ON “BRIGHT COLOURS” IN KANT’S ARGUMENTATION: ANALOGIES, METAPHORS AND THOUGHT EXPERIMENTS

Sobre um “brilhante colorido” na argumentação kantiana: Analogias, metáforas e experimentos de pensamento

Rômulo Eisinger Guimarães

Universidade Federal de Santa Maria – Santa Maria, Brasil
r_eisinger_guimaraes@yahoo.com

Abstract: In order to carry out “a great and important piece of work, and that in a complete and lasting way”, Kant claims that one must join one’s “effort with that of the author”. The reading of a work, therefore, must try to embrace “the articulation or structure of the system, which yet matters most when it comes to judging its unity and soundness”. This structure and articulation, however, may not be so easily accessible; it may, rather, be latent under “bright colours [that] paint over and make unrecognizable” the argument’s essence, often confusing the reader, who “cannot quickly enough attain a survey of the whole” (*KrV*, A XIX). Now, it would be expected, then, that Kant, condemning such “ornaments”, would not employ those resources I n his writings. However, it is curious that not infrequently he makes use of such “bright colours”. The present paper aims at pointing out some passages of Kant's theoretical and practical philosophy in which it is possible to identify such “colouring”.
Keywords: Bright colours; Example; Analogy, Metaphor; Thought Experiment.

Resumo: A fim de se “realizar inteiramente e de maneira duradoura uma obra grande e importante, de acordo com o plano que lhe é proposto”, afirma Kant que deve o leitor “juntar o seu esforço ao do autor”. O ato de leitura de uma obra deve, portanto, tentar abraçar a “articulação ou a estrutura do sistema, que é o mais importante para se poder julgar da sua unidade e do seu valor”. Esta estrutura, contudo, pode não ser de tão fácil acesso; pode, antes, estar latente sob um “brilhante colorido [que] encobr[e], por assim dizer, e torn[a] invisível” o essencial da argumentação, confundido frequentemente o leitor, que não pode “alcançar, com suficiente rapidez, uma visão desse conjunto” (*KrV*, A XIX). Ora, seria de se esperar, então, que Kant, condenando tais “ornamentos”, não fizesse uso destes recursos em sua argumentação. Contudo, curioso é que, não raras vezes, serve-se Kant de um “brilhante colorido” em seus escritos. O presente artigo visa apontar algumas passagens da filosofia teórica e prática de Kant, nas quais é possível identificar tal “colorido”.
Palavras-chave: Brilhante colorido; Exemplo; Analogia; Metáfora; Experimento de Pensamento.

1.

Either to broaden a theory’s domain (by evidencing it) or refute a preceding one (either by deducing fallacies or deriving absurdities from its premises, or by providing counterexamples), philosophers and scientists often employed narrative resources in their works. From Plato’s parallel between artists and sophists (*Soph,* 235 a–b)[1] to Rutherford’s

[1] For quotations of works that were not originally written in the English language, I shall use the translations indicated in the bibliography (with some changes, when judged necessary). Regarding texts that either are not available in English language or whose translations I did not have access to, all translation is my

"planetary model" (cf. Rutherford, 1911, p. 688), from Locke's "white paper" (cf. *E*, PWL 1: B.II, C.I, § 2:205), to Einstein's clocks (cf. Einstein, 1905, pp. 894 ff), reasoning based on an existing relationship between relevant characters of different things – one might argue – cannot be separated from scientific development itself (whether in natural or human sciences).

It's not up to this modest paper – for I here would have neither the space nor the intention – to offer an austere separation between the different *modi operandi* of "relational thinking" – i.e., Analogy, Metaphor, Thought Experiment,[2] etc. Rather, I limit myself to a minimally sufficient (and, in some respects, arbitrary[3]) differentiation to only then point out as Immanuel Kant, a philosopher who believes that the "bright colours" of an argument "paint over and make unrecognizable the articulation or structure of the system,[4] which yet matters most when it comes to judging its unity and soundness" (*KrV*, A XIX), employs precisely such resources – now more, now less "colourful" – in his own writings.

Since this paper aims, for matter of order, to establish a sort of "landmark" among those thinking resources, it would seem important to point out, promptly, two things. First (and what seems more obvious), if even today a precise terminological distinction between such resources is problematic, at the time of Kant's writings, such demarcation could encounter even greater difficulties and some terms, one might argue, did not even exist – at least, not as we understand them today.[5]

---

responsibility (reason why "my translation" will not be indicated at the end of each quote). In order to facilitate the location of passages of already classic philosophical texts, for Aristotle's work, I shall use the Bekker pagination, and abbreviation *Poet* for his Poetics; for Avicenna's work, I took as reference the Arabic text Avicenna's De Anima, the psychological part of *Kitāb al-Najāt,* edited by Fazlur Rahman. I shall use the following citation system: (*De Anima*, c. (chapter), § (section): page in Rahman's edition)); for Descartes' work, I shall use the abbreviation *Med* for Meditations on First Philosophy and the following citation system: (*ABBREVIATION*, AT (volume number): page); for Kant's works, I follow the *Akademie-Ausgabe* and its citation system; for Locke's work, I shall use the abbreviation *E* for An essay concerning human understanding and the following citation system: (*ABBREVIATION*, PWL (volume): B. (book), c. (chapter), § (section): page); finally, for Plato's work, I shall use the Stephanus pagination and the abbreviations *Soph* for Sophist and *Rep* (followed by the book number) for Republic.

[2] Hencefoth: *TE*

[3] It would be to some extent "arbitrary" because, one might argue, either everything could be framed under the broad concept of "Metaphor", or that the metaphorical reasoning itself is just a sort of Analogy. However, for the sake of order, I distinguish Analogy from Metaphor in order to consider the former as the relationship between distinct objects and the latter, the phrasal construction that implies this relationship (cf. sections 2 and 3 of this paper).

[4] I.e., "that which first makes ordinary cognition into science" (*KrV*, B 860).

[5] If, on the one hand, is true both that "there is neither the expression 'TE' nor an explicit theory of the TE in Kant" (Buzzoni, 2011, p. 93), and that "Hans-Christian Ørsted introduced the term 'TE' to philosophy in 1811" (Stuart et al., 2018, p. 1), on the other hand, it seems undeniable that the practice of such reasoning

Secondly, Kant talks about "examples" <*Beispiele*>[6] – and that in a very broad sense. "Sharp[ening] the Judgment <*Urteilskraft*>", he claims, the examples, however,

> more usually do some damage, since they only seldom adequately fulfil the condition of the rule (*as casus in terminis*) and beyond this often weaken the effort of the understanding to gain sufficient insight into rules in the universal and independently of the particular circumstances of experience, and thus in the end accustom us to use those rules more like formulas than like principles. (*KrV*, B 173)

Curious, however, is: although for Kant are not the examples, but the Judgement's[7] healthy use what is indispensable for the cognition – so that he attributes to the examples the title of walkers or "walking frame <*Gängelwagen*> of the Judgment" (*KrV*, B 174) – Kant often employs to such "walkers" to illustrate (and, sometimes, in a very fruitful way) what he presents in his writings.

In light of this, I try, from now on, to highlight – and, as far as possible, to distinguish – some argumentative resources that, although do not substitute the cognition through the Judgment *per se*, are extremely valuable to understand – often, in their subtleties – some fundamental passages of the Kantian text. And to do so, in a walk-through of those heuristic thought acrobatics that we perform whether in order to address (by more or less evident comparison) a theory, an emerging problem, or, still, a problem that otherwise could not be satisfactorily illustrated, I approach, in this paper, Analogies (section 2), Metaphors (section 3), and TEs (section 4). And this in order to, hopefully, show how (and if) some passages of Kant's writings could fit the here presented

---

in both science and philosophy dates back much earlier. Indeed, Brown and Fahige draw attention to the fact that "we can go back even further and find in the work of the German philosopher-scientist Georg Lichtenberg (1742–1799) a tacit theory of a 'experiments with thoughts and ideas'" (Brown & Fehige, 2019). Furthermore, Plato's lent weapon (*Rep* I, 331 c ff), Avicenna's floating body in a void (cf. *De anima*, c. I, § 1: R 16), Galilei's falling bodies (Galiliei, 1954, p. 61 ff) and Newton's bucket (Newton, 1999, p. 412 ff) – to take a few examples – could easily argue for the idea that the history of those disciplines itself seems to be merged with the usage of such narrative arguments.

[6] It might be interesting to point out that the German language has, in addition to *Beispiel*, the word *Exempel* and that both, by common agreement between German native speakers, could be translated as "*example*". The (subtle) difference, I guess, rest on the fact that *Beispiel* [*bei*: by; *Spiel*: game, play] does not seem to have the same exemplary, the same model aspect of *Exempel*. This becomes clear with the use of the latter – and only this one! – in the sentence "*ein Exempel statuiren*", which could be translated as "*to set an example to someone*" (in the sense of the example that parents or guardians give or serve their children). On the other hand, *Beispiel* – precisely, I believe, because of the term *Spiel* – seems to include the flexibility (and the fictionality) – and, why not, the playful <*spielerisch*> character – of relationships that I intend to address here.

[7] Here understood as "the faculty of subsuming under rules, i.e., of determining whether something stands under a given rule (c*asus datae legis*) or not" (*KrV*, B 171).

definitions and to verify (how far could one go with) the use of such resources in these passages (section 5).

2.

Among the aforementioned reasoning resources, Analogies are those which Kant ostensibly employs. Broadly understood either as an "equality of relations" (Abbagnano, 2007, p. 55) – which could be considered its meaning for Logic and for ancient Philosophy –, or as a

> probable extension of knowledge, passing from a proposition that expresses a certain situation to another proposition that expresses a generally similar situation, or as an extension of the validity of a proposition, from a certain situation to a generally similar situation [so that there is no equality of relations, but only a generic similarity [i.e., its meaning for modern and contemporary Philosophy] (Abbagnano, 2007, p. 57),

an Analogy involves comparison between different objects based on accepted similarities between their relevant characters (cf. Bartha, 2019).

From demonstrative conclusions – evidencing relevant similarities "between two objects, or systems of objects" – to those that offer "minimal plausibility" (Bartha, 2019) – i.e., that show that the conclusions are not completely impossible (even if improbable) –, the reasoning by Analogy maintains "that since things are alike in some ways, they will probably be alike in others" (Blackburn, 2005, p. 13) and it comes to inferences – more or less valid – based precisely on these aspects common to objects or systems analogously compared. It's because these common aspects are relevant ones[8] that an analogical reasoning works: they are, as it were, the touchstone of its functioning.

Now, Kant's usage of Analogies – especially those he calls *philosophical* ones – is of particular importance since it's the (only) resource that our intellect has to (legitimately) think about what it cannot effectively cognize, i.e., about the supersensible.[9] His conception of philosophical Analogy must be distinguished both from

[8] One cannot infer, e.g., that because they live in the sea and have caudal fins, whales are fish since sharks also live in the sea, they also have causal fins and are, indeed, fish. And this is because there are *more relevant differences* between whales and sharks than anatomical similarities and the habitat common to both.

[9] It would be worth stressing here that, as van den Berg correctly points out, because of the insufficiency of our limited intellect to fully understand organisms in nature, reasoning by Analogy is also employed by

the *(immediate) inferences of the Understanding* – e.g., "*all humans are mortal* [...] [in which] lie already the propositions 'some humans are mortal', 'some mortal beings are human beings', [and] 'nothing immortal is a human being' [...] [as] immediate conclusions from the first one" –, the *(indirect) inferences of Reason* – e.g., "all scholars are mortal", which "can be concluded [...] only by means of an intermediate judgment" (*KrV,* B 360), since all scholars are humans and all humans are mortal –, and the *constitutive-qualitative Analogies of Mathematics*, in which, "if three members of the proportion are given the fourth is also thereby given, i.e., can be constructed" (*KrV*, B 222) – e.g., 2 is to 4, as 5 is to x, so that x is equal to 10.

In turn, the Analogy that Kant has in mind here concerns an

> identity of the relation between grounds and consequences (causes and effects), insofar as that identity obtains in spite of the specific difference between the things or those of their properties that contain in themselves the ground for similar consequences (i.e., their difference outside of this relation). (*KU*, 5: 464, footnote)

It's not, therefore, a constitutive Analogy; rather, it's a *regulative-qualitative* one: it's, on one hand, *regulative*, because it allows us to think about what we do not have sufficient means to cognize; it's, on the other hand, *qualitative*, because this thinking resource "is not the identity of two *quantitative* but of two *qualitative* relations, where from three given members I can cognize and give *a priori* only the *relation* to a fourth member but not *this* fourth *member* itself" (*KrV*, B 222).

Such thinking would be legitimate since, at some level, a link is maintained with the cognizable, i.e., "by the maintenance of its determinant [...] in the relationship with the beings of the sensible world and their respective qualities" (Hamm, 2013, p. 79). Now, if this is Kant's criterion of the legitimacy of a philosophical Analogy, it's not only illegitimate but apparently it seems very unjustified to compare, among themselves, by Analogy, things that I am completely unaware of, e.g., between what determines the behaviour of mermaids and what determines the behaviour of dogs.

What seems, yes, legitimate for me to establish, is a relationship between what serves as a guide to animals and what serves as a guide to human beings (as rational

---

Kant in his philosophy of biology. For an analysis on this aspect of analogical arguments in Kant, cf. van den Berg, 2018, pp. 67 ff.

beings), i.e., between instinct and Reason. For, even if I was not completely aware of the ground of the animals' actions, I could draw such a relationship in view of the effect that instinct and Reason have, respectively, on animals and human beings. And what, presumably, would justify "the usage of such a conclusion by Analogy is the *paritas rationis*, i.e., the sameness of the ground <*Einerleiheit des Grundes*>, for counting as members of the same genus animals and human beings" (Hamm, 2013, p. 72.): since "in spite of their specific difference, they [the animals] are still of the same genus as human beings (as living beings)" (*KU*, 5: 464, footnote).

Nevertheless, even if one cannot count as being "members of the same genus", it's legitimate to draw a degree of Analogy between, e.g., human beings and the highest being <*das höchste Wesen*>. For even though it lacks, here, "the possibility [...] for counting the highest being as part of one and the same species along with human beings" (*KU*, 5: 464, footnote), one could trace an analogical relationship between both, as causes, at least as to their effects perceived by me and which I attribute to them as causes – e.g., between works of art (in the case of human beings) and the world (considered as the work of the highest being).

Indeed, although

> the causality of the highest cause is that, with respect to the world, which human reason is with respect to its works of art [,] thereby the nature of the highest cause itself remains unknown to me: I compare [by Analogy] only its effect (the order of the world), which is known to me, and the conformity with Reason <*Vernuftmäßigkeit*> of this effect, with the effects of human Reason that are known to me. (*Prol*, 4: 360, footnote)

If so, such thinking, Kant claims, "concerns only language and not the object itself", considering that, he goes on, "I do not thereby cognize [the object-of-judgment] according to what it is in itself, but only according to what it is for me, that is, with respect to the world of which I am a part" (*Prol*, 4: 357).

Therefore, this *modus operandi* of reasoning cannot aim at the same demonstrative character as the logical inferences mentioned above.[10] For, Kant explains,

[10] Which concern what objects are, not what they are to me; because "all human beings *are* mortal" and not "all human beings are mortal *to me*".

> one can, of course, *think* of one of two dissimilar things, even on the very point of their dissimilarity, by means of an *Analogy* with the other; but from that respect in which they are dissimilar we cannot *draw an inference* by means of the Analogy, i.e., transfer this characteristic of the specific difference from the one to the other. (*KU*, 5: 464)

Now, from the *Critique of Pure Reason* we learned that, in the cognition realm – i.e., the cognizable – Reason[11] takes us to the objective limits of experience (cf. *KrV*, B 166). There, however, by Analogy, we could cognize (in a weak sense) what lies beyond these limits, "but only in relation to Reason's own complete use in the field of possible experience [...]. This is, however, all of the benefit that can [and should] reasonably even be wished for here, and there is cause to be satisfied with it" (*Prol,* 4: 361-62): for, since we cannot effectively cognize, we do the best with what we have.

Thus, since it is a "cognition" that has less to do with *cognizing* something objectively than with *understanding subjectively* – or making it understandable for ourselves how it behaves –, Kant maintains that "the expression suitable to our weak concepts will be: […] *as if*" (*Prol*, 4: 359) and not "the copula *is*", which distinguishes "the objective unity of given representations from the subjective [...] i.e., a relation that is *objectively valid* [...] from the relation of these same representations in which there would be only subjective validity, e.g., in accordance with laws of association" (*KrV*, B 142).

Although it "concerns only language and not the object itself", analogical reasoning, therefore, as "a perfect similarity between two relations in wholly dissimilar things" (*Prol*, 4: 357), remains far from a "letter-by-letter" identification between what is compared. Furthermore, one might argue, it is not equivalent to Metaphor – so that, although the latter has similarities with the former, they could be distinguished from each other.

3.

Between Analogy and Metaphor, Aristotle seems to offer the following distinction:

[11] Understood here as the set of our intellectual faculties.

> Analogy or proportion is when the second term is to the first as the fourth to the third. We may then use the fourth for the second, or the second for the fourth. Sometimes too we qualify the Metaphor by adding the term to which the proper word is relative. Thus the cup is to Dionysus as the shield to Ares. The cup may, therefore, be called 'the shield of Dionysus,' and the shield 'the cup of Ares.' Or, again, as old age is to life, so is evening to day. Evening may therefore be called, 'the old age of the day,' and old age, 'the evening of life,' or, in the phrase of Empedocles, 'life's setting sun.' For some of the terms of the proportion there is at times no word in existence; still the Metaphor may be used. For instance, to scatter seed is called sowing: but the action of the sun in scattering his rays is nameless. Still this process bears to the sun the same relation as sowing to the seed. Hence the expression of the poet 'sowing the god-created light.' There is another way in which this kind of Metaphor may be employed. We may apply an alien term, and then deny of that term one of its proper attributes; as if we were to call the shield, not 'the cup of Ares,' but 'the wineless cup'. (*Poet*, 1457 b 16 ff)

In view of this, we could be tempted to discriminate Analogy from Metaphor in the following way: the first would consist of a mere relation between terms (such as *a* is to *b,* as *c* is to *d*), and the second, of an "appropriation" thereof by language. As a figure of speech, Metaphor does not abandon that relationship established by Analogy; rather, it constructs it grammatically,[12] i.e., in a sentence. Thus, the Analogy pointed out is: *cup* is to *Dionysus*, as *shield* to *Ares*. The Metaphors, in turn, are: *cup of Ares* is equal to *shield*; *shield of Dionysus,* to *cup*.

However, the use of metaphorical language, as pointed out in Aristotle's example, does not exhaust the possibilities of this resource. That is because a Metaphor, in addition to establishing a relationship between two things in terms of their similarities or similes – i.e., to saying that Ares's cup is like, as it were, Dionysus' shield – enables *a much broader semantic field*. For, non-rarely, a relationship is not established as direct as in the relationship between similes. Rather, a Metaphor seems (or could) serve as a resource that instigates indirect reasoning through phrasal construction. Indeed,

> the most obvious semantic difference between simile and Metaphor is that all similes are true and most[13] Metaphors are false. The earth is like

[12] Indeed, § XXI of Poetics (where the aforementioned passage is inscribed), as well as two paragraphs that precede it and the one that follows it (§ XIX, on thought and modes of expression; § XX: on utterance and its parts; and § XXII, on poetic utterance and criticisms of it in Homeric poems), they seem to refer more to grammar than poetry itself (cf. Souza, in: Aristóteles, 1986, pp. 185-86).

[13] This is because a Metaphor could consist not only of an evidently false construction – e.g., in the utterance of the king's philosopher, in José Saramago's *Tale of the Unknown Island*: "*each man* is an island" (cf.

> a floor, the Assyrian did come down like a wolf on the fold, because everything is like everything. But turn these sentences into Metaphors, and you turn them false; the earth is like a floor, but it is not a floor; Tolstoy, grown up, was like an infant, but he wasn't one. We use a simile ordinarily only when we know the corresponding metaphor to be false. We say Mr S. is like a pig because we know he isn't one. If we had used a Metaphor and said he was a pig, this would not be because we changed our mind about the facts but because we chose to get the idea across a different way. (Davidson, 1984, p. 257)

Now, such possibility to access an information "across a different way" is what seems proper to the Metaphor. Indeed, although the Aristotle's definition seems to be quite broad – encompassing different figures of speech[14] – one could easily say that, in a metaphorical speech, expressions like "*so to speak*", "*as it were"*, "*as if"*, "*just as"*, and so on, are implied.

It's difficult, however, to establish how such a "*so-to-say* thinking" could lead us to know something as metaphorical reasoning does. And since, presumably, "no theory of metaphorical meaning or metaphorical truth [seems] can help explain how Metaphor works" (Davidson, 1984, p. 259), it could be quite fruitless to dwell on that. Much more profitable, in turn, it would be to emphasize that, as "a poetically or rhetorically ambitious" (Hills, 2017) language resource, a Metaphor seems to serve to illustrate something that, otherwise, would not be so visible. Nonetheless, such a "making-visible" is not quite straightforward; rather, what seems to concern Metaphor is that it "inspires or prompts the insight" (Davidson, 1984, p. 263), making "easier or more transparent in some respects" (Davidson, 1984, p. 264) something more complex. It's precisely this character that Kant reserves for some fundamental passages of his texts: to illustrate or make more evident what is – or would be – already sufficiently grounded on a theoretical level – albeit often in a very dry way.

Saramago [1997], 1999, p. 27) –, but also, although more rarely, in an obviously true sentence – e.g., in John Donne's meditation excerpt "*no man* is an island" (cf. Sparrow, 1923, p. 98).

[14] As a matter of fact, Aristotle claims that "Metaphor is the application of an alien name by transference either from genus to species, or from species to genus, or from species to species, or by analogy, that is, proportion". However, the transference "from genus to species, as: 'There lies my ship' [...] [, where] lying at anchor is a species of lying", could be called a *Metonymy*; the one "from species to genus, as: 'Verily ten thousand noble deeds hath Odysseus wrought' [...] [, where] ten thousand is a species of large number, and is here used for a large number generally", a *Hyperbole*; and, finally, the one "from species to species, as: 'With blade of bronze drew away the life,' and 'Cleft the water with the vessel of unyielding bronze' [...] [, where] ἀρύσαι, 'to draw away' is used for ταμεῖν, 'to cleave,' and ταμεῖν, again for ἀρύσαι – each being a species of taking away", could easily qualify as a *Euphemism* (*Poet*, 1457 b 9ff).

Already widespread, I presume, is the Metaphor used by Kant in the *Introduction* to his *Critique of Pure Reason*. If "*critique*", in its Greek origin – *κρίνω* – essentially means "*separate*", this is precisely what corresponds to Kant's endeavour in his first *Critique*: the separation between what is legitimate in the theoretical sphere and what is not – i.e., "what can I know?" (*KrV*, B 833). Therefore, dealing with the limits of our possible cognition, Kant claims that

> there is no doubt whatever that all our cognition begins with experience; for how else should the cognitive faculty be awakened into exercise if not through objects that stimulate our senses and in part themselves produce representations, in part bring the activity of our understanding into motion to compare these, to connect or separate them, and thus to work up the raw material of sensible impressions into a cognition of objects that is called experience? (*KrV*, B 1)

Indeed, Kant asserts in the *Introduction* to his *Transcendental Logic,*

> our cognition arises from two fundamental sources in the mind, the first of which is the reception of representations (the receptivity of impressions), the second the faculty for cognizing an object by means of these representations (spontaneity of concepts); through the former an object is *given* to us, through the latter it is *thought* in relation to that representation (as a mere determination of the mind). Intuition and concepts therefore constitute the elements of all our cognition, so that neither concepts without intuition corresponding to them in some way nor intuition without concepts can yield a cognition. [...] Neither of these properties is to be preferred to the other. Without Sensibility no object would be given to us, and without Understanding none would be thought. Thoughts without content are empty, intuitions without concepts are blind. [...] Further, these two faculties or capacities cannot exchange their functions. The Understanding is not capable of intuiting anything, and the senses are not capable of thinking anything. Only from their unification can cognition arise. (*KrV*, B 74-76)

Since human cognition has this "double dependence" (both on an activity of thought (Understanding) and the need for an intuition – i.e., a matter for this thought –, which is invariably obtained through Sensibility), Kant puts in check, e.g., the legitimacy of all (supposed) cognition based on pure thinking, which abandons Sensibility – i.e. every doctrine that claims to know the objects of Metaphysics.

Indeed, referring to this presumptuous cognition, "encouraged by such a proof of the power of Reason", Kant offers us the image of "[a] light *dove*, [which,] in free flight

cutting through the air the resistance of which it feels, could get the idea that it could do even better in airless space" (*KrV*, B 8-9). And, he follows, "likewise, Plato abandoned the world of the senses because it posed so many hindrances for the Understanding, and dared to go beyond it on the wings of the ideas, in the empty space of pure Understanding" (*KrV*, B 9). The metaphorical reasoning that Kant suggests seems to be none other than: *just as* the dove is unable to fly in a vacuum – considering that it's precisely the resistance of the air that allows its displacement in flight –, *so* cognition cannot progress in the absence of the world of the senses – since it is from there that we receive the matter of cognition!

So, Plato, according to Kant, "did not notice that he made no headway by his efforts, for he had no resistance [like the friction of the air, in the case of the dove], no support, as it were, by which he could stiffen himself, and to which he could apply his powers in order to put his Understanding into motion" (*KrV*, B 9). The figure of the dove aiming to take flight in a vacuum illustrates properly one of the conclusions of the first *Critique*: the impossibility of expanding our cognition *beyond* the world of the senses.

Also to illustrate what is legitimate for us to call "cognition", Kant presents, in the third and final chapter of his *Analytic of Principles,* one of his most fruitful Metaphors. Assuming that the *Transcendental Deduction* results are satisfactory – i.e., that the categories apply only to the phenomena and the latter, already in perception, are submitted to the categories (*KrV*, B161), since

> [if] the categories are *a priori* applicable to the objects of empirical intuition [, it is] because these objects, insofar as they are apprehended in space and time, are (intuitively) synthesized in such a way that they can be (discursively) reflected under empirical concept according to the logical forms of judgment (Longuenesse, 2000, p. 243),

and, furthermore, only what enters our Sensibility and is subsumed under the pure concepts of the Understanding could be legitimately called "cogniton", Kant reiterates this conception by stating that the Understanding, as "land of truth", is an island "enclosed in unalterable boundaries by nature itself" (*KrV*, B 294).

Such an island, he continues, would be

> surrounded by a broad and stormy ocean, the true seat of illusion, where many a fog bank and rapidly melting iceberg pretend to be new lands

> and, ceaselessly deceiving with empty hopes the voyager looking around for new discoveries, entwine [them][15] in adventures from which [they] can never escape and yet also never bring to an end. (*KrV*, B 294-295)

Now, Kant's "maritime" Metaphor takes up what had been exposed in the opening lines of the first edition of the first *Critique*: that

> human Reason has the peculiar fate in one genus of its cognition <*Gattung ihrer Erketnisse*> that it is plagued <*belästigt*> with questions which it cannot dismiss, since they are given to it as problems by the nature of Reason itself, but which it also cannot answer, since they transcend every faculty <*Vermögen*> of human Reason. (*KrV*, A VII)

Now, that to which Reason *can* provide answers – our legitimate cognition – refers to what affects our Sensibility, i.e., objects of sensible nature, "enclosed in unalterable boundaries". Everything remaining *beyond* these "sensible" boundaries of possible experience, obviously, is *super*sensible. Furthermore, if Understanding's concepts have their use have their use (for cognizing an object) restricted to sensible intuitions, everything beyond those boundaries of possible experience cannot be object-of-cognition.

It is in this sense that Kant tells us of a dense fog circumscribing the island that falsely gives us the idea of solid land. Indeed, we have no security to extend our cognition beyond that "land of truth"; still, due to its "peculiar fate" – i.e., due to a need <*Bedürfnis*> felt by virtue of its natural "*drive to cognition*" <*Erkenntnistrieb*> (*WDO*, 08: 139, footnote) –, Reason entices us to "the battlefield of these endless controversies [which] is called *Metaphysics*" (*KrV*, A VIII).

Considering this, Kant himself even warns us: "before we venture out on this sea, to search through all its breadth and become certain of whether there is anything to hope for in it, it will be useful first to cast yet another glance at the map of the land that we would now leave" (*KrV*, B 295). It means that *if* coherent thinking is to be sought about what goes beyond sensible nature, *then* this thinking must develop according to the same rules governing the cognition formation on objects of possible experience. Ultimately,

[15] Taking into account both male and female gendered persons, in the absence of a generic third-person pronoun in English and due to both stylistic strangeness of *s\he*, and the prolixity of *he or she*, I will use, whenever the noun's gender cannot be identified, the pronoun *they* with the plural verb, even when referring to a person in the singular – as well its derived forms: *them*, *their*, *theirs*, and the non-standard *themself* –, including in quotations.

this is nothing more than saying that we should think about the field of the incognizable in formal terms of possible cognition – hence Kant's mention of the "map of the land of truth".

This is undoubtedly one of the most prolific sections, not only for the first *Critique* but also for the closing of Kant's philosophical system of Reason as a whole. For it is where he lays the first cornerstones to ground our capacity to judge about what we cannot, effectively, cognize; an argument that would be taken up in the final paragraphs of the *Prolegomena*, in his essay *What does it mean to orient oneself in thinking?*, and, above all, in his *Critique of Judgment*.

It would be unnecessary (and rather burdensome), therefore, to mention here all the unfoldings of such a profuse passage. Instead, I limit myself to pointing out that, as in the aforementioned dove's Metaphor, the island's one does not seem to add anything we didn't already know to the problem of human cognition. Rather, Kant seeks, through such a particularly poetic language resource to highlight "what can I [surely] know?" (*KrV*, B 833). And he does so by illustrating that it's impossible for our cognition to extend safely beyond the limits of sensible nature itself, when we dare to dive into that "broad and stormy ocean, the true seat of illusion" which is the supersensible.

Nonetheless, it's fascinating how Kant's argument occasionally go beyond the mere analogical or metaphorical resource. In some cases, especially with regard to his practical Philosophy, he suggests that we mentally and narratively construct hypothetical situations – i.e., TEs – in order to prove, in different conjectures, the validity of what is *a priori* grounded.

4.

There is no consensus about a rigid definition of what TEs are,[16] nor about their true function as a "substitute for real experiment, or as a reliable device for discerning possibilities" (Blackburn, 2005, p. 365).

From what we might call an empiricist perspective, TE are nothing more than arguments surrounded by fanciful elements and "can be reconstructed as arguments based on tacit or explicit assumptions [, so that] the belief in the outcome-conclusion the TE is

[16] Brown draws attention to the fact that "people often use the expression 'TEs', but hardly anyone has thought seriously (or at least written extensively) about them" (Brown, 1991, p. x). And this could be justified considering the great heterogeneity of TEs found both in the scientific and philosophical fields.

justified insofar as the reconstructed argument can justify the conclusion" (Norton, 1996, p. 339). Therefore, the empiricist approach argues that, by eliminating any fantasy element – which, strictly speaking, is "not to be found in the conclusion" extracted from a TE –, we arrive at the same conclusions that we would reach only through premises: all the narrative elements of TEs, then, "are always eliminable without compromising our ability to arrive at the conclusion, although elimination may make the argument for the conclusion considerably more complicated" (Norton, 1996, p. 336).

Such an empiricist approach does not see, therefore, any "epistemic miracle" achieved through TEs: "they cannot do more than can ordinary thinking with its standard tools of assumption and argument [, and] they open no new channels of access to the physical world" (Norton, 1996, p. 336), serving rather as enablers in presenting such arguments. It does not follow, however, that, by providing no "new knowledge" – since, as any self-respecting empiricist theory would argue, it can only be drawn from experience, which in turn cannot be replaced by fantastic elements of narrative language –, TEs should be rejected; rather, the empiricist perspective maintain that an argument by TE is only valid insofar as its narrative structure can be translated into valid (i.e., not fallacious) arguments.

If, on the one hand, such empiricism recognizes the guiding role of TEs, especially when scientists "found the vehicle of the TE as an especially comfortable medium in which to apply [their] verificationism" (Norton, 1996, p. 365), on the other hand, the criticism that this approach makes to TEs corresponds to the application of the "Ockham's razor" – i.e., the adoption of the shortest path – in favour of an argument reduced to premises. For,

> [whether] the success of reconstruction thesis shows that the argument view is viable [...] [in a way that] argumentation is uncontroversially a part of many TEs, any reasonable account of TEs must give some place to it. Given that the argument is also sufficient to provide a complete account of the Epistemology of TEs, why should we consider calling in any further epistemic machinery? Ockham's razor instructs us not to. The magician materializes a dove. We know he can do this by ordinary trickery. Why would we assume that real magic was used? (Norton, 1996, pp. 357-58)

Now, such empiricist perspective, in turn, is under a strong objection: that arguments cannot explain what happens epistemologically in TEs: for to construct a

logical argument does not seem to be the same as to construct a TE, since this "reductionism" seems to miss something that is quite particular to the latter. This criticism – which we could call a rationalist, idealistic one – claims then that the reasoning involved in TEs cannot be entirely derived from logical arguments; rather, reasoning by TE would open a kind of "special window" (Norton, 1991, p. 129), through which we would have access to knowledge.

One could agree that no rationalist approach that claims to be taken seriously would completely forgo reasoning through arguments – even whether it were somehow possible. Rather, in such a rationalist conception the TE itself acts as an argument – either when it "destroys or at least presents serious problems for a theory, usually by pointing out a shortcoming in its general framework" (Brown, 1991, p. 34), or establishes the positive results thereof (Brown, 1991, p. 36).[17]

The divergence between an empiricist and a rationalist approach appears, it seems, when the TE intends to do more than provide answers to "'what if' questions" (Cooper, 2005, p. 336). In such cases, the rationalist perspective attributes to TEs the ability to enable a sort of *a priori* "inductive leap" (Brown, 1991, p. 45) that, although "it is not based on new empirical evidence nor is it merely logically derived from old data", it leads (or seems to lead) us to a theory "better than the predecessor theory" (Brown, 1991, p. 76). Nonetheless, even a rationalist approach of TEs could recognize that the latter class is among "the most controversial" (Brown, 1991, p. 45), objectionable ones.

A common point, however, between both approaches on such a curious way of arguing is that "TEs are basically devices of the imagination" (Brown & Fehige, 2019) which, through a narrative structure, allow or suggest a rational (non-subjective!) analysis of an issue. Unlike Analogy (which compares objects to each other) and Metaphor (which appropriates this analogue comparison through language), a TE does not seems to be a comparison in the sense of equivalence between objects. Rather, it seems to be the case of the analysis of the nature of the object-of-investigation in a narrative way, that uses

[17] In this sense, the TE could serve both to "facilitat[e] a conclusion drawn from a specific, well-articulated theory" (Brown, 1991, p. 36) – and this seems to be the case, e.g., with Descartes' argument about considering a Deceiving God or Evil Genius in order to question the reliability of our thoughts (but not our thinking!) (cf. *Med*, AT VII: 21 ff]) –, to formulate a hypothetical theory in order to explain an imagined phenomenon (cf. Brown, 1996, p. 40) – a resource employed, e.g., in moral dilemmas, as the so-called *Trolley Problem* (cf. Foot, 1967, p. 8) –, or even for, starting "with some vague general principle" (Brown, 1996, p. 42), to propose a well-formulated theory – like, one might assume, the Putnam's Twin Earth (cf. Putnam, 1973, pp. 700 ff).

elements, so to speak, "*touchable*",[18] which are created by those who perform the TE itself.[19]

Thus defined, one could seamlessly find some "TEs invitations" also in Kant's writings. Nevertheless, whereas Metaphors and Analogies were more prominent in their theoretical philosophy, it is in Kantian practical Philosophy, one might argue, where we could find such narrative resources.[20]

After maintaining that among all commands "the categorical imperative alone can be stated as a practical *law*" (*GMS*, 4: 420), Kant claims, in the *Second Section* of his *Groundwork of Metaphysics of Morals,* that such an imperative "is thus only a single one, and specifically this: *Act only in accordance with that maxim through which you can at the same time will that it become a universal law*". And, follows, if it is "from this one imperative [that] all imperatives of duty can be derived as from their principle" (*GMS*, 4: 421), through the sieve of the categorical Imperative it is possible to distinguish between perfect and imperfect duties. Such a distinction, however, Kant reserves for his *Metaphysics of Morals* (in particular, his *Doctrine of Virtue*) (cf. *MSTL*, 6: 415 ff). For now, he deals with presenting four cases in which, testing the maxims of the agents, he aims to verify the moral value of their actions.

In the first case, we are led to imagine a person who "through a series of evils that have accumulated to the point of hopelessness, feels weary of life but is still so far in possession of [their] Reason that [they can ask themself] whether it might be contrary to the duty to [themself] to take [their] own life" (*GMS*, 4: 421-22). In order to justify suicide, the agent's maxim would be: "from self-love, I make it my principle to shorten my life when by longer term it threatens more ill than it promises agreeableness" (*GMS*, 4: 422).

The second situation is elaborated in terms of a person who "[see themself] pressured by distress into borrowing money [However,] [they know] very well that [they] will not be able to pay, but [they also see] that nothing will be lent [them] if [they do] not

[18] I.e., appealing to the sensible side of the subject, building, most of the time, a scenario in which this subject sees a hypothetical situation.
[19] If, on the one hand, such a definition seems too broad to encompass a range of reasoning and fictions in general, on the other hand, it's precisely this broadness that favours the heterogeneity of TEs – and, in this perspective, the more wide the definition, the more satisfactory it appears to be.
[20] Buzzoni, in turn, argues that one recognisably finds what might be called TEs in Kant's philosophy of nature. For his analysis, cf. Buzonni, 2011, pp. 99 ff.

firmly promise to pay at a determinate time". It would be then a case of false promise, whose maxim is: "if I believe myself to be in pecuniary distress, then I will borrow money and promise to pay it back, although I know this will never happen" (*GMS*, 4: 422).

The third case describes someone who find in themself "a talent, which could, by means of some cultivation, make [them] into a human being who is useful for all sorts of aims". However, Kant follows, they see themself "as in comfortable circumstances and sooner [prefer] to indulge in gratification than to trouble [themself] with the expansion and improvement of [their] fortunate natural predispositions". In this case, it would be a "maxim of neglecting [their] gifts of nature with [their] propensity to amusement" (*GMS*, 4: 423).

Finally, Kant suggests that we consider a person "for whom it is going well, while [they see] that others have to struggle with great hardships (with which [they] could well help them)". However, indifferent to the scourges of a third party, such a person thinks: "what has it to do with me? Let each be as happy as heaven wills, or as [they] can make [themself], I will not take anything from [them] or even envy [them]; only I do not want to contribute to [their] welfare or to [their] assistance in distress!" (*GMS*, 4: 423).

Immediately after presenting each of these situations, Kant compares their maxims with the "supreme principle of morality" (*GMS*, 4: 392) moulded in the categorical imperative formula. As for a maxim justifying suicide in view of self-love difficulties, he claims: it "could not possibly obtain as a universal law of nature, and consequently it entirely contradicts the supreme principle of all duty"; as an alleged principle of self-love "whose law it was to destroy life through the same feeling whose vocation it is to impel the furtherance of life", it "would contradict itself". Nor could a false-promise maxim claim universalisation, since its universality "would make impossible the promise [itself] and the end one might have in making it, since no one would believe that anything has been promised [the,], but rather would laugh about every such utterance as vain pretence" (*GMS*, 4: 422).

However, regarding the "hedonistic behaviour" maxim, Kant refutes it, in turn, with some reservation. He claims:

> although a nature could still subsist in accordance with such a universal law [i.e., a law of neglect of natural gifts and pursuit of amusement], though then the human being […] would think only of letting [their] talents rust and applying his life merely to idleness, amusement,

> procreation, in a word, to enjoyment; yet it is impossible for [them] to *will* that this should become a universal law of nature, or that it should be implanted in us as such by natural instinct. For as a rational being [they] necessarily [will] that all the faculties in [them] should be developed, because they are serviceable and given to [them] for all kinds of possible aims. (*GMS*, 4: 423)[21]

Finally, concerning the maxim of indifference towards the difficulties of third parties, Kant replies that, in view of such universal law, "the human race could well subsist". However, he follows,

> although it is possible that a universal law of nature could well subsist in accordance with that maxim, yet it is impossible to will that such a principle should be valid without exception as a natural law. For a will that resolved on this would conflict with itself, since the case could sometimes arise in which [they need] the love and sympathetic participation of others, and where, through such a natural law arising from [their] own will, [they] would rob [themself] of all the hope of assistance that [they wish for themself]. (*GMS*, 4: 423)

Now, Kant's intention with each of the aforementioned situations – it seems clear – is to collate their maxims with the categorical imperative: if what gives moral value to action is "never to conduct myself except so *that I could also will that my maxim become a universal law*" (*GMS*, 4: 402), such collating aims to deny the possibility of universalizing these maxims. Thus, on the one hand, Kant denies their genuine moral value – since these episodes have as their background the sensible and selfish inclination of self-love and their maxims have no such value (at least, not quintessentially, so to speak) –, and, on the other, he highlights what, indeed, *does have* such value in his Ethics. For,

> [once] robbed the will of every impulse that could have arisen from the obedience to any law [of sensible nature], there is nothing left over except the universal lawfulness of the action in general [i.e., a universal principle for actions of every rational being] which alone is to serve the will as its principle. (*GMS*, 4: 402)

[21] It's remarkable, one might argue, that Kant's answer sounds somewhat artificial - betting, above all, on a teleology of the subject's own faculties. A similar teleological argument is also employed, e.g, to justify the purpose of Reason in human beings (cf. *GMS*, 4: 395-6). This, it could also be argued, takes place on the horizon of his own critical-transcendental system, which has the rational human being at its centre: hence the teleological argument about the subject's faculties; hence the need for them to develop their talents. I admit this, within the critical Kantian project, as well as I recognize that an in-depth analysis of this subject would exceed the limits of this paper – which is why I pass abeam thereof.

More complex, however, than any of the aforementioned "examples" (*GMS*, 4: 429) of the *Groundwork*, seems to be the hypothetical situation elaborated by Kant, which resulted in the debate with Benjamin Constant expressed in the 1797 essay *On a supposed right to lie from philanthropy*.

In response to Constant's position, according to which

> the moral principle: it is a duty to tell the truth, would make any society impossible if it were taken singly and unconditionally [...] [, since] to tell the truth, is thus a duty; but it is a duty only towards to one who has a right to the truth [, and] no one has, however, a right to a truth that harms others (*VRML*, 8: 425),

Kant claims that "firstly it must be noted that the expression: to have a right to truth, is meaningless <e*in Wort ohne Sinn*>" (*VRML*, 8: 426). Therefore, given the absurdity of the expression, it's not legitimate to ask whether there is a right to the truth, rather

> whether a human being, in cases where [they] cannot avoid answering yes or no, has the warrant <*Befugniß*> (right) to be untruthful [...] [, and] whether [they are] not actually bound to be untruthful in a certain statement which [they are] unjustly compelled to make in order to prevent a threatening misdeed against [themself] or someone else. (*VRML*, 8: 426)

Now, Kant does not make explicit here the passage that motivated Constant's answer.[22] However, one can read in his *Metaphysics of Morals* the following casuistical pondering:

> Can an untruth from mere politeness (e.g., the 'your obedient servant' at the end of a letter) be considered a lie? No one is deceived by it. An author asks one of [their] readers, 'How do you like my work?' One could merely seem to give an answer, by joking about the impropriety of such a question. But who has [their] wit always ready? The author will take the slightest hesitation in answering as an insult. May one, then, say what is expected of one? If I say something untrue in more serious matters, having to do with what is mine or yours, must I answer for all the consequences it might have? For example, a householder has ordered his servant to say 'not at home' if a certain man asks for [them]. The servant does this and, as a result, the master slips away and commits

[22] For, although he grants that he said it "somewhere or other", he himself "cannot recall [...] where" (*VRML*, 8: 425, footnote).

> a serious crime, which would otherwise have been prevented by the guard sent to arrest [them]. Who (in accordance with ethical principles) is guilty in this case? Surely the servant, too, who violated a duty to [themself] by [their] lie, the results of which [their] own conscience imputes to [them]. (*MSTL*, 6: 431)

It's possible that the passage about which Kant and Constant quarrel is not exactly this one.[23] Nonetheless, the core problem seems to be the same: if one would have the right to lie in order to avoid harming others or not. And, regarding that, Constant is emphatic: we would be entitled to lie since "no one has [...] a right to a truth that harms others" (*VRML*, 8: 425). He considers – quoted by Kant – that

> it is a duty to tell the truth [, however,] the notion of duty is inseparable from the notion of right [, so that, whether] a duty is what in one being corresponds to the right of another [,] where there are no rights there are no duties. To tell the truth then is a duty, but only towards him who has a right to the truth. (*VRML*, 8: 425)

So, according to Constant, there would be no problem in lying to anyone who harms others. For, in any case, whoever tries to harm someone is, for this very reason, deprived of a right to any information that makes it possible to harm that person. Curiously, however, Kant maintains that, *even in such a situation, there would be no right to lie* – even if it was done to protect (i.e., out of love for) humanity. For, he insists, "it is [...] a sacred, unconditional command of Reason <*Vernunftgebot*>, limited by no expediency <*durch keine Convenienzen einzuschränkendes*>: in all declarations to be truthful (honest)" (*VRML*, 8: 427).

Now, to defend truthful statements, even when they may put someone's life at risk, sounds rather unusual for a philosopher who argues that "humanity, as much in [my] own person as in the person of every other, [must be seen] always at the same time as end [i.e., as something that has absolute and unconditional value, i.e., in itself] and never merely as means [i.e., as that which has conditional value]" (*GMS*, 4: 429). Nevertheless, Kant takes into account the principle: "truthfulness <*Wahrhaftigkeit*> in statements which cannot be avoided is the formal duty of a human being to everyone, however great the disadvantage that may arise from it to [them] or another" (*VRML*, 8: 426).

[23] Although both Kant's texts and Constant's essay date from the same year.

Firstly, it's noteworthy that this is not a matter of duty of Ethics, but of *Right* <*Rechtspflicht*>. In this regard, Kant clarifies: "the doctrine of virtue [i.e., Ethics] sees in that transgression [i.e., in the lie] only *worthlessness* <*Nichtswürdigkeit*>, whose reproach the liar draws on [themself]" (*VRML*, 8: 426, footnote); Right, in turn, would involve the contractual dimension of "a free community of competent subjects" (Höffe, 1992, p. 171). And it's due to such contractual competence that

> all Right consists solely in the restriction of the freedom of other, with the qualification that their freedom can co-exist with my freedom within the terms of a general law, and public Right (in a commonwealth) is simply a state of affairs <*Zustand einer Wirklichen*> regulated by a real legislation which conforms to this principle and is backed up by power, and under which a whole people live as subjects <*Untertanen*> in a lawful state (*status iuridicus*) in general, namely of the equality of the effect and counteraction of a mutually restraining choice <*einander einschränkenden Willkür*> in accordance with the general law of freedom (which is called the civil state <*der bürgerliche Zustand*>): thus the *innate Right* <*das angeborne Recht*> of everyone in such a state (i.e., before all legal acts of the same) with regard to the authority to compel everyone else <*in Ansehung der Befugniß jeden andern zu zwingen*> so that [they always remain] within the limits of the agreement of the use of [their] freedom with mine, is consistently *equal*. (*VATP*, 8: 292-93)

Based on such a notion of *Right* Kant states that

> although to someone who unjustly compels me to a statement, I am not doing wrong <*nicht Unrecht tue*> if I falsify it, nevertheless by this falsification, which (though not in the jurists' sense) may be called a lie, I am doing wrong in the most essential part of duty *in general* <*im wesentlichsten Stücke der Pflicht überhaupt*>: i.e., I do, so far as in me lies, that statements (declarations) find no faith at all, and consequently all rights based on contracts cease to exist and lose their power; which is a wrong inflicted on humankind in general. (*VRML*, 8: 426)

Thus, in Kant's exposition, even if the most allegedly unobjectionable lie is (apparently) harmless, "it always harms another; whether not some other particular human being, still it harms humankind in general, since it makes the source of Right unserviceable <*indem sie die Rechtsquelle unbrauchbar macht*>" (*VRML*, 8: 426). Surely, one could ask, however, about the cases in which the lie is used in order to save someone from some evil (like the episode mentioned above): is it also here, a false testimony, harmful at any rate? Now, it's anything but surprising that Kant's answer is categorical:

> This benevolent lie *may*, however, by *accident* (*casus*) become punishable even by civil laws <*nach bürgerlichen Gesetzen*>; and that which escapes punishability <*Straffälligkeit*> only by accident may also be condemned as a wrong by external laws <*nach äußeren Gesetzen*>. For instance, whether you have prevented in the act one who is just now dealing with murderous craving <*Mordsucht*> *through a lie*, you are for all the consequences, which might arise from it, responsible in a legal way. Whether you have, however, strictly remained with the truth, then public justice cannot harm you <k*ann dir die öffentliche Gerechtigkeit nichts anhaben*>; be the unforeseen consequence what it may. (*VRML*, 8: 426-427)

Returning, then, to the situation that sparked Constant’s response, of lying to someone who might harm others, Kant’s position could not be different:

> It is possible that after you have honestly answered yes to the murderer’s question, whether [their] intended victim is in the house, the latter may have gone out unobserved, and so not have run into the murderer’s spear <*dem Mörder nicht in den Wurf kommen*>, so that the deed, therefore, would not be done; whether you lied, however, and said [they were] not in the house, and [they] had really (though unknowingly to you) gone out, so that the murderer met [them] leaving, and committed [their] deed on [them]: then you can rightly be accused of being the causer <*Urheber*> of the death thereof. For if you had told the truth, as well as you knew it: the murderer, perhaps, while seeking for [their] enemy in the house, would have been seized by neighbours who were approaching <*herbeigelaufenen Nachbarn*>, and the deed would have been prevented. (*VRML*, 8: 427)

Now, despite an apparent *naïveté* of Kant’s argument, what should be taken from this is: more important than the outcome of our actions (either for us or for others), is what *ground* them. That is what we must ultimately respond to, for the outcome of our actions *is not* within our reach; what *is*, instead, is the principle we adopt to act. And even if, by this principle, “with [...] greatest effort nothing [...] were accomplished, [...] it would shine like a jewel for itself, as something that has its full worth in itself” (*GMS*, 4: 394).

5.

From the fact that Kant considers “examples” harmful to understanding a theory, since “they only seldom adequately fulfil the condition of the rule (*as casus in terminis*)” (*KrV*, B 173), it does not follow – as, hopefully, it became clear in the course of the present

paper – that he himself does not employs such exemplifying argumentative figures. And this – what sounds quite curious – not so "seldom" as one might think.

Now, the clarity of an argument, Kant considers, may be both discursive and intuitive. The former, logical, is achieved "*through concepts*"; the latter, aesthetic, "through *intuitions*, that is, through examples or other illustrations *in concreto*" (*KrV*, A XVII). With regard to the discursive clarity of his argument, Kant flatters himself: he has "taken sufficient care for [it]", since "that was essential to [his] undertaking". And in the fulfilment of such an undertaking, on the one hand, Kant recognizes the "dry, merely scholastic manner" of his means (*KrV*, A XVIII); on the other hand, however, he believes that such *straightforwardness* of his argument favours "a survey of the whole", as well "the articulation or structure of the system, which yet matters most when it comes to judging its unity and soundness" (*KrV*, A XIX): it's this systematic articulation that Kant prizes so much; it's this, presumably, the essence of his "undertaking".

Nonetheless, as Kant himself is keen to remind us, the reader has a right to demand, in addition to a logical clarity, an aesthetic one – a visual one: a "less strict", but "still fair" clarity. And even though he claims he cannot satisfy such a demand – since the logical arguments "would suffice to fill an extensive work", so that it would seem "inadvisable to swell it further with examples and illustrations, which are necessary only for a *popular* aim" (*KrV*, A XVIII) –, it still seems quite obvious that he *is* concerned with providing "examples and illustrations", making more intuitive – and, why not, more "popular" – what he had already exposed in a discursive way. But why, after all? Why does Kant seem – to use his own words – to paint his logical argument with such "bright colours"? Would it be to make it more "palatable" to the large mass, or wouldn't Kant trust the logical-discursive clarity of his own text enough?

On the first point, the "colourful clothes" of his argument cannot (or at least does not seem to) aim to make it more popular. This is because Kant himself is quite aware that his work (or, at least, a considerable part of it) "could never be made suitable for popular use, and real experts in this science do not have so much need for things to be made easy for them; although this would always be agreeable, […] it could also have brought with it something counterproductive" (*KrV*, A XVIII). Regarding the second point, in turn, i.e., about an eventual distrust of Kant in relation to the discursive, merely logical clarity (and sufficiency) of his own argument, I would be quite doubtful that this

is the case. And this taking into account the weight that Kant himself gives to "logic" in his work.

As we know, it is the "Logic", in general, that "contains the absolutely necessary rules of thinking, without which no use of the understanding takes place" (*KrV*, B 76); pure and *a priori*, the Logic "is a *canon of the Understandin*g and Reason" (*KrV*, B 77). However, if, on the one hand, it's true that logical, merely formal arguments could be sufficient from a rational point of view, it's also true, on the other hand, Kant is aware that we, as rational beings, are rational *human* beings, and that our "human" aspect cannot be separated from our Sensibility, i.e., from our sensible, intuitive side, which brings its limitations and demands.

A similar concern with our peculiar rational human nature could be seen in the *Preface* to *Religion*. After reaffirming that moral laws

> command absolutely, no matter what their result may be; indeed, they even compel us to abstract entirely from the result when a particular action is at issue; and through this they make duty an object of the greatest respect, without putting before us and assigning to us a purpose (as well as a final purpose) that would have to amount, say, to the commendation of these laws and to the incentive for fulfilling our duty (*RGV*, 6: 6, footnote),

Kant ponders that such laws could "be sufficient for all human beings if only they abided (as they should) by the precept of pure reason in the law". Nonetheless, right after that, he calls into question: "why do they need to know the outcome of their moral doing and refraining, which the course of the world will bring about"? For, he follows, "it suffices that they do their duty, even if with [the end of] earthly life everything were to be over and even in it happiness and worthiness were perhaps never to coincide" (*RGV*, 6: 6, footnote). "It is one of the *unavoidable limitations of the human being*", Kant concludes, and of their

> (perhaps also of all other world beings') practical power of Reason to look out, in all actions, to the result issuing from them, in order to discover in it something that could serve as a purpose for [them and could also prove the purity of the intention – which result, though last in performance (*nexu effectivo*), is first in presentation and intention (*nexu final*). (*RGV*, 6: 6, footnote)

Thus, even if, on the one hand, Kant claims that an "*a priori* thought" is sufficient for something to be "unconditionally commanded as a law without borrowing anything from experience" (*KpV*, 5: 31), on the other hand, he seems aware that we, human beings, do not operate at a merely formal level, but also at a *sensible* one. And that at this sensible level, an intuitive clarity – a clarity, which, by the way, we have the right to demand, and such a right would be (at least, it seems to be) guaranteed by our own sensible nature – could also serve us very well. But in what sense could it serve us?

For sure, not to ground an argument. This is because, for such a valid groundwork, nothing "intuitive" would be necessary. If one takes into account, for instance, the aforementioned hypothetical situations of the *Groundwork* and the 1797 essay, one could see that, by using narrative structures, Kant does not intend to suggest something "new" or complement his Ethics or Doctrine of Right – inserting, as it were, subparagraphs for particular cases. Rather, those cases serve to demonstrate the validity of what is grounded for both Morality and Right. Through such argumentative resources, Kant does not conjecture what we should do, if a given situation were the case. Instead, he shows that *even if* a given situation were the case (i.e., in *any* case), it would absolutely not harm the validity of his argument.

Now, if we were to intend to classify such passages as TEs, we would have to situate them at an intersection between the empiricist and rationalist approaches: for we would have to rely both on the definition that EPs "are always eliminable without compromising our ability to arrive at the conclusion" (Norton, 1996, p. 336) – since awareness of the Categorical Imperative and the general principle of Right consist of would be enough to know how to proceed in any situation – and on the definition that an EP "facilitates [or, at least, might make more evident] a conclusion drawn from a specific, well-articulated theory" (Brown, 1991, p. 36).

This common point,24 concerns those TEs that work as cases and countercases for an already well-grounded theory, not substituting what supports the theory itself, since

[24] Such a compromise between Norton and Brown's conceptions for TEs in Kant also seems to have been sought by Buzzoni (cf. Buzzoni, 2018, p. 333), although his conception of TE as an operational-methodological thinking, as something that one experiences merely through reasoning, seems to lead him to consider that "the whole [*Critique of Pure Reason*] can be seen as an experiment of pure reason, a philosophical TE whose truth is guaranteed by the fact that human understanding loses itself in antinomies when it ventures beyond the limits of possible experience and attempts to deal with things in themselves" (Buzzoni, 2018, p. 329). Instead, I limited myself and took as a criterion the narrative (and not merely mental) character of the TEs – otherwise, one might argue, the philosophical practice itself could be framed

it is the arguments of the theory, i.e., how it is constructed and what solidifies it (and not its fictional, narrative, colourful wrapper), that lead to the conclusions. This without any major problems applies to the here discussed excerpts of Kant: for if the narrative structure could even facilitate access to the conclusion or validity of his theory, the former could *never replace* the latter.

However, if, on the one hand, Kant considers that for a valid argument – here understood as a necessary and strictly universal one – just a logical, formal, *a priori* reasoning would be required (cf. *KrV*, B 3-4), on the other hand, he seems to be quite aware of the possible *harshness* of a text that rests solely on such arguments. Therefore, one might argue, it is above all due to Kant's awareness of a certain *dryness* or, at least, of the subtlety and high degree of complexity of *his own texts* that, although he claims that

> *many a book would have been much clearer if it had not been made quite so clear* [,] for the aids to clarity help in the *parts* but often confuse in the *whole*, since the reader cannot quickly enough attain a survey of the whole; and all their bright colours paint over and make unrecognizable the articulation or structure of the system, which yet matters most when it comes to judging its unity and soundness (*KrV*, A XIX),

he himself – and not a few times! – employs such "bright colours" in his texts: not – I stress out – to replace a theory properly structured and grounded in an argument, in itself, coherent; rather, to make this same theory more evident (at least, more tactile *for us*, as not only intelligible *but also* sensible beings) even in threshold situations.

References

Abbagnano, N. (1999). *Dicionário de filosofia.* Martins Fontes.

Aristóteles. (1986). *Poética* (E. d. Sousa, Translation, preface, introduction, comments and appendices). Imprensa Nacional – Casa da Moeda.

Bartha, P. (2019). Analogy and analogical reasoning. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Spring Edition. https://plato.stanford.edu/archives/spr2019/entries/reasoning-analogy/

Blackburn, S. (2005). *The Oxford dictionary of philosophy*. Oxford University Press.

under a broad (whether not the broadest) concept of TE. And this, I presume, would add (unnecessary) difficulties to the discussion here conducted.

Brown, J. R. (1991). *The laboratory of mind: thought experiments in natural science*. Routledge.

Brown, J. R., & Fehige, Y. (2019). Thought experiments. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Winter Edition. https://plato.stanford.edu/archives/win2019/entries/thought-experiment/

Buzzoni, M. (2018). Kantian accounts of thought experiments. In M. T. Stuart, Y. Fehige, & J. R. Brown (Eds.), *The Routledge companion to thought experiments* (pp. 327–341). Routledge.

Buzzoni, M. (2011). Kant und das Gedankenexperiment: Über eine kantische Theorie der Gedankenexperimente in den Naturwissenschaften und in der Philosophie. *Deutsche Zeitschrift für Philosophie, 59*(1), 93–107

Cooper, J. M. (Ed.). (1997). *Plato*. Complete works. Hackett.

Cooper, R. (2005). Thought experiments. *Metaphilosophy*, *36*(3), 328–347.

Davidson, D. (1984). What metaphors mean. In D. Davidson, *Inquiries into truth and interpretation* (pp. 245–264). Oxford University Press.

Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, *322*(10), 891–921.

Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford review, 5*, 5–15.

Hamm, C. (2013). "Conhecimento por analogia": sobre a função da argumentação indireta em Kant. In A. Serrão, C. Simões, E. M. de Souza, F. Afonso, M. L. R. Ferreira, P. Calafate, & U. R. de A. Marques (Eds.), *Poética da razão: homenagem a Leonel Ribeiro dos Santos* (pp. 71–80). Centro de Filosofia da Universidade de Lisboa.

Galilei, G. (1954). *Dialogues concerning two new sciences* (H. Crew and A. de Salvio, Trans.). Dover Publications.

Hills, D. (2017). Metaphor. In Edward N. Zalta (Ed.). *The Stanford encyclopedia of philosophy*. Fall Edition. https://plato.stanford.edu/archives/fall2017/entries/metaphor/

Höffe, O. (1992). *Immanuel Kant* (M. Farrier, Trans.). State University of New York Press.

Locke, J. (1854). *The philosophical works of John Locke* (2 vols.). Preliminary essay and notes by James Augustus St. John. Henry G. Bohn.

Longuenesse, B. (2000). *Kant and the capacity to judge: sensibility and discursivity in the transcendental analytic of the Critique of Pure Reason*. Princeton University Press.

Newton, I. (1999). *The Principia: mathematical principles of mathematical philosophy* (I. B. Cohen, A. Whitman and J. Budenz, Trans.). University of California Press.

Norton, J. D. (1991). Thought experiments in Einstein's work. In T. Horowitz & Gerald J. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 129–148). Rowman and Littlefield.

Norton, J. D. (1996). Are thought experiments just what you thought? *Canadian journal of philosophy*, *26*(3), 333–366.

Putnam, H. (1973). Meaning and reference. *Journal of philosophy, 70*(19), 699–711.

Rahman, F. (Ed.). (1959). *Avicenna’s De anima*. Being the psychological part of *Kitāb al-Najāt*. Oxford University Press.

Rutherford, E. (1911). The scattering of α and β particles by matter and the structure of the atom. *Philosophical Magazine, 6*(21), 669–688.

Saramago, J. (1999). *The tale of the unknown island* (M. J. Costa, Trans.). The Harvill Press.

Stuart, M. T., Fehige, Y., & Brown, J. R. (2018). Thought experiments: state of the art. In M. T. Stuart, Y. Fehige, & J. R. Brown, (Eds.), *The Routledge companion to thought experiments* (pp. 1–28). Routledge.

Sparrow, J. (Ed). (1923). *Devotions upon emergent occasions by John Donne*. Bibliographical note by Geoffrey Keynes. Cambridge University Press.

van den Berg, H. (2018). Kant and the scope of analogy in the life sciences. *Studies in History and Philosophy of Science, 71,* 67–76.

Received: 4 August 2021
Revised: 19 February 2022
Approved: 20 February 2022

This is an open-access article distributed under the terms of the Creative Commons Attribution License.