

Unsafe reasoning: a survey

Paulo Faria
UFRGS/CNPq
paulo.faria@ufrgs.br

abstract Judgments about the validity of at least some elementary inferential patterns (say *modus ponens*) are *a priori* if anything is. Yet a number of empirical conditions must in each case be satisfied in order for a particular inference to instantiate this or that inferential pattern. We may on occasion be entitled to presuppose that such conditions are satisfied (and the entitlement may even be *a priori*), yet only experience could tell us that such was indeed the case. Current discussion about a perceived incompatibility between content externalism and first-person authority exemplifies how damaging the neglect of such empirical presuppositions of correct reasoning can be. An externalistic view of mental content is ostensibly incompatible with the assumption that a rational subject should be able to avoid inconsistency no matter what the state of her empirical knowledge may be. That fact, however, needs not be taken (as it often is) as a *reductio* of externalism: alternatively, we may reject that assumption, adding to the agenda of a philosophical investigation of rationality an examination of the vicissitudes of logical luck. I offer an illustration and defense of that alternative.

keywords Inference; Presupposition; Content externalism; Logical luck

This is how the story goes: you should not be held responsible for what evades your control. It is not for you to stop the autumn leaves from falling. ‘Ought’, we are told, implies ‘can’.

The story has been challenged. Moral luck has been a lively issue in ethics since the exchange between Bernard Williams and Thomas Nagel which sparked the debate over thirty years ago (see WILLIAMS 1976,

NAGEL 1976). The reckless driver who passes a red light may be lucky enough that no pedestrian is crossing the street; but, red lights or no, if a child comes running after a ball and gets hit, that will make all the difference. The *moral* difference: 'We feel sorry for the driver, but that sentiment co-exists with, indeed presupposes, that there is something special about his relation to the happening, something which cannot merely be eliminated by the consideration that it was not his fault.' (WILLIAMS 1976: 28).

Jurisprudence, a realm where consequences *always* matter, has long known the doctrine of strict liability, which features in judicial settlements of torts and criminal accountability. In the English case of *Rylands v. Fletcher* (1868), the defendants' underground water reservoir caused an old mine shaft owned by the plaintiff to collapse; although the court found that the defendants were not negligent, they were still strictly liable for damages.¹

As we'll see in a while, it's not so much that 'ought' will not imply 'can', just like that. Rather, there will be more to 'can' than meets the eye. Specifically, there are some tricky questions to be faced concerning how actual control (or lack thereof) relates to control in what we keep calling, for want of better words, close enough possible worlds. The unlucky driver who hit the child had actually no control over the outcome; but *it might have been otherwise*: that's how issues of negligence bear on such matters.

Now moral luck itself is not, you guess, my topic; but the background assumption about the connection between responsibility and control lies at the heart of what I am after here.

My aim, in brief, is to bring into sharper focus what Roy Sorensen, in a highly perceptive (if, sad to say, largely unheeded) paper, proposed to call 'logical luck' (SORENSEN 1998). It was Sorensen's original insight that current debates on content externalism (specifically as they bear on the apriority of our logical abilities) are fruitfully illuminated when set against the framework provided by comparison with the *prima facie* unrelated topic of moral luck. I may eventually have some reservation to voice about Sorensen's unabashedly consequentialist approach to blameworthiness; unlike Sorensen, moreover, I will make essential use of a distinction between excusable and inexcusable ignorance: something which the prevailing approach to the 'externalism and inference' debate

(relying, as it does, on the ‘slow switching’ thought experiments introduced by Burge in 1988) has made all but invisible.

Grub first, then ethics, recommended Brecht. Let me follow the advice by starting with the daily bread of some very elementary, ground floor pieces of reasoning. Somewhat fancier scenarios, most notably involving switches between Earth and Twin Earth, will show up as we go along. And then, when the ease of such flights of fancy will have made things really impossible for us (as impossible they have actually become in the present state of the art) we’ll dutifully turn to ethics for some guidance. First things first, though.

If this paperweight is a stone from Walden Pond, then something is a stone from Walden Pond. Likewise, if everything must perish, then Vanina must perish. Or so it seems; but appearances mislead.

For, come to think, what exactly (if anything, that is) was meant by ‘this paperweight’ in the preceding paragraph? Suppose I wrote that at home, where I actually have a paperweight on my desk; and suppose further that, as I wrote it, I was actually attending to that (one and only) paperweight on my desk. Now if *that* is a stone from Walden Pond, then, sure, something is a stone from Walden Pond. Yet mark how much you had to assume.

Or how little: for it all boils down to ‘this paperweight’ *meaning something* – to the predicate ‘is a stone from Walden Pond’ being, truly or falsely, ascribed to something.

Again, if everything must perish, so must Vanina – provided, that is, ‘Vanina’ names something. (The only Vanina I can think of is a character in a Stendhal short story. And sure I know better than infer from ‘Everything must perish...’ that a fictional object must perish.)

That is rather trite, yet enough to get us going. For it would seem that whether an entailment holds should be settled by logic alone; yet haven’t we just seen it to be otherwise? After all, what is at stake in both my examples is a matter of truth – of contingent truth at that: the existence and uniqueness of a certain piece of stone, the identity of Vanina, whoever she may be. And what has logic to do with that? ‘Logic’, wrote Wittgenstein, ‘precedes every experience – that something is so. It is before the How, not before the What.’ (*Tractatus Logico-Philosophicus*, 5.552, in the Ogden translation). Whether something exists – let alone

whether it remains the same or undergoes change – is supposed to be no concern of logic at all. In Hegel’s epigram, logic is ‘the exposition of God as he is in his eternal essence before the creation of nature and a finite mind’ (1812: 50). Both remarks are apt to sound as dark sayings of old, representative of their authors’ shared penchant for the oracular, yet here is what Quine has to say about the inference rules we know as \exists -introduction and \forall -elimination:

“The principle embodied in these two operations is the link between quantifications and the singular statements that are related to them as instances. Yet it is a principle only by courtesy. It holds only in the case where a term names and, furthermore, occurs referentially. It is simply the logical content of the idea that a given occurrence is referential. The principle is for this reason, anomalous as an adjunct to the purely logical theory of quantification”. (QUINE 1953: 146)

What exactly Quine (of all people) may have meant by ‘the logical content’ of an idea is a matter for conjecture. But if he is right about \exists -introduction and \forall -elimination, then much the same would seem to apply to pretty any inference rule you could think of. From ‘Laura is Canadian’ and ‘Laura is blonde’ we are entitled to infer ‘Laura is a blonde Canadian’ – provided, that is, the name ‘Laura’, in both occurrences, denotes the same individual. The rule which licenses the inference ‘is simply the logical content of the idea’ that the two occurrences are co-referential.

These ‘ideas’ (about existence, uniqueness or what have you) are *empirical assumptions* – they may well turn out to be false; and we should count ourselves rather lucky that, on the whole, ‘we live in a world in which objects and stuffs are generally stable, and we are not switched by evil demons from one planet to another’ (FALVEY 2003: 238).

Yet, if once in a while, shifts happen. Suppose as I’m coming home in the afternoon I notice this beautiful Golden Retriever dog playing around in my neighbor’s front yard. I stop for a while to pet my new acquaintance, who turns out to be very amiable. As I walk home I think ‘That’s a very friendly dog.’

Now, a couple of days later, same scene – or so it seems. Here’s the nice front yard with its blooming bushes and this beautiful golden dog running around. Again I stop, hoping to attract the dog’s attention, resort-

ing perhaps to whistling or finger clapping, yet this time to no avail: the dog keeps running nonstop around the yard, heedless of my inviting moves, barking up every other tree. Maybe he's spotted a cat, who knows. I walk home thinking: 'That's a very restless dog'.

Am I now entitled to infer that there is a dog in my neighborhood who is both friendly and restless?

1. Fa
2. Ga
3. $\exists x (Fx \wedge Gx)$

Well, suppose my neighbor is a breeder of Golden Retrievers and what I successively spotted on those two occasions were a pair of siblings from the same litter – call them Harry the Friendly and Barry the Restless. As things go, Harry is not excitable at all, while Barry is of a rather unfriendly disposition. Suppose further there are no other dogs in the neighborhood. So my conclusion is just false, and my reasoning is unsound – a plain fallacy of equivocation. Its form is not 1-3 but rather:

4. Fa
5. Gb
6. $\exists x (Fx \wedge Gx)$

And the trouble lies in the way my mistaken empirical assumption (that there was one single dog which I encountered twice) impinges on my grasp of the logical form of the inference I performed – specifically in my taking 4-6 to be of the form 1-3. For it's not as if I inferred validly, except that my inference relied on a tacit (and false) identity premise (namely that Harry = Barry) – so that my reasoning was really an enthymeme:

7. Fa
8. Gb
9. a=b
10. $\exists x (Fx \wedge Gx)$

I'll come back to that suggestion; in the meantime, let me just remark that whatever I may have presupposed (about existence, uniqueness, or what have you), is better *not* construed as amounting to an additional premise in my reasoning. That way, we'll see, madness lies.

Now there is a powerful recipe to make such uneventful mishaps as my unlucky inference about the dogs in the neighborhood sound really dramatic – so much so that what will seem to be at stake is nothing less

than ‘the apriority of our logical abilities’ (BOGHOSIAN 1992: 17). And that is to indulge in some science fiction – or, as philosophers of an analytic bent prefer to say, do some thought experiments.

Now I won’t rehearse once again Putnam’s Twin Earth or Burge’s ‘arthritis’ fictions. I just want to focus on this methodological feature shared by both: the appeal to pairs of epistemically indiscernible counterpart worlds.

Here is then the subject in Putnam’s scenario, who thinks that the glass she holds is filled with the insipid, odorless etc. liquid she calls ‘water’. And here are the two counterpart worlds: in *Possible World 1*, ‘water’ denotes H₂O; in *Possible World 2*, ‘water’ denotes XYZ. That, as we all know, will impinge on where meanings are to be found; but, for current purposes, here is what really matters to me: the relevant difference (H₂O / XYZ) is *stipulated* to be inaccessible from the first person perspective: that’s built *into* the thought experiment as we tell it.²

And this is what engaging in thought experiments (stipulating, as Kripke says, a possible world), amounts to: we, the runners of the thought experiment, know what our characters are stipulated to ignore. *Their* ignorance (among other things, of what may, on occasion, impinge on the soundness of their reasoning) is, accordingly, excusable if any ignorance is.³

That’s, of course, what raises the vexed problem about externalism and first person authority. As Tyler Burge aptly sums it up: “How can one individuate one’s thoughts when one has not, by empirical methods, discriminated the empirical conditions that determine those thoughts from empirical conditions that would determine other thoughts?” (BURGE 1988: 116)

The slow switching thought experiments introduced by Burge in ‘Individualism and Self-Knowledge’ (1988), in which Earth and Twin Earth coexist in the same possible world, are meant to bring into focus the *diachronic* dimension of the First Person Authority Problem: namely, the impact of shifting contextual parameters on a subject’s ability to retain and redeploy conceptual contents. Here’s Burge, and bear with me to have him quoted at some length:

“Suppose that one underwent a series of switches back and forth between actual earth and actual twin earth so that one remained in

each situation long enough to acquire concepts and perceptions appropriate to that situation. Suppose occasions where one is definitely thinking one thought, and other occasions where one is definitely thinking its twin. *Suppose also that the switches are carried out so that one is not aware that a switch is occurring. The continuity of one's life is not obviously disrupted.* So, for example, one goes to sleep one night at home and wakes up in twin home in twin bed – and so on. (Your standard California fantasy.) Now suppose that, after decades of such switches, one is told about them and asked to identify when the switches take place. The idea is that one could not, by making comparisons, pick out the twin periods from the ‘home periods’”. (BURGE 1988: 115. *My emphasis*, PF).

As we should expect, slow switching will have a bearing on ‘any reasoning that takes place over time, hence any reasoning’ (BURGE 1998: 363). For one thing, the subject’s ability to assess rightness of inference would seem to be jeopardized by his unawareness that he’s been switched.

That was the problem raised by Boghossian in ‘Externalism and Inference’: externalism ‘is inconsistent with the thesis that our thought contents are epistemically transparent to us (...) this is true in a sense that falsifies another important and traditionally held view – that we can detect *a priori* whether our inferences are logically valid or not.’ (BOGHOSSIAN 1992: 13)

And, to be sure, that would be very disturbing indeed. After all, the main interest (and, just possibly, the main promise) of externalism lies in its acknowledgement of the impact of exposure to changing contexts on the constitution of thought contents. That’s precisely what explains the interest aroused, in the literature about externalism, by cases of context-switch, often illustrated with such elaborate fantasies about space-traveling from Earth to Twin Earth and back, interplanetary abductions and like exercises in science fiction. At the end of the day, such exercises should have been just a device to graphically describe a range of much less extraordinary phenomena which, if externalism is right, take place in a variety of situations prompted by exposure to differences (for instance, though not only, of a cultural nature) between the contexts in which the rational capacities of a single subject must be exercised.

Anyway, that's the setting of Boghossian's problem about externalism and inference (BOGHOSSIAN 1992). Boghossian's argument has the form of a *reductio ad absurdum*: the truth of externalism implies the possibility of undetectable errors in reasoning, due to unperceived shifts in propositional content. That possibility clashes with the transparency of mental content; therefore, externalism is false.

Suppose that, having enjoyed a happy childhood on Earth, I am someday carried away to Twin Earth. And suppose further, as Burge asks us to, that 'that the switches are carried out so that one is not aware that a switch is occurring. The continuity of one's life is not obviously disrupted.' (BURGE 1988: 115). I just wake up on Twin Earth in twin bed and everything looks exactly as before. (Remember, that was built into the thought experiment). In due time, according to externalist common wisdom, my usage of the term 'water' comes to mean what the linguistic community to which I now belong uses it to mean: namely, XYZ. And here I am, inferring from the conjunction of true premises

11. I enjoyed playing in water (=H₂O) when I was a kid (*from memory*)

12. This glass is full of water (=XYZ) (*from current perception*) the false conclusion

13. This glass is full of the same liquid I enjoyed playing in when I was a kid.

Worse still, the fallacy I fall prey to is not comparable to the usual fallacy of equivocation, in which an ambiguity is neglected, and the reasoner is in principle in a position to detect and rectify, on a wholly *a priori* basis, the flaw in her reasoning. In the slow switching scenarios, there is just nothing the subject can do to prevent or fix the irrationality, short of undertaking an empirical investigation of the environment, and of her own personal history.

Worse still, given the conditions built into the slow switching stories, it is (to put it mildly) unclear what such an 'empirical investigation' could possibly look like. It's not as if there might be traces, like a flight ticket from Earth to Twin Earth in the inside pocket of your jacket, or custom papers attached to your passport, or a message from home in the answering machine. The two worlds were stipulated to be indiscernible, exact duplicates (apart from the single 'external' difference): were it not so, we would have learnt nothing from the fictions.

That's how logical luck turns out to be, as much as its moral equivalent, a matter of indiscernible counterparts. Here is a putatively sound reasoning:

14. Fa

15. Ga

16. $\exists x(Fx \wedge Gx)$

Now add subscripts to your non-logical constants, if only to mark that they are different tokens of the same type, hence not logically assured to be co-referential:

17. Fa_1

18. Ga_2

19. $\exists x(Fx \wedge Gx)$

And there we are: in *Possible World 1*, $\delta('a_1') = \delta('a_2')$, and the inference is sound. Switch to *Possible World 2*, where $\delta('a_1') \neq \delta('a_2')$ and 17-19 is (again) a fallacy of equivocation. Trouble is, possible worlds 1 and 2 are not like, say, Barcelona and Valencia. They're way more like a world in which I am 'sitting by the fire, dressed in a winter cloak, touching this paper with my hands, and the like' and one which is an artifact of 'some evil spirit who is supremely powerful and cunning and who has expended all his energy in deceiving me'.⁴

The perception that there's not much that the victims of slow switching can do to detect the external sources of their possible logical shortcomings – a perception which I find to be widespread (if mostly tacit) in the literature – answers for the surprising willingness, displayed by friends and foes of externalism alike, to devise exculpating moves as a response to these scenarios of logical misfortune.

After all, the *only* difference between the lucky and the unlucky reasoners lies wholly beyond their ken. I suppose Boghossian speaks for most writers in the field when he writes: 'It seems to me that there is an immediately recognizable sense in which there can be no difference *in respect of rationality* between [*the reasoner in Possible World 1 and that in Possible World 2*]. It seems implausible in the extreme to say that they differ in their capacity to *reason*' (BOGHOSSIAN 1992: 27).

That is of course an eminently plausible appraisal, given what the differences between the two contrasting worlds are supposed to be. Small wonder, then, that the choices on offer are (again, with the remarkable

exception of SORENSEN 1998) a variety of exculpating moves, designed to shield the rationality of the possibly unlucky reasoner against the contingencies of context-shifting. (See, e.g. SCHIFFER 1992, BURGE 1998, LUDLOW 2004, SOSA 2005, COLLINS 2008.)

The *internalistic* exculpating move will predictably recoil from the broader and unsafer landscape to an inner domain, a.k.a. narrow content, sealed off from the contingencies of external causation, accident, and luck. Here is how Boghossian introduces it: “If, then, it is also true that there is an important sense in which [*the reasoner’s*] behavior *makes sense from his point of view*, we would appear to have here an argument for the existence of a level of intentional description which conserves that sense’. (BOGHOSSIAN 1992: 28).⁵

A bit more surprising are the *externalistic* exculpations, paramount among which is the Schiffer-Burge “anaphoric” view of content preservation.⁶

The main idea here is that the reiteration, in an occurrent episode of thinking, of the content of a past thought is made possible by a dependency relation comparable to that which holds between relative pronouns, and other anaphoric expressions, and their antecedents in the linguistic constructions in which they feature. In ‘Laura was confident that she would get the prize’, the pronoun ‘she’ designates Laura: its semantic value is determined by the anaphoric antecedent which is the proper name – as the value of a bound variable in first order quantification is determined by the quantifier which is its anaphoric antecedent. Hence, the unhappy reasoning 11–13 would get reinterpreted as:

20. I enjoyed playing in water when I was a kid

21. This glass is full of *that* ↑.

22. This glass is full of the same liquid I enjoyed playing in when I was a kid.

Here I resort to ‘that ↑’ to mark the anaphoric dependence of ‘water’, as it was (presumably) tokened by the reasoner in premise 21, on its occurrence in premise 20. And what we have as a result is, small wonder, a valid argument with a false premise: as ‘water’ in 20 denotes H₂O (the thought content here being supplied by preservative memory), premise 21 amounts to the false judgment that the Twin Earth glass is full of H₂O.

A more surprising externalistic path to exculpation is provided by Peter Ludlow’s “Orwellian” theory of content preservation. On Ludlow’s

theory, ‘it is not the job of memory to record contents, but rather to provide information about past episodes relative to current environmental conditions.’ (LUDLOW 1996: 316).⁷ As in Orwell’s *Nineteen Eighty-Four*, the past is rewritten from the standpoint and priorities of the present. Hence when I, on Twin Earth, recall having played in water as a kid, the content of my “Orwellian” memory is the false judgment that I played in twater (XYZ). The outcome is, again, a valid judgment with a false premise:

23. I enjoyed playing in water (XYZ) when I was a kid. (*False*)

24. This glass is full of water (XYZ). (*True*)

25. This glass is full of the same liquid I enjoyed playing in when I was a kid. (*False*).

Now there’s no question that all these different construals somehow manage to take into account the fact that, as Boghossian says, the reasoner’s behavior ‘makes sense from his point of view’. What is not so conspicuous is the sheer absence, *built into the very terms of the slow switching thought experiments*, of any other point of view against which the subject might try and assess the soundness of her reasonings. And that is my complaint against the freewheeling use of so many thought experiments in contemporary analytical philosophy: we end up losing some of our grip on what things look like in some real life, down to earth surroundings.

The shared assumption underlying all the exculpating moves we have considered is explicitly stated by David Sosa in a recent, very illuminating paper: ‘*Ignorance is insufficient for incoherence*: inferring subjects are in principle in a position to avoid invalidity, no matter what their state of knowledge (indeed, no matter what the truth of their premiss beliefs).’ (SOSA 2005: 219).

As the attentive reader will not have failed to notice, there is a further assumption at work here: namely, that ignorance is always excusable. Which, as I stressed, makes perfect sense in the fancy setting of the slow switching thought experiments. *There* was indeed virtually nothing that the switched subjects could do to prevent the fallacies of equivocation they were prone to – hence the appeal of the exculpating moves we briefly reviewed.

When it comes to down to earth reasoning, though, there’s much that a subject is justifiably expected to know and duly take into account.

Suppose I left an empty glass on my desk as I went out for a haircut. Back home an hour later, I notice that the glass on the desk is now full and infer, at my own peril, that the glass which was previously empty is now full. At my own peril, I stress, as I am neglecting the likelihood that the housemaid, profiting from my absence to tidy up my office, brought a new, full glass to replace the previous, dusty one, as she often does. Neglect of this easily available piece of information is comparable to my not taking into account the well-known fact that my neighbor is a breeder of Golden Retriever dogs, which increases the likelihood that it was not the same dog I met on those two occasions.

Or think, for a change, of reasonings involving tensed statements. I may know for sure that Laura is singing; and then, with equal certainty, that she is dancing. But to infer from this pair of true statements that someone is both singing and dancing betrays an inexcusable neglect of well-known facts about the truth-conditions of tensed statements.⁸

Of course, in down to earth reasoning, we take a lot for granted. As Strawson wryly remarks: "Our methods, or criteria, of reidentification must allow for such facts as these: that the field of our observation is limited; that we go to sleep; that we move. That is to say, they must allow for the facts that we cannot at any moment observe the whole of the spatial framework we use, that there is no part of it that we can observe continuously, and that we ourselves do not occupy a fixed position within it." (STRAWSON 1959: 32)

At t_1 , I see an object a_1 and think 'This is F '. At t_2 , I see an object a_2 and think 'This is G '. Then I draw the conclusion: 'Something is both F and G '. When am I entitled to that conclusion? Well, of course, whenever $a_1 = a_2$. But was that a tacit premise, and my inference, therefore, an enthymeme? That way, I suggested, madness lies. For suppose

26. Fa_1

27. Ga_2

are not enough to for you to infer ' $\exists x(Fx \wedge Gx)$ '. After all, you need to make sure that ' a_1 ' and ' a_2 ' are co-referential. That is, you need the further premise:

28. $a_1 = a_2$

But that's not gonna do, either. For now you have to make sure that ' a_1 ', as it occurs in 26, and ' a_1 ' as it occurs in 28 are also co-referential;

and ditto for ‘ a_2 ’ as it occurs in 27 and ‘ a_2 ’ as it occurs in 28. At which point it is manifest that you are embarked on a vicious, Lewis Carroll-style, regress.⁹

The moral is, to be sure, that reasoning about changing objects (reasoning for which existence, uniqueness or permanence assumptions may be crucial) is an essentially unsafe business, in which we are bound to *riskily* take an awful lot for granted. That’s what makes us subjects of logical luck, and the extent of our accountability will be a matter of the excusableness of our ignorance. Which brings us to the point at which, as I suggested, we may usefully turn to ethics for some guidance.

The legal philosopher H. L. A. Hart opens a paper on criminal responsibility with the following imagined dialogue: ‘I didn’t *mean* to do it: I just didn’t think.’ ‘But you ought to have thought.’ (HART 1961: 136)

‘Ought’ implies ‘can’, to be sure. It is not, in such cases as Hart discusses, as if there was *nothing* the defendants could possibly have done to avoid the unfortunate outcome. Rather: what *actually* evades the control of the subject *would* be under her full control in close enough (and, moreover, *epistemically accessible*) possible worlds. The owner of a rundown house may, as a matter of fact, ignore that the ceiling is about to collapse. *Yet that would be a manifest case of inexcusable ignorance.* The relevant information was fully available, had she only cared to go after it.

In such down to earth cases, the information which the subject actually lacks *is* available: the subject would be apprised of it if only she cared enough to know. Not so on the slow switching scenarios – hence the exculpating moves of which I was complaining.

In what is likely to be one of the most gnomic elucidations of the concept of inexcusable ignorance, Wittgenstein writes: ‘That I am a man and not a woman can be verified, but if I were to say I was a woman, and then tried to explain the error by saying I hadn’t checked the statement, the explanation would not be accepted.’ (WITTGENSTEIN 1969, § 79)

Epistemology, including the epistemology of reasoning, has been, for most of its history, this cantankerous discipline in which we keep asking whether and how we possibly know anything. Much of what I was arguing for may be summed up in the idea that it is high time for a change of gear. As I draw some encouragement to think so from Elizabeth Anscombe, let me give her the last word: ‘The interesting enquiries about

knowledge, once one has given up the attempt to characterize it as a sort of belief that satisfies certain conditions, concern what everyone, or everyone in certain cultures, can be assumed to know once grown-up and reasonably competent: and the relation of some claims of knowledge to questions like: “How do you know?” (ANSCOMBE 1993: 32).¹⁰

¹ Compare Article 927, single paragraph, of the Brazilian Civil Code: ‘The defendant will be liable for damages, *independently of guilt*, in such cases as are stipulated in statute, or when the defendant’s usual activity implies, by its own nature, risk for the rights of others.’ [*My emphasis, PF.*]

² And here is, for the sake of comparison, Burge’s scenario: the subject thinks that the painful ailment in his thigh is arthritis. In *Possible World 1*, ‘arthritis’ denotes an inflammation of joints (and the subject has a false belief). In *Possible World 2*, ‘arthritis’ denotes a number of rheumatoid ailments (and the subject has a true belief). Again, the two possible worlds are indiscernible “from the first person perspective”.

³ That thought experiments are stipulations; that these stipulations are such that it will often be the case that we, who make them, will know an awful lot that their protagonists are designed to ignore: these points have been forcefully made, as was fairly pointed out by an anonymous referee to this journal, in SILVA FILHO 2006, to which the reader is referred. Where I seem to part company with Silva Filho is in my disinclination to take the workings of an ‘omniscient narrator’ as casting suspicion on the methodology of the externalist (or, for that matter, any) thought experiments. There’s nothing wrong with stipulating possible worlds, no matter how contrived, and nothing to regret about omnisciences which are as much artifacts of the stipulating game as is, say, the ignorance of the slow-switched subjects about their surroundings. I have no qualms with the narrator’s omniscience; what troubles me is her all but unavoidable (logical) leniency: again, let there be no doubt, a matter of *how the story is told*.

⁴ *Honni soit qui mal y pense!* I hope to elaborate on these dangerous liaisons elsewhere. (For those who may care: the Descartes quotes come from Harry G. Frankfurt’s translation in FRANKFURT 1970: 10, 13).

⁵ The move bears comparison with Kant’s forceful “shrinking” of the proper domain of moral assessment to the inner realm where a pure will operates by itself, sealed off alike from the vagaries of causation, contingency and luck. See Bernard Williams’ perceptive discussion in WILLIAMS 1976.

⁶ See SCHIFFER 1992, BURGE 1998.

⁷ The theory is further articulated in LUDLOW 2004.

⁸ ‘Our inferential practice exploits a convenient insensitivity to the specific context of our judgments. As a result, we sometimes run into trouble. One example may be the philosophi-

cal tendency to suppose that change is contradictory: the disposition to connect judgments inferentially (...) might overflow into the inference from the true judgement 'Judy is dancing' and the true judgement made a second later 'Judy is not dancing' to the self-contradictory judgement 'Someone is both dancing and not dancing'. (WILLIAMSON, 1997: 652-3)

⁹ Argument credit: John Campbell (see CAMPBELL 1987; and, for a brief restatement, CAMPBELL 1994: 75-6).

¹⁰ Earlier drafts of this paper were presented at the Porto Alegre colloquium 'Semantics and Cognition' (July 2009), and at the Fortaleza conference 'Thought Experiments and the *A Priori*' (August 2009). I thank audiences at both venues for their responses. I am especially grateful to André J. Abath and Rogério Severo (in Porto Alegre), and to Ernest Sosa, Christopher Hill, Jonathan Ichikawa, Anand Vaidya, Albert Casullo and Jens Kipper (in Fortaleza) for their comments. My gratitude extends to the students who attended a graduate seminar on Cognitive Dynamics I gave at my home university in 2009; to an anonymous referee for this journal, who very perceptively singled out the most pressing task with respect to which I remain an insolvent debtor (namely that of going beyond my vague and blatantly insufficient appeal to the notion of 'epistemically accessible possible worlds' in articulating criteria for inexcusableness of ignorance); and, last and not least, to the editor of this special issue, André Porto, for the unfailing patience and good humor with which he met my delinquency in complying with renewed submission deadlines.

Referências bibliográficas

ANSCOMBE, G. E. M. 1993. 'Knowledge and Essence' in Josep-Maria Terricabras (ed.), *A Wittgenstein Symposium* (Amsterdam / Atlanta, GA: Rodopi): 29-35

BOGHOSSIAN, Paul A. 1992. 'Externalism and Inference', *Philosophical Issues* 2: 11-28

BURGE, Tyler. 1988 'Individualism and Self-Knowledge', *Journal of Philosophy* 85: 649-663, reprinted in LUDLOW & MARTIN (1998): 111-127

_____. 1998. 'Memory and Self-Knowledge', in LUDLOW & MARTIN (1998): 351-370

CAMPBELL, John. 1987. 'Is Sense Transparent?', *Proceedings of the Aristotelian Society* 88 (1987/1988): 273-92

_____. 1994. *Past, Space, and Self* (Cambridge, MA: The MIT Press).

- COLLINS, John, (2008). 'Content Externalism and Brute Logical Error', *Canadian Journal of Philosophy* 38: 549-574
- FALVEY, Kevin. 2003. 'Memory and Knowledge of Content', in Susana Nuccetelli, ed. (2003). *New Essays on Semantic Externalism and Self-Knowledge* (Cambridge, MA & London: The MIT Press): 219-240
- FRANKFURT, Harry G. 1970. *Demons, Dreamers, and Madmen: The defense of Reason in Descartes's Meditations* (Indianapolis / New York: Bobbs-Merrill)
- HART, H. L. A. 1961. 'Negligence, Mens Rea and Criminal Responsibility', *Oxford Essays in Jurisprudence*, reprinted in *Punishment and Responsibility: Essays in the Philosophy of Law* (Oxford: Clarendon Press, 1968): 136-157
- Hegel, G. W. F. 1831. *Science of Logic*, transl. A.V. Miller, preface by J. N. Findlay (London, Routledge, 2004)
- LUDLOW, Peter. 1998 'Social Externalism and Memory: a Problem?', *Acta Analytica* 14, reprinted in LUDLOW & MARTIN (1998): 311-317
- _____. 2004. 'What Was I Thinking? Social Externalism and Shifting Memory Targets', in R. Shantz (ed.), *The Externalist Challenge* (Berlin: de Gruyter): 419-426
- LUDLOW, Peter & MARTIN, Norah, eds. 1998. *Externalism and Self-Knowledge* (Stanford: CSLI Publications)
- NAGEL, Thomas. 1976 'Moral Luck', *Proceedings of the Aristotelian Society*, supplementary volume 50, reprinted in *Mortal Questions* (Cambridge: Cambridge University Press, 1979): 24-38
- QUINE, W.V. 1953. 'Reference and Modality', in *From a Logical Point of View*, 2nd ed. (Cambridge, MA: Harvard University Press, 1980)
- SCHIFFER, Stephen (1992). 'Boghossian on Externalism and Inference', *Philosophical Issues* 2: 29-38
- SILVA FILHO, Waldomiro José da. 2006. 'O Autoconhecimento, o Narrador Onisciente, a Vida Comum', *Philósophos* 11: 287-303

SORENSEN, Roy A. 1998 'Logical Luck', *The Philosophical Quarterly* 48: 319-334

SOSA, David (2005). 'The Inference that Leaves Something to Chance', in Sanford Goldberg (ed.), *Internalism and Externalism in Semantics and Epistemology* (Oxford: Oxford University Press): 219-234

STRAWSON, P. F. 1959 *Individuals: an Essay in Descriptive Metaphysics* (London: Routledge)

WILLIAMS, Bernard. 1976. 'Moral Luck', *Proceedings of the Aristotelian Society*, supplementary volume 50, reprinted in *Moral Luck* (Cambridge: Cambridge University Press): 20-39

WILLIAMSON, Timothy. 1997. 'Sense, Validity and Context', *Philosophy and Phenomenological Research* 58: 649-654

WITTGENSTEIN, Ludwig. 1969. *On Certainty*, ed. G.E.M. Anscombe & G.H Von Wright, transl. Denis Paul & Elizabeth Anscombe (Oxford: Blackwell)