episteme to the magnificent episteme of an omniscient person; I have to do the thinking myself.

I am aware that there are still weaker versions of anti-mentalism. One is certainly "social behaviorism."[13] Characteristic of these weaker versions is the acceptance of the mental coupled with the minimization of its role. Take a theoretician who derives, say, the normativity of logic from the requirements of the dialogue without going into the mentalistic story about how each individual participant complies with them.[14] Perhaps one such theory might account for some aspects of TEs, but I doubt it can do it for all of their most important features.

Let me conclude with a dilemma for Ule, contrasting scientific and philosophical modeling. Assume that philosophical modeling is all grammatical. What about the scientific one? Not grammatical, for certain. But then, can philosophers not help themselves to similar, non-grammatical techniques? Ule is trying to dodge the issue. As we

---

13 See Gier (1982).

14 See Dutilh Novaes (2015) (I would like to thank Nenad Smokrović for pointing out this article to me).

saw, for him "[m]ental modelling in science thus seems to be an integral part of public practices of the usage of non-linguistic or para-linguistic representations." (Ule, 2017: 231, the fuller context quoted above)

What about any original proposal for modeling? When Nash proposed his revolutionary game-theoretical models for various human practices, his innovation emphatically was not just "internalization of the public practice"! And it came as a result of hard thinking. In no way was it merely the result of the blind following of already existing rules. Moreover, most ironically, Wittgenstein's revolutionary ideas have themselves been quite far from being merely part of already existing "public practices" he could have internalized in his native Vienna. He was a thinking and thoughtful innovator, not a blind follower of pre-existing rules and public practices.

References

Baker, G. P., and Hacker, P. M. S. (2005). *Wittgenstein: Understanding and Meaning, Part I: Essays.* Second extensively revised edition. Oxford: Blackwell.

Barry, D. K. (1996). *Forms of Life and Following Rules: A Wittgensteinian Defence of Relativism.* Leiden: E. J. Brill.

Brown, J. R. (2010). *Laboratory of the Mind: Thought Experiments in the Natural Sciences. Second edition.* London: Routledge.

Brown, J. R., and Fehige, Y. (2011). "Thought Experiments." In Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy* (Fall Edition 2011). https://plato.stanford.edu/archives/fall2011/entries/thought-experiment/.

Budd, M. (1989/1991). *Wittgenstein's Philosophy of Psychology.* Oxford: Routledge.

Burnyeat, M. F. (1987). "The Inaugural Address: Wittgenstein and Augustine's *De Magistro*." *The Proceedings of the Aristotelian Society, Supplementary Volumes*, 61, pp.1– 24.

Cohen, M. (2004). *Wittgenstein's Beetle and Other Classic Thought Experiments.* Oxford: Wiley-Blackwell.

Diamond, C. (2002). "What if X isn't the Number of Sheep? Wittgenstein and Thought-Experiments in Ethics." *Philosophical Papers*, 31 (3), pp. 227–250.

Dutilh Novaes, C. (2015). "A Dialogical, Multi-Agent Account of the Normativity of Logic." *Dialectica,* 69 (4), pp. 587–609.

Garver, N. (1990). "Form of Life in Wittgenstein's Later Work." *Dialectica*, 44 (1–2), pp. 175–201.

Gendler, T. Szabó (2004). "Thought Experiments Rethought – and Reperceived." *Philosophy of Science*, 71, pp. 1152–1163.

Gier, N. F. (1982). "Wittgenstein, Intentionality, and Behaviorism." *Metaphilosophy,* 13 (1), pp. 46–64.

Miščević, N. (2013). "Wittgenstein's Thought-Experiments." *Concept*, 42 (1–2); pp. 1–21.

Miščević, N. (2017). "Reply to Andrej Ule." In Borstner, B., and Gartner, S. (eds.), *Thought Experiments between Nature and Society: A Festschrift for Nenad Miščević*, Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 234–238.

Rhees, R. (1959). "Wittgenstein's Builders." *The Proceedings of the Aristotelian Society,* 60, pp. 171–186.

Roux, S. (2011). "Introduction: The Emergence of the Notion of Thought Experiments." In Ierodiakonou, K., and Roux, S. (eds.), *Thought Experiments in Methodological and Historical Contexts,* Leiden: Brill, pp. 1–33.

Savickey, B. (1999). *Wittgenstein's Art of Investigation*. Oxford: Routledge.

Stern, D. (2007). "The Uses of Wittgenstein's Beetle: *Philosophical Investigations* §293 and Its Interpreters." In Kahane, G., Kanterian, E., and Kuusela, O. (eds.), *Wittgenstein and His Interpreters: Essays in Memory of Gordon Baker*, Oxford: Blackwell, pp. 24 –268.

Travis, C. (2006). *Thought's Footing: A Theme in Wittgenstein's Philosophical Investigations*. Oxford: Oxford University Press.

Ule, A. (1990). *Filozofija Ludwiga Wittgensteina: od* Traktata *do* Filozofskih raziskav. Ljubljana: Znanstveni inštitut Filozofske fakultete.

Ule, A. (2005). "How Can One Catch a Thought-Bird?: Some Wittgensteinian Comments to Computational Modelling of Mind." *Synthesis Philosophica* 40 (2), pp. 373–388.

Ule, A. (2017). "Mental Models in Scientific Work." In Borstner, B., and Gartner, S. (eds.), *Thought Experiments between Nature and Society: A Festschrift for Nenad Miščević*, Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 219–233.

van Gennip, K. (2011). "Wittgenstein on Intuition, Rule-Following, and Certainty: Exchanges with Brouwer and Russell." In Kuusela, O. and McGinn, M. (eds.), *The Oxford Handbook of Wittgenstein*, Oxford: Oxford University Press.

von Savigny, E. (1991). "Common behaviour of many a kind: *Philosophical Investigations*, section 206." In Glock, H.-J., and Arrington, R. L. (eds.), *Wittgenstein's Philosophical Investigations: Text and Context*, Oxford: Routledge.

Wittgenstein, L. (1963). *Philosophical Investigations*. Oxford: Blackwell.

Wittgenstein, L. (1965). "A Lecture on Ethics." *The Philosophical Review*, 74 (1), pp. 3–12.

Wittgenstein, L. (1969). *The Blue Book*. In Wittgenstein, L., *The Blue and Brown Books*, Oxford: Blackwell, pp. 1–74.

Wittgenstein, L. (1975). *Philosophical Remarks*. Edited by Rhees, R. Oxford: Blackwell.

Wittgenstein, L. (1981). *Zettel* (2nd ed.). Oxford: Blackwell.

Wittgenstein, L. (1993). "Philosophy." In Klagge, J., and Nordmann, A., (eds.), *Philosophical Occasions 1912–1951*, Indianapolis: Hackett Publishing Company.