Toward a Kantian theory of prudential irrationality: between intellectual error and volitional failure

Rumo a uma teoria kantiana da irracionalidade prudencial: entre o erro intelectual e a falha volitiva

Tales Yamamoto[1]
Universidade Federal de Santa Catarina (UFSC)/Universität Vechta
talesyamamoto@hotmail.com

Abstract: This article investigates the conditions for a Kantian theory of prudential irrationality. Against Merle (2023), it argues that intellectual errors, such as incorrect beliefs, do not suffice to generate irrational actions. The analysis focuses instead on whether volitional failures can lead to prudentially irrational actions. To examine this, two interpretive models are considered: the negative model, inspired by Timmerman (2022), which denies prudential irrationality, and the positive model, developed by Korsgaard (2008), which affirms it. While Timmerman strips the Hypothetical Imperative of normativity, Korsgaard subordinates it to the Categorical Imperative. Both models face limits: the first excludes the possibility of instrumental irrationality; the second risks expanding the moral domain or weakening the link between freedom and the moral law. The challenge remains to explain how instrumental rationality is possible without reducing it to morality, while preserving its intrinsic tie to freedom.

Keywords: categorical imperative; hypothetical imperative; instrumental reason; Kant; normative ethics; rational agency.

Resumo: Este artigo investiga as condições para uma teoria kantiana da irracionalidade prudencial. Contra Merle (2023), defende-se que erros intelectivos, como crenças incorretas, não são suficientes para gerar ações irracionais. A análise concentra-se, em vez disso, na possibilidade de falhas volitivas conduzirem a ações irracionais prudenciais. Para examinar essa hipótese, consideram-se dois modelos interpretativos: o modelo negativo, inspirado em Timmerman (2022), que nega a irracionalidade prudencial, e o modelo positivo, desenvolvido a partir de Korsgaard (2008), que a afirma. Enquanto Timmerman esvazia o Imperativo Hipotético de normatividade, Korsgaard o subordina ao Imperativo Categórico. Ambos os modelos enfrentam limites: o primeiro exclui a possibilidade de irracionalidade instrumental; o segundo arrisca ampliar excessivamente o domínio moral ou enfraquecer o vínculo entre liberdade e lei moral. O desafio permanece em explicar como a racionalidade instrumental é possível sem reduzi-la à moralidade, preservando ao mesmo tempo sua ligação intrínseca com a liberdade.

Palavras-chave: imperativo categórico; imperativo hipotético; razão instrumental; Kant; ética normativa; agência racional.

[1] I had the privilege of discussing earlier versions of this paper with many colleagues. A thank you to Frank Rettweiler, Gehad Marcon Bark, Luciana Martinez, Marina Guimarães Back, Nicole Martinazzo and Sulamith Weber. Very substantial contributions were offered by some colleagues and professors. A very special thank you to Cristina Foroni Consani, Eduardo de Oliveira da Costa, Eduardo Estevão Quirino, Jean-Christophe Merle and Joel Thiago Klein. I would also like to thank the two anonymous reviewers for their contributions. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and is part of the PROBRAL Project "Kantian Perspectives on Social Irrationality" (CAPES-DAAD), process number 88887.078809/2024-00.

Recebido em 31 de julho de 2025. Aceito em 23 de novembro de 2025.

1. Introduction

If we want to ask *how* an irrational action is possible, firstly we must ask *whether* irrational actions are possible. This skepticism is valid insofar as Kant refers multiple times in the *Groundwork* (therefore GMS) and in the *Critique of Practical Reason* (therefore KpV)[2] to an action, but not to an irrational action. More specifically, in this case, we want to ask whether *prudential* irrational actions are possible. The question about prudential irrational actions is, I believe, a bit more tricky than the question about moral irrational actions. If we want to take Kant's philosophy and create a theory about moral irrationality, I see that as an easy path to take. In the case of prudential irrationality, I do not think that it would be that simple.

My aim in this paper is modest: to lay out some conditions for a possible Kantian model of prudential irrationality. The problem, then, is to understand how we can ideally make sense of all the relevant parts of Kant's work in order to address this issue. In this reconstruction, my main concern is not to offer an account of how Kant might be used to philosophically engage with contemporary problems of irrationality. Rather, I wish to sketch a model of how Kant himself would have thought about irrationality. The former type of work could freely abandon, or at least set aside, some very important theses and arguments of Kant's philosophical system. In this paper, however, I will attempt to make sense of Kant's (hypothetical) position on irrationality while considering all the relevant aspects of his philosophy. This entails that the most accurate reading of Kant's work may lead us to conclude that his account of this topic is not suitable for today's debates on irrationality.

For methodological clarity, I provide in section two an account of the meaning of certain concepts central to the discussion, such as 'rationality,' 'irrationality', 'action,' and 'prudential action'. In section three, I argue that irrational actions can only be conceived in opposition to practical rationality, and that they cannot be understood as mere products of intellectual mistakes or errors. Finally, in section four – which I take to be the central part of the paper – I compare current interpretative models of Kant's account of prudential irrationality, highlighting both their advantages and, especially, their limits. My conclusion will be that the two predominant models fail to provide an account that both preserves key elements of Kant's philosophy and allows us to further elaborate a theory of instrumental irrationality.

2. Some initial conceptual presentations

2.1. Rationality

Rationality has at least two central meanings in Kant's work: the rationality derived from (i) the good use of understanding and the one derived from (ii) the practical use of reason. In this sense, being theoretically rational means being able to make a good use of the (general) faculty

[2] In this paper, I adopt conventional models of citation and employ a series of abbreviations. The *Critique of Pure Reason* (KrV) is cited according to the A/B pagination. All other works of Kant are referenced following the *Akademie-Ausgabe* (AA) convention: *Critique of Practical Reason* (KpV, AA 05); *Critique of the Power of Judgment* (KU, AA 05); *Groundwork of the Metaphysics of Morals* (GMS, AA 04); *The Metaphysics of Morals* (MS, AA 06); *What Does It Mean to Orient Oneself in Thinking?* (WDO, AA 08); and *An Answer to the Question: What Is Enlightenment?* (WA, AA 08). For the sake of readability, I also abbreviate Categorical Imperative as CI, Hypothetical Imperative as HI, and hypothetical imperatives as HIs.

of knowledge (KU AA 05: 174-176, 195)[3]. Therefore, in its theoretical aspect, a rational agent is one that is constantly using his general faculty of knowledge in a critical way. A rational agent is also one that is not in a position of cowardice of laziness and can use his own understanding without the guidance of others (WA, AA 08: 35). In other words, he is one capable of thinking for himself[4] (WDO, AA 08: 146)[5]. Now let us consider practical rationality.

In the KpV, Kant argues for the practical reality of the freedom (KpV AA 05: 3-4). The argument is possible only because Kant developed a very distinct framework to deal with the relation with the object. While the general faculty of knowledge only seeks to represent an object, the faculty of desire, beyond the representation of the form of the a priori moral law, also seeks to create the object of desire through this representation (KpV AA 05: 44-45; KU AA 05: 178). This central distinction implies also a central differentiation between the practical rationality and the theoretical one. While the latter wants to represent correctly the object (i.e., according to the good use of the faculty of knowledge), the former seeks to create it[6].

For now, let us follow the *Groundwork* and grant that there are two general practical principles: one in which the desired end is conditional (the Hypothetical Imperative – hence HI), and another in which the desired end is unconditional (the Categorical Imperative – hence CI). Now, let us *assume, only provisionally*, that a rational agent, in their *practical realm*, is one that *(i) recognize his/hers desired ends* (whether conditional or unconditional) (*intellectual aspect*)[7], and *(ii) is capable of acting as a cause* in bringing about the objects in accordance with the representation of those desired ends (*volitional aspect*)[8].

[3] In a more strict sense, it also means being disciplined against the temptation of taking the ideas of reason as having a constitutive theoretical reality (B 421, B 737-740), and, therefore, not to take as objective what is merely subjective (KrV B 353–354).

[4] "Thinking for oneself means seeking the supreme touchstone of truth in oneself (i.e., in one's own reason); and the maxim of always thinking for oneself is enlightenment. Now there is less to this than people imagine when they place enlightenment in the acquisition of information <*Kenntnisse*>, for it is rather a negative principle in the use of one's faculty of cognition, and often he who is richest in information is least enlightened in the use he makes of it. To make use of one's own reason means no more than to ask oneself, whenever one is supposed to assume something, whether one could find it feasible to make the ground or the rule on which one assumes it into a universal principle of the use of reason. This test is one that everyone can apply to himself; and with this examination he will see superstition and enthusiasm disappear, even if he falls far short of having the information to refute them on objective grounds. For he is using merely the maxim of reason's self-preservation" (WDO, AA 08: 146).

[5] Klein (2023, p. 108-109) also correctly highlights the relation between this minority position and logical egoism (see V-Lo/Blomberg, AA 24:187; V-Lo/Blomberg, AA 24:151; V-Lo/Dohna, AA 24:740). Such logical egoism fosters authoritarian thinking, and 'thinking for oneself' cannot be equated with, nor provide any grounding for, such a stance. As Klein (2023, p. 108) observes: "although striving for enlightenment implies not accepting arguments from authority, it also entails a rejection of logical egoism, which can only lead to relativism and skepticism." Finally, being a rational agent, in its theoretical aspect, also means acting under the two other maxims of *Sensus Communis*, i.e., "to think from the standpoint of everyone else"; and "to think always consistently" (KU AA 05: 294).

[6] As Kant states (KU AA 05: 174; 176-179), these are completely different territories, which can only be unified by the faculty of judgment.

[7] Although it may seem relatively uncontroversial, this aspect is relevant for establishing the conditions of a Kantian theory of instrumental irrationality. As Martínez (2023) points out, Kant understood that not all representations are clear and evident; there are also obscure <*dunkel*> representations.

[8] In this definition, a rational agent may act against her own ends. Such an action would be considered irrational, yet she would still count as a rational agent. This means, of course, that a rational agent is not defined as one who never acts contrary to the ends she has set for herself, but rather as one who is capable of acting in that way. Within this not-so-intuitive framework, an "irrational agent" is not the opposite of a rational agent but instead a particular case of one. A person is called "irrational" when she performs a series of irrational actions, yet she is so only insofar as she also retains the capacity to act rationally. Accordingly, while the volitional aspect of rational agency consists in the possibility of acting either in pursuit of or in disregard of the ends one

With this initial account of what it means to be a (practical) rational agent, we can now introduce the concept of *irrationality* as having two possible meanings: *a failure* either in (i) the *intellectual aspect or* in (ii) the *volitional aspect*. As I intend to argue, *only volitional irrationality can be counted as a form of practical irrationality* and the intellectual aspect of rational agency will need a critical review.

2.2. Action

Distinct from other concepts, Kant offers few elucidative comments on how we can best present the concept of action. In the critical treatment offered in the KpV, there is no theory of non-moral or merely prudential actions. Since the investigation is specifically concerned with the foundations of morality, most of the commentary on the notion of action occurs in comparisons between good and bad actions, or, in other words, between actions that are legally and morally permissible and those that are not. At the beginning of Chapter 3 of the KpV (AA 05: 71-72), for example, he affirms that an action may contain legality, and among those, some may contain morality (if the will's determining ground is only the moral law itself). From this we can certainly conclude that (i) there are illegal actions (in the sense of forbidden according to reason), and (ii) there are legal actions that are not moral (actions performed in accordance with duty, but not from duty). Later in this chapter, he defends the thesis that some actions are called duties (AA 05: 80). These actions (duties) are conceived through the "exclusion of every determining ground of inclination." Again, by contrast, we can infer that actions which have some determining ground of inclination cannot be called duties[9].

For our purposes, it suffices to *concede hypothetically* (and also *provisionally)* that a given action can either fall within the moral domain or outside it (non-moral, as I will call it). We know well that within the moral domain, (i) actions may stem from a will whose maxim is universalizable to every rational agent, or (ii) they may stem from a will whose maxim, when universalized, entails a contradiction. The former, if also performed from duty (and not merely in accordance with duty), are moral actions. The latter are immoral and contrary to duty. The irrationality arising in this domain can easily be understood as a transgression of the moral law, and thus all type-(ii) actions are irrational. Hence, we understand here moral irrational actions as those that transgress the moral law. I believe this constructs a theory of irrationality without major controversial theses or conclusions.

Within the non-moral domain, there are actions whose subjective maxims cannot be taken as absolutely necessary and whose ends are not unconditional, but rather conditioned and relative to a particular agent. All of them would be counted as actions whose determining ground is inclination and, therefore, grounded not in a command of reason, but in a sensible element. This leads us to the question of whether a failure in an action of this field could be counted as a kind of irrationality; in other words, the central question is whether there is such a thing as irrational non-moral actions. As introduced previously, such a failure could hypothetically be (i) a failure arising from an intellectual error or mistake, or (ii) a failure caused by a volitional ground.

has previously set for oneself, any rational or irrational action is itself an expression of rational agency. When the action accords with the ends one has established, it is a rational action; when it does not, it is irrational.

[9] Another key feature of action is the existence of a double standpoint, which makes it possible to regard the very same action both as naturally determined and as transcendentally and practically free (KrV B XXVIII, B 566, B 579, B 585; KpV AA 05: 94, 99, 100, 104). I will return to this point in Section 4.2.3.

Once again, let us *assume provisionally* that prudential actions are a specific type of non-moral action, and they can be conceived as actions that promote (or at least attempt to promote) happiness. In other words, prudential actions are actions created by the faculty of desire taking the representation of happiness as its cause. In order to fully understand prudential actions, we must first understand the distinction between the two types of imperatives that Kant draws. In the next section, I will conduct an analysis based solely on the GMS to demonstrate that – even without the critical revisions Kant later introduced in the KpV and the KU – prudentially irrational actions cannot stem from intellectual errors.

3. The Theory of Imperatives in the Groundwork

3.1. Irrational actions outside imperatives?

Kant states that an imperative is the "formula of the command" which is, in turn, "the representation of an objective principle, insofar as it is necessitating for a will" (GMS AA 04: 413). Furthermore, "all imperatives are expressed through an ought" and they indicate that something ought to be done or omitted because it is good in a practical sense. Kant explains what is good in a practical sense and immediately contrasts it with the "agreeable" <*angenehm*>:

> Practical good <*Praktisch gut*>, however, is that which determines the will by means of representations of reason, hence not from subjective causes, but objectively, i.e., from grounds that are valid for every rational being as such. It is distinguished from the *agreeable* <*Angenehmen*>, as that which has influence on the will only by means of sensation from merely subjective causes, those which are valid only for the senses of this or that one, and not as a principle of reason, which is valid for everyone. (GMS AA 04: 413).

From this, it follows that every imperative operates according to the good, which is, in a practical sense, objectively good and stands in opposition to the subjectively good, namely, the agreeable. Therefore, this type of action falls outside the scope of imperatives, as it does not take the good as its foundation. Here arises the question: could irrationality be associated with what is subjectively agreeable but is not otherwise objectively good? If understood in a sense of *absence of reason*, then the answer is affirmative, since only what is practically good is good in a rational and objective sense. Thus, it is the case that guiding oneself by what is merely agreeable is irrational to the extent that there is no participation of reason in the course of action.

Certainly, this type of action (which aims at the agreeable) cannot be an action that conforms to any imperative. As Kant indicates in the footnote, it was proven in Section I of the *Groundwork* that an action done from duty cannot be one interested in the production of its object, but only for the sake of the action itself according to reason. It is also clear in the footnote that the agreeable is merely the anticipated object of the action. Thus, a will that acts out of interest in the object of the action is a will that acts according to a pathological interest (GMS AA 04: 413). Again, if we understand here the *absence of reason as irrationality*, then this type of action is certainly irrational. This certainly has a very different meaning from how we use 'irrationality' in *everyday language*, where it is generally understood more as *something contrary to reason* than as merely the absence of reason. Nonetheless, it does not seem to be the case that this type of action is irrational in the sense that we stated above, i. e., as a *volitional or an intellectual failure*. Textually, there is no indication that an action performed for the sake of producing the agreeable is irrational as contrary to a command of reason. One clear reason for this is the following: if an action is guided solely by the feeling of the agreeable and, therefore, bears no relation whatsoever to any duty, then it cannot be regarded as transgressing any form of command (or even coun-

sel) of reason. In such a case, reason is simply absent, and if it plays any role, it does so merely as an instrument of a will pathologically affected.

3.2. The Concept of Hypothetical Imperative

As we know, Kant distinguishes the imperatives between hypothetical and categorical. According to him,

> The former represent the practical necessity of a possible action as a means to attain something else which one wills (or which it is possible that one might will). The categorical imperative would be that one which represented an action as objectively necessary for itself, without any reference to another end. (GMS AA 04: 414).

I would like to stress the fact that both categorical and hypothetical imperatives are here stated as bounded to the duty and, therefore, to a normative account of action. However, as we shall see through Kant's critical revisions in the KpV and KU, this will no longer be the case. Furthermore, both imperatives are "formulas of the determination of action, which is necessary in accordance with the principle of a will which is good in some way" (GMS AA 04: 414). This passage makes it clear that a good will contains a principle guided by some source of imperative. As we know by the following line, this source of command could be either categorical or hypothetical. It is hypothetical only if "the action were good merely as a means to *something else*" while the categorical imperative represents an action that is good *in itself*. Kant then distinguishes hypothetical imperatives between imperatives of skill and imperatives of prudence (GMS AA 04: 414-415).

The imperative of skill is problematic because it guides action toward "any possible intention" (GMS AA 04: 415). It consists in finding the best means to achieve a given or conditioned end. The example of the doctor and the murderer makes this clear. According to Kant, the precepts for a doctor to cure someone and for a murderer to poison them with the same drug are the same and have the same value. Furthermore, the imperative of skill (or technical imperative) is possible according to an analytical principle of willing[10]:

> Whoever wills the end also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power. This proposition is, as regards the volition, analytic, for [...] the imperative extracts the concept of actions necessary to this end merely from the concept of the volition of this end [...] (GMS, AA 04: 417).

The imperative of prudence is assertoric, as it guides action toward a "real intention" (GMS AA 04: 415). According to Kant, beyond particular ends, all rational-sensitive beings have a common end, namely, happiness. Thus, the imperative of prudence is constituted by the "choice of means to one's own happiness" (GMS AA 04: 416). Furthermore, happiness is an (indirect) duty and an "ideal [of imagination]" from which "all inclinations are united in a sum" (GMS AA 04: 399). In terms of its necessity, the imperatives of prudence possess a greater necessity than the imperatives of skill, but not as great as the categorical imperatives, since only the latter have an unconditional necessity. As Kant indicates,

> The giving of counsel contains necessity, to be sure, but can be valid merely under a subjective, pleasing

[10] With this, Kant points out that we are not talking about the "execution of the action", but only about "how to think the necessitation of the will that the imperative expresses in the problem" (GMS AA 04: 417). As we will see, this is relevant because if one intends to develop a theory of Kantian irrationality based on the Groundwork, one cannot consider the necessitation that operates in the will but rather the execution of the action. This is counterintuitive and will be further elaborated upon later.

> <*gefälliger*>[11] condition, whether this or that human being counts this or that toward his happiness (GMS AA 04: 416).

In addition to carrying greater subjective necessity, the imperatives of prudence also entail a greater indeterminacy. The reason for this is that the imperatives of prudence are based on the concept of happiness, which (i) consists only of empirical elements and, at the same time, (ii) requires an "absolute whole, a maximum of welfare, (...) in my present and in every future condition" (GMS AA 04: 418). As is known from the KrV, the absolute of conditions cannot be known either transcendentally or empirically, and thus a sophism regarding the ultimate conditions of the world – whether with speculative interests or practical interests – is illegitimate and leads to a set of antinomic conflicts. In the GMS, this results in the impossibility of a finite being knowing for certain what one ought to do to be happy, since the totality of conditions is not given in the series of space and time. Happiness is not even an "ideal of reason," but merely one of the "imagination" <*einbildungskraft*>, and thus could not even constitute an object of practical reality (GMS AA 04: 418). For this reason, the imperatives of prudence are constituted only as "empirical counsels, e.g., diet, frugality, politeness, restraint," instead of "determinate principles" and they are in no way a "command of reason" (GMS AA 04: 418). Still in this same paragraph, Kant raises a serious doubt about the practical possibility of this imperative:

> the problem of determining, certainly and universally, what action will promote the happiness of a rational being, is fully insoluble, hence no imperative in regard to it is possible, which would command us, in the strict sense, to do what would make us happy (GMS AA 04: 418; emphasis added)

Later, however, he concedes that they would be "analytically practical propositions if one assumes that the means to happiness could be specified with certainty," and thus, "there is also no difficulty in regard to the possibility of such an imperative" (GMS AA 04: 419). Moreover, unlike the Categorical Imperatives, "their reality is given in experience" (GMS AA 04: 419-420).

3.3. Do Intellectual Failures Lead to Prudential Irrational Actions?

We now turn to the question of whether irrational actions can arise from intellectual mistakes inside the hypothetical imperatives. According to Merle (2023, p. 10-11), one source of instrumental irrationality lies in holding *incorrect beliefs*. Regardless of whether reason is present or absent, an agent may err in selecting the means to their ends. As Merle (2023, p. 10) notes, such mistakes can occur in two ways: (i) Misjudging the best means for given ends, or (ii) erring about whether one actually possesses the necessary means to achieve those ends.

The objection I wish to raise at this point concerns the possibility of instrumental irrationality – and, as I understand it here, the possibility of a prudential irrational action arising from an intellectual errors. I believe that the subject's error in these cases does not constitute prudential irrational action in either of the senses I proposed above. Let us take an example to illustrate this. Suppose I am walking down the street and suddenly feel thirsty. Therefore, my end is to quench my thirst. I consider buying water at the store or drinking the water I have at home. Since (i) my thirst is not great, (ii) I am close to home, and (iii) I want to save money, I decide

[11] It is important to state that the word "pleasing" <*gefälliger*> was changed by the editors in the Academy Edition to "contingent" <*zufälliger*>. Nevertheless, whether in the original or in the editors' version, Kant makes it clear that it is not the same subjective feeling that occurs prior to duty, namely, the "agreeable" <*angenehm*>. As noted, this term arises to distinguish an action according to an imperative from an action determined merely by the subjectively agreeable. On the other hand, "*gefälliger*" was used to differentiate the imperative of prudence from the categorical imperative, understanding the former as subjectively conditioned and the latter as unconditioned.

that the best course of action is to drink the water I have at home. However, when I arrive home, I discover that the city water company is conducting maintenance on my street and as a result I have no water at home. Since I am without water, I cannot drink it as planned. The conclusion is that the means I chose (namely, drinking water at home) proved to be erroneous because it was based on an incorrect *external belief*. Yet, it can be said that at the time of my decision, the action taken was based on the best possible conditions and, therefore, was rational. I was not actually in possession of the necessary means, but as shown, I had no way of knowing at the time that these means were unavailable to me. At that time, I took the best course of action based on the information I had and the critical analysis of the situation was sound. Therefore, not being in effective possession of the means does not imply the irrationality of the action, even if it implies an *external* wrong belief.

I call this an *external* wrong belief because I was led to a course of action over which I had no control or agency. If I am deceived by an external factor, can it really be said that I acted irrationally? It seems difficult to assert that a subject who has been deceived is irrational, because their course of action no longer depends on them. They acted based on the best possible presuppositions and conditions at the time, and it just so happened that, in this particular case, they were deceived by an external factor. I believe that this argument rules out the possibility of irrational action motivated by *external* factors or *wrong beliefs*.

Now let us examine the second possibility, namely, an *internal wrong belief*. Let us modify the thirst example a bit. In this example, let us suppose that I am already at home and have access to water. Let us also assume that I am thirsty, but in addition, I need to take a shower because I am sweaty from the walk. Finally, let us assume that I am short on time and therefore want to save time. In this case, instead of drinking filtered, good-quality tap water, I decide to drink water from the shower while bathing, because by doing this, (i) I quench my thirst, (ii) take a shower, and (iii) save time, as I am drinking water during the shower. A few days later, I end up getting sick, and the diagnosis indicates a likely causal link with the shower water I drank. Could it be said that I made an irrational decision? It does not seem so, because I fulfilled all the conditions and ends I set for myself at the time. However, something seems to suggest that the action was irrational.

The only condition under which the action described above could be considered irrational is to take into account not just this or that particular end within a given timeframe, but the totality of my ends across all time. This leads us to conclude that an *instrumental irrational action is impossible unless it considers the totality of ends across the entirety of time*. In the example above, my health is also a relevant end, and thus I did not use the best means to achieve the totality of my ends, given that this end, even if latent, was not fulfilled. This is a *necessary condition* for an action to be considered irrational with respect to the instrumental use of reason[12]. Furthermore,

[12] Since from this point on I carry out an analysis only from the point of view of practical rationality, and not theoretical rationality, it is nevertheless possible to consider here the case of an error stemming from theoretical irrationality. Let us suppose that I drank water from the shower because I saw a video of someone recommending this as a way to save time. Certainly, I did not make good use of my general faculty of knowledge, and I did not adopt a *critical attitude toward the information* presented to me. In other words, I used another's understanding to guide my life and remained in a state of minority. *From the point of view of theoretical rationality, I performed an irrational action*, since I did not carry out a proper procedure to investigate my beliefs. It is of utmost importance to emphasize that this differs from the notion of an intellectual failure in practical rationality, since the latter refers only to the failure in recognizing one's own self-imposed ends (for instance, a confusion). It is not about the procedure through which beliefs are assessed and subjected to criticism, but only about a misapprehension of the ends I desire (in this

as we have seen, imperatives of skill only concern a specific set of ends at a particular time, rather than the totality of ends across all time. An analysis based on the totality of ends would be subordinate to the imperatives of prudence and not to the imperatives of skill. Thus, from the standpoint of the execution of action[13] (where irrationality might *problematically reside*), there is no possibility of an irrational action arising from skill-based imperatives that are erroneously guided by false beliefs (whether internal or external). These imperatives lead us to a false positive irrational action, that is, an action that is in fact irrational if we take all the ends of an agent into consideration, but that seems to be rational if we consider only a specific set of those ends. Therefore, in a Kantian framework, this requires *conceptualizing non-moral irrationality through prudential imperatives rather than skill imperatives.*

In the case of imperative of prudence, it is difficult to conceive how someone could act contrary to such imperative. One clear statement that Kant does is that it is impossible to formulate universally valid objective rules as imperatives for happiness. This does not mean, however, that subjectively one cannot know through trial and error what brings him closer to or further away from his own happiness. Let us suppose I love studying and doing philosophy, but I am afraid of the future in terms of employment. Since I see that I am young, I might still give up and start over in another field, but I can also stay in the field and bet my chips on philosophy. Let us suppose I decide to stay and, in doing so, give up another career. This course of action will certainly lead to future outcomes, which are, at the moment, unpredictable. I may have good future outcomes in philosophy, or, even if I try hard, I might end up in a job I do not like and which, so to speak, makes me unhappy. Let us say this is the scenario and I end up in a job I do not like. Let's also say that this leaves me with a degree of happiness, on a scale from 0 to 10, of 5 points. Nevertheless, if I had taken a different course of action and decided to abandon philosophy and enter another profession, I still would not have the guarantee of a degree of happiness higher than 5. It could very well be the case that my degree of happiness is even lower, like 3 or 4. Thus, choosing one course of action or another cannot constitute an irrational action, as one cannot know what the best means to happiness (and its highest degree) are[14].

What can perhaps be said is that pursuing a course of action that I think may make me unhappy is irrational. Thus, if among the choices A, B, and C I reasonably believe that A and B have good chances of making me happy and C has no chance of making me happy, then it can be said that deliberately acting according to C is irrational. Therefore, an action is irrational in relation to the imperative of prudence only if we understand it not as a mistake to choose the best possible

case, conditioned ends). A possible objection might take the following form: in the case of the shower, could it not be said that I had a misunderstanding of the ends I desired? That is, I ignored health due to a miscomprehension of that specific end? In response, I consider that we must take into account, for the analysis, the totality of ends – something the skill imperative does not do, but only the prudential imperative can. The central point here is that, in the case of a rational agent's actions over time, one must always consider the totality of ends and not an isolated subset of specific ends. To isolate specific ends in such a way would amount to a grave error regarding the very notion of practical rational agency, as it would split the agent's ends arbitrarily – without reason – which is not only superfluous but certainly far from how we understand what it means to be an *unified agent* who desires or sets ends.

[13] Here, we are conducting an analysis based on the execution of an action derived from a technical imperative, and it is worth remembering that what constitutes each imperative, rather than the execution of each action, is simply its commanding relationship with the necessitation of the will. This is entirely consistent with the spirit of Kant's philosophy, since if we analyzed only the necessitation of each imperative in the will, we would end up merely reproducing the tripartite formulation of imperatives.

[14] Kant's examples are even better because they demonstrate that even desires for commonly valued goods – such as the pursuit of wealth, knowledge, longevity, and health – can lead to harmful ends (GMS AA 04: 418).

course of action, but merely as the deliberation in intentionally choosing a course of action that I conceive as having high chances of bringing me unhappiness. In this way, rather than irrational action according to this hypothetical imperative being a failure to maximize one's happiness, it is simply an act of choosing in favor of one's own unhappiness. This, however, is certainly not an intellectual error but a volitional failure. Such potentially irrational action would thus stem not from an intellectual mistake, but from a volitional one – that is, a case where despite clear representations, the agent's chosen course of action fails to maximize their own happiness. *Thus, we conclude that in no case do intellectual failures lead to prudential irrational actions*[15].

3.4. Setting conditions to Prudential Irrational Actions

As we can see, prudential irrational actions depend on two necessary conditions: (i) that happiness is conceptually distinct from instant pleasure[16], and (ii) that it is possible to rationally pursue happiness. If (i) is false, then no prudential action could be considered irrational, since the agent would always and immediately be choosing solely in favor of their own happiness (that is, every non-moral choice would simply be a choice aimed at increasing one's own happiness). If (ii) is false and reason does not play an active role in the pursuit of happiness – even if happiness and pleasure are indeed understood as non-identical – then it is not possible to judge an agent's action (or omission) in this pursuit as rational or irrational. Furthermore, I would like to stress that non-moral irrational actions should always be considered in light of all the ends of an agent and thus through prudential imperatives. As we have seen, taking into account only a specific subset of an agent's ends is problematic, insofar as it may lead to a *false positive* irrational action, i.e., an action that appears to be rational but is, in fact, irrational.

4. Two models of prudential irrational actions

If the argument presented in Section 3 is sound, we reach the following conclusion: *on a Kantian account, no practical irrationality arises from intellectual error*. We may now proceed to examine whether non-moral irrational actions can arise from volitional failures. For this investigation, I will employ two distinct interpretive models of Kant's theory of action. The first model, which I shall call the "negative prudential irrationality model", denies the possibility of irrational prudential actions. According to this model, all prudential actions must be considered at least as non-irrational. The second model – called the "positive prudential irrationality model" – makes the contrasting claim: prudential irrational actions are indeed possible. As we shall see, the for-

[15] This conclusion closely aligns with Klein's (2023) framework of irrationality. His approach rejects conceptualizing irrationality as mere failure or mistake, instead defining it as "a manner of thinking and acting grounded in erroneous principles" (Klein, 2023, p. 102).

[16] The example of a patient with gout serves as evidence for this distinction (GMS AA 04: 399). Kant tells the story of a gout sufferer who chooses to exchange his health for momentary pleasures, reasoning that he may not live much longer anyway. According to Kant, if happiness does not compel him through inclination, then it must compel him through duty. In such case, momentary pleasure must not prevail over the totality of one's inclinations – that is, over happiness itself. In this case, an irrational action understood as a transgression of the imperative of prudence is only possible insofar as it is an action that takes immediate pleasure as its foundation, rather than happiness. Perhaps this is what Kant meant when he referred to the "*agreeable*" as a purely subjective condition, in opposition to a *practical good*. As we shall see (4.1.4), for Timmerman (2022), both are happiness, and while the latter is the *broad* concept of happiness, the former is the *narrow* one. Moreover, we could say that happiness in its *broad* concept is a rational disposition toward the means of what could become a happy state of mind and therefore considers the means in the long term. This would represent the aggregate set of all satisfactions (an ideal of imagination). Meanwhile, in its *narrow* concept, happiness is the immediate satisfaction required by sensibility and thus considers the means in the short term. Here, the satisfaction arises from a single inclination.

mer model finds its basis primarily in Timmermann's interpretation of Kant, while the latter derives from Korsgaard's reading. As I shall argue, both accounts ultimately render non-moral irrational actions impossible – though they differ crucially in that Timmermann classifies prudential actions as non-moral, whereas Korsgaard treats them as steaming from unconditional reasoning and, therefore, from the CI.

4.1. The negative prudential irrationality model

In *Kant's Will at the Crossroads: An Essay on the Failings of Practical Rationality* (2022), Timmermann advances two central claims: (i) all prudential failures are *Socratic* – and thus intellectual failures – and (ii) all moral failures are *akratic* and consequently volitional. As established in Section 3, intellectual failures can never generate practical irrationality. Timmermann's central challenge, therefore, lies in demonstrating that prudential failures are exclusively intellectual, with no volitional component whatsoever. On his account, the HI is merely a principle that explains how non-moral actions come about in the world, not one that normatively evaluates those actions. His argument relies especially on Kant's reconsiderations in the two Introductions to the KU and in the KpV

In the KpV (AA 05: 25-26, 26n), principles of self-love are understood as principles that employ rules (or means) to achieve their ends. However, these principles are strictly theoretical. Kant's example is that if one wants to eat bread, one must build a mill. This is merely a rule of understanding that states the *intrinsic relation between cause and effect*. In the KU (AA 05: 171-173), Kant restates even more clearly the assimilation of rules of skill to theoretical principles and to theoretical philosophy. A common mistake, he claims, is to conflate the practical as a concept of nature (theoretical philosophy) with the practical as a concept of freedom (practical philosophy). The concept of "will" <*Willkür*> analytically contains the concept of the "practical", but we cannot know whether this "practical" refers to a rule of causality that derives from a concept of nature or from a concept of freedom. If it derives from the former, the principles are technically practical and belong to theoretical philosophy; if from the latter, they are morally practical and belong to practical philosophy.

What is of key interest to us follows directly from this distinction. Kant explicitly states that technically practical rules belong to theoretical philosophy, and that all rules of skill and all rules of prudence are technically practical. In other words, Kant is asserting that hypothetical imperatives do not belong to practical philosophy because their rule of causation does not originate in freedom, but in nature[17].

This point may seem subtle, but its philosophical implications for Kant's system of action are profound. When the concept guiding the will in the hypothetical sphere is nature, it follows that no free agency is possible. In fact, the entire context of the first part of the KU's Introduction is meant to distinguish theoretical from practical philosophy. While theoretical philosophy

[17] In the First Introduction (AA 20:199-200), Kant's philosophical revision regarding HIs becomes even more explicit. He states that it is time to "correct a mistake I made in the Groundwork of the Metaphysics of Morals." In his reformulation, Kant considers that he should have called the imperatives of skill technical imperatives and included the pragmatic (or prudential) imperatives within the set of technical imperatives. They still retain a proper name because, whereas technical imperatives (insofar as they are imperatives of skill) have a purpose <*Zweck*> that is given or presupposed as known, the pragmatic imperatives must still determine what this purpose consists in (namely, happiness) and not merely deliberate on how to achieve it. In other words, as an end, happiness is given to the agent, but its determination remains uncertain.

is grounded in the concept of nature, practical philosophy – through the concept of freedom – enables the possibility of a metaphysics of morals. Such a philosophical enterprise is only possible because free action is possible for the human will. When Kant isolates hypothetical imperatives to the domain of theoretical philosophy, he is implicitly denying that any free action can originate from them. Now, I would like to stress two major implications that follow from this critically revised version of the HI presented in KpV and in KU.

4.1.1. Determinism of the action

The first and most obvious implication is that one could not act differently when acting on the basis of HIs alone. Since HIs are conceived as concepts of nature, there is no space for freedom to be imperative or authoritative in the actions that arise from them. To be more precise, practical reason always plays a role in the process of setting ends. However, reason can never override the power of sensibility (and therefore inclination) in this realm. The key concept attached to the practical meaning of will is not freedom, but nature and its natural mechanism. This might not be convincing enough, so I would like to offer another argument in support of this claim. Instead of appealing to the critically revised version of the HI that Kant presents especially in the KU, let me defend this view only by reconstructing Kant's position in his theory of action as laid out in the GMS and KpV.

Kant states in the Third Section of the GMS (AA 04: 447) that "a free will and a will under moral laws are the same. Thus if freedom of the will is presupposed, then morality follows together with its principle from mere analysis of its concept." Let us first understand why morality always entails freedom. When my dog attacks my cat, I can certainly scold him for it. I might even say to him: "Bad dog! You're grounded for misbehaving." However, deep down I know he is not morally responsible for his behavior. I discipline him in a specific way to reinforce the behaviors I want from him (like not attacking my cat, for example). The situation would be completely different if, instead of my dog, a friend of mine attacked my cat for no reason. At the very least, I would certainly reconsider our friendship or even call the police.

What is the key difference at work here? It is that animals are not free and therefore their actions cannot be morally evaluated, whereas humans are. This distinction is grounded (at least in modern times) in our common-sense morality. As we can see, acting morally clearly entails acting freely. That is why Kant says that the *ratio essendi* of morality is freedom (KpV AA 05:04). Without freedom as the metaphysical foundation, morality would be nothing more than a "chimerical idea" or a "figment of the mind" (GMS AA 04: 445).

What is more difficult to grasp is why every free act is also a *morally relevant* or a *moral-type* act[18]. The answer lies in the relation between sensibility and reason. In its negative meaning, freedom can only be conceived as a detachment from sensibility (GMS AA 04: 446). One who acts solely according to sensibility acts exactly like an animal. No morality is derived from such action, for it is conceived merely as a mechanically caused response to the interplay of sensible forces. One detaches from sensibility only insofar as one considers one's maxim not as determined by sensible inclination, but as something that could be universally willed.

[18] Here I take morally relevant and from a moral-type as the same. As Hill (2013, p. 20) points out, immoral actions, i.e., actions against the unconditional duty are also free (otherwise they would not be imputable). Nevertheless, the *expression of freedom* can only be conceived as actions performed in accordance with the moral law.

We cannot know freedom apart from the moral situation. It is only through morality that we initially become aware of freedom. This is why Kant calls the moral law the *ratio cognoscendi* of freedom (KpV AA 05:04). Outside the moral sphere, we know nothing about freedom, for that would amount to trying to discern an idea beyond the bounds of practical experience. Detached from its practical use, the attempt to know freedom theoretically is the product of an undisciplined reason that falls into transcendental illusion. Without morality, the practical domain would be understood merely as a theory *explaining* actions. As Korsgaard (2008) points out (and as we will see in more detail later), it would not be possible to assess an agent's actions, and thus we would not be able to say whether a given action was right or wrong. In other words, normativity arises only insofar as there is a higher-order evaluative principle that allows actions to be judged as correct or incorrect. Through HIs alone, we would never reach the concept of freedom, for there would be no normativity enabling us to judge actions as coherent with any external principle. It is only the moral situation that grants legitimate entry for the concept of freedom into the realm of practical experience.

But does this prove that all freedom is exclusively moral and can never be instrumental? One might suggest that we gain initial access to freedom through morality but that freedom is not thereby limited to it. In response, I would ask: how could freedom pertain to actions that are not moral? For what reason should we believe such a claim? It seems like a dogmatic assertion that attempts to extend the use of a concept into illegitimate terrain. In other words, it may be metaphysically possible that the spontaneity of practical reason extends beyond morality, so that we might also be free in the purely instrumental domain (i.e., apart from morality). However, this would be solely a problematic proposition from the standpoint of theoretical philosophy. In the Kantian spirit, we must ask what would lead us to treat such a proposition as possessing practical validity – that is, as being an assertoric judgment for the practical domain. Such a move would be an illegitimate leap, for we have no grounds to claim that we know the idea of freedom outside its moral use. Hence, this metaphysically problematic judgment cannot become an epistemically assertoric one within the practical domain.

Now we can clearly see that (i) if an action is conceptually conceived as free, it is always also conceptually conceived as morally relevant. Thus, (ii) every free action is a morally relevant action. (iii) If an action is not free, then there is also no free agency. And (iv) where there is no free agency, there can be no irrationality. (v) If we represent an action as purely instrumental, then it is certainly not moral. It follows that (vi) *if an action is purely instrumental, it cannot be considered irrational, since it is not morally relevant and therefore not free*[19].

4.1.2. Analyticity vs. syntheticity

Another significant shift introduced by Kant's critical revision of the HI, particularly in the KpV and the KU, concerns how we can understand the concept of analyticity in relation to HIs. As Timmermann (2022, §19, p. 56) states, HIs do not require any motivational status to determine the agent's will. It is precisely because the agent is already determined by a specific end that she finds herself motivated to will the means to achieve it. Since there is no end beyond sensibility, the agent desires the end immediately and without any volitional resistance. As Timmermann

[19] Here of course I am considering one hidden premise, i.e., that every action can only be thought as purely instrumentally if and only if it arises directly from HIs. The CI cannot generate purely instrumental actions.

(2022, p. 59-61) explains, the hypothetical imperative, being an analytic practical principle, requires only that the agent has a sufficient theoretical understanding of the connection between means and ends in order to act. Once the agent knows all the necessary means to attain her end, she acts in accordance with the desired end. There is no room for alternative courses of action or genuine choice. To say that the proposition is analytic means that the agent acts immediately upon the representation of her end, as long as the means to it are sufficiently known[20].

Another way to understand the distinction between synthetic and analytic imperatives is by comparing them with the divine will. A divine will acts according to the moral law in a purely analytic manner – that is, it is sufficiently and always motivated merely by the representation of the law itself. In Kant's terms, the objective necessity of the moral law is identical to the subjective necessity of this law. In contrast, human beings have a will that is not purely intelligible but also sensibly affected. For this reason, the moral law is not immediately taken as the foundation of our actions. Thus, for us, the moral law can only be represented through a *synthetic* proposition, not an *analytic* one. The *syntheticity* of the CI indicates that sensibility stands between us and the representation of the moral law and, for this reason, the moral law is not immediately sufficient to motivate our actions. We are always divided into a two-fold possibility, whereas for a divine will there is always only one possible mode of acting. As previously argued, because we possess freedom of the will, this two-fold possibility is not merely illusory, but real. If we represent ourselves as lacking freedom of the will, two options remain regarding the status of the moral law:

(i) Action would always follow analytically and immediately from the representation of the law (as in the case of divine wills); or

(ii) The moral law would be empty for the will – even if it were represented, it would not be possible to act against sensibility. In this second scenario, action would be entirely determined by the sensible nature of the agent.

The first case describes divine wills; the second corresponds to the empiricist thesis that Kant criticizes from the KrV onward, especially in the third section of the GMS. In contrast, the synthetic character of the CI signals the obstacle that exists between the representation of the moral law and the sufficient motivation for producing the object demanded by it (which, according

[20] Timmerman points out that the normativity of HIs lies in the requirement of knowing a series of theoretical rules in order to achieve ends. Ignorance of the means implies failure in the fulfillment of ends. The notion of an imperative as a rule that can possibly be violated by the agent is thus preserved, since the agent may fail to follow the theoretical rules that he does not know in order to reach the desired end. Such a violation, however, does not necessarily occur knowingly or willingly. As becomes clear, this error is merely intellectual rather than volitional, for it does not indicate any lack (or weakness) of motivation of the will, but only a will that does not know the necessary means to achieve its ends. If I want to bake a cake but do not know the ingredients, then there is a transgression of the hypothetical imperative insofar as I want the end but do not know the means. Nevertheless, even within the commentator's argumentative framework, this seems mistaken: if I want the end but do not know the means, I cannot truly want the end, for I cannot want the means. If I want to travel to Switzerland but have no money, then I merely have a wish <*Wunsch*> and not a desire <*Begierde* (*appetitio*)> (Anth AA 07: 251). The latter is connected to the capacity (or *faculty*) to desire <*Begehrungsvermögen*> (MS AA 06: 211), whereas the former, when it does not become a ground for the faculty of desire, is a peevish wish <*launische Wunsch*> (Anth AA 07: 251). The HI entails the following structure: If I want to go to Switzerland, then I also want the money to do so; and if I am not willing to pursue the money I lack, then I do not truly want to go to Switzerland. I might wish for it, but I do not genuinely desire it. Wishes arise in the mind and may by themselves motivate action. Desire, however, is expressed in the world as a form of causality. I may have many wishes, but I desire only one thing (and therefore must abandon contradictory wishes). As I understand it, this allows only a *post factum* explanation of what the agent really desired in the past. This is very close to how Schopenhauer (1960) conceptualized the will. In his account, I can only know what I truly wanted in the past and never know exactly what I want in the present. Willing can only be seen in the world retrospectively, through actions that have already taken place.

to KpV, is the good). The agent can therefore either succeed in following the moral law or fail morally. This dual possibility, opened by the freedom of the will, is crucial for moral imputation. Without it, morality would be no more than a "figment of the mind" or a "chimerical idea" (GMS AA 04: 445). As we can see, synthetic propositions are those in which there is a volitional obstacle between the representation of the end and the action that would make that end real.

Furthermore, the analytic character of HI also correlates with the impossibility of contradiction as presented by Kant in the KrV. If in analytic propositions the predicate and the subject are thought through a relation of identity, then to affirm one while denying the other generates a contradiction. It would be like saying, "It is not true that every bachelor is unmarried." This statement is contradictory because the concept of "unmarried" is already contained within the concept of "bachelor." Someone who genuinely tries to defend the truth of such a proposition surely does not grasp at least one of the two concepts involved.

In the practical field, this would be like wanting to make coffee but refusing to fetch the coffee grounds, filter, and dripper from the drawer. This is a practical contradiction and, in this case, if one does not fetch the filter, the grounds, and the dripper, then he does not really want to make coffee. The act of making coffee and fetching its required elements from the drawer is a single act, and thus there are no extrinsic or instrumental desires derived from the desired end (making coffee). The means are immediately desired from this end and, in this way, the agent intrinsically wants to get the grounds, the filter, and the dripper. If he receives some friends at home and says that he will make coffee for them, but instead of using the coffee grounds he grabs the guarana powder that was right next to it in the drawer, we would surely interpret this oversight as a mistake or failure in the faculty of judging the appropriate means (that is, as an *intellectual* error)[21]. I find it difficult – following Timmerman (2022 §21–23) – to think we would judge the case as one involving weakness of will. In other words, we would judge the case as involving an intellectual mistake, but it would not follow from that that one acted irrationally in taking the means to his end. Thus, the failure to comply with the HI means solely an intellectual error on the part of the agent in understanding the actual causal relation between means and ends. If at time T2 I realize that means M2 were worse than M1 for achieving end E1, then there was an intellectual error, which, however, does not imply an irrational action. The HI becomes a theoretical apparatus for evaluating actions from a non-moral and, in a certain sense, non-volitional standpoint[22].

Analytic practical propositions indicate the immediate necessity between an agent's ends and the action undertaken to achieve them. Given their analytic relation to the moral law, the divine will can never fail to conform to it, and thus no command is required. The same holds when we, as human beings, regard ourselves as situated outside the moral domain. Kant understands HIs as analytic propositions, since anyone who genuinely wills the end faces no volitional obstacle to also willing the means. This, however, does not exclude the intellectual obstacles we may encounter in relation to the means required for our ends.

This is especially the case with happiness. We cannot be certain whether M1, M2, … Mx will

[21] This is very similar to the conclusion we reach in section 3.3, i.e., irrational actions can never stem from imperatives of skill.

[22] In Timmerman's words (2022, p. 57): "We violate hypothetical imperatives if and only if we fail to grasp the instrumental connection they contain, i.e. out of ignorance. If we act differently when we are clearly aware of the means and it is known to be at our disposal, we are left to conclude that we do not desire the end with the requisite force, that we want something else more than the end we think or say we prioritize; and this hardly counts as rational failure, even less as practical irrationality."

make us happier, and we may become intellectually confused in this realm. We choose M1 at T1, later realize it was not as fulfilling as we expected, and then switch to M2 at T2. At T2, however, we might reflect again and conclude that the former was actually better, leading us to return to M1. Let me illustrate this: suppose someone pursues a career as a teacher, abandons it to make a living as an artist, and then realizes that being a teacher was in fact more satisfying, returning to the classroom. This shifting of means to achieve happiness is constitutive of human life and, according to this model, the change is due to intellectual errors made along the way rather than to volitional weakness. This is why Kant states in GMS (AA 04: 418) that imperatives of prudence are, at least with regard to the determination of their object, impossible.

4.1.3. Abandoning ends

Yet, making an intellectual mistake cannot explain how we abandon ends in our daily lives. The coffee-guarana example is illustrative in showing how, given a specific end, an error in the choice of means does not imply irrationality. It does not explain, however, the abandonment of ends I once set for myself in the past. Instead of the explanation that there may be means-end irrational actions, Timmerman defends the hypothesis that the abandonment of an end is merely the abandonment of a means in relation to the totality of ends, that is, happiness. In other words, abandoning a specific end simply means abandoning an end that no longer makes sense within the overall calculation of happiness. Put differently, this end no longer promotes the same degree of happiness as was once expected. In this case, according to Timmerman, there is nothing irrational about the abandonment of ends, and a theory of irrationality could not be derived from such cases.

Let us take an example to illustrate this. Suppose Gabriel is a young university student who has set himself the goal of improving his academic performance this semester and, therefore, plans to engage more deeply with assignments and activities that contribute points in the courses he is enrolled in. Moreover, he has also decided to take better care of his health this semester and is thus aiming to lose fat and gain lean muscle through a specific diet. Now, at the end of the semester, he knows he must focus especially on writing an essay for his ethics class and maintaining a healthy diet. On Thursday morning, he thinks: "This week I haven't done everything I needed to. But today will be different. Today I'll commit." Around 11:30 a.m., contrary to what he had planned for himself, he no longer wants to have lunch and then write his paper; instead, he now prefers to play video games. *He is fully aware of the consequences*: (i) he will feel hungry and weak throughout the afternoon; (ii) he will not meet his nutritional goals for gaining lean muscle; (iii) he will have to write the essay later in the week; and (iv) the quality of the essay may suffer, possibly lowering his grade. Timmerman's (2022, p. 58) account would say that Gabriel simply abandoned the end of writing his essay because he now considers that skipping lunch and playing video games all afternoon would yield a greater total amount of happiness. Gabriel would be committing a HI mistake only and only if he considered something false as if it were true (for instance, that not having lunch would not make him feel hungry).

For Timmerman, the intellectualist position readily explains why Gabriel acted as he did: Gabriel now considers that his overall happiness is greater than it would have been had he followed through with his previous plans. In this way, Timmerman (2022, p. 58-59) reduces the HI to an explanatory principle of non-moral actions, one that does not serve to normatively assess

the rationality of actions, but merely to indicate and clarify *post factum* what the agent's actual reasoning was at that moment in terms of maximizing his overall happiness. The advice, for example, "You should not make a lying promise, so that if it were revealed then you would lose your credit" (GMS AA 04: 419), contains a conditioned duty, but one that can ultimately be abandoned by the agent, even without knowing for sure whether such abandonment will bring about greater happiness or unhappiness. Timmerman (2022, p. 58) claims that it will never be possible to know whether the abandonment of one particular end in favor of another has, in fact, generated more or less happiness than would have been the case in the counterfactual scenario.

4.1.4. The indirect duty to happiness

Although well-articulated and textually well-grounded, this negative model faces a very specific challenge. In a few passages of his work, Kant refers to a duty to promote one's own happiness (GMS AA 04: 399; KpV AA 05: 53; MS AA 06: 387-388). This is potentially problematic for the negative model, insofar as a moral duty to act in ways that enhance one's own happiness can only make sense if actions concerning happiness are understood as involving a rational aspect. Thus, happiness cannot, in this case, be a concept that immediately drives the agent toward the most pleasurable action (within the totality of the sensible play), but must instead be understood as a possible object that the agent may choose not to pursue in favor of another. In other words, if the agent must choose happiness as a duty, what would be the alternative choice? What other course of action remains when one fails to fulfill the duty to act for one's own happiness? As we have seen, one possible response would be to dissociate the feeling of the *agreeable* from the *practical good*, and to understand happiness as residing only in the latter. However, another condition is necessary for duty to arise. A duty can only emerge if the agent has the possibility both to follow and to transgress it. Without this dual possibility, there is no choice, and therefore no duty. This implies that the agent could deliberately choose against their own happiness – something that, up to this point, the negative model cannot accommodate.

According to Timmerman (2022 §25, §36), imperatives of skill are sometimes also necessary for the use of the CI. That the happiness of others is a duty is clear from the CI. However, the complexity of the world makes its application difficult. HIs enter here as tools for the application of the duty of beneficence. The case of the indirect duty to promote one's own happiness also falls within this scope. Promoting one's own happiness is merely a means to attain the moral end, and thus the hypothetical imperative to promote one's own happiness is only a tool used by the agent in order to reach that moral end. *Instrumental reason is here employed as a tool of moral reason and, outside of it, possesses no normativity*[23]. As we will see, Korsgaard will defend the same conclusion.

In this way, perfect duties have a clear application in the world. Imperfect duties, however, may require hypothetical imperatives to guide action. Normativity stems solely from pure practical reason, and instrumental reason is subordinated to it in this specific case. Mistakes in the execu-

[23] In contrast to Kohl (2018), Timmerman argues that not all hypothetical imperatives derive solely from one's own happiness. However, Timmerman maintains that the same instrumental procedure operates both (i) in the pursuit of one's own happiness (which is not in itself a duty) and (ii) in the promotion of others' happiness (which is a duty). For instance, taking a taxi to go to the bank and pay my debt requires the use of technical imperatives for the fulfillment of this duty. The action will have moral worth if it is carried out by taking reason as the sufficient cause of the action – that is, by acting from duty – and it will lack moral worth if the action is merely in accordance with duty.

tion of imperfect duties must be understood as errors in instrumental deliberation, rather than volitional moral failures.

The case of the gout sufferer (GMS AA 04: 399) pushes the difficulty raised here to its limit. The gout sufferer faces the following problem: do I pursue the ideal of happiness (through health) and thereby renounce immediate pleasure, or do I pursue immediate pleasure and thereby renounce the ideal of happiness? The ideal of happiness, let us recall, is merely an ideal of the imagination, and the "sum of all inclinations" can never in fact be realized in this world. This is the *broad* sense of happiness. In contrast, the *narrow* sense of happiness is the pursuit of a single inclination at the expense of the imaginary sum of all inclinations. The dilemma lies in the fact that the agent does not truly know what will make him happier: perhaps it is the *immediate* (*agreeable*) pleasure, or perhaps it is the *ideal* of happiness (*practical good*). He acts only according to the guiding principle of pleasure and, thus, will take the course of action that he sees as having the greatest chance of making him happy. At this point, there is no normativity yet, and he acts merely in accordance with *felicific happiness*. So far, neither option is rationally preferable or dismissible.

Following Timmerman (2022, §36), we can argue that normativity arises only insofar as the agent encounters the indirect duty to promote material conditions that facilitate the fulfillment of duty. Among these conditions there is a subcase of happiness, namely, health. In this specific case, happiness is taken as a means for promoting the unconditioned duty. That is, *when I am confronted with duty*, I can then choose between a *broad* or *narrow* conception of happiness. Outside the representation of an unconditioned duty, such a choice cannot be made deliberatively. Agents may always be playing the game of choosing between the ideal of happiness and happiness as mere pleasure, but this game is not grounded in reason. Inclination has the final word. Under the conditions established in section 3.4, criterion (i) is satisfied whereas criterion (ii) is not. However, when reason enters the game with the concept of a duty (in specific cases such as not placing oneself in a position that undermines the fulfillment of morality), the agent then has a real choice between the imaginative ideal of happiness and momentary pleasure. This choice between happiness and momentary pleasure does not occur in every case, but only hypothetically in cases where happiness is taken as a means to the fulfillment of some unconditioned duty. This is of great importance, as it explains why Kant affirms that happiness *may* become a duty in some cases, yet he never claims that it *is* a duty. Thus, *the standard position is that happiness is not a duty, but in particular and specific cases it may become an indirect duty*. The gout sufferer's mistake in choosing momentary pleasure instead of happiness as an ideal becomes a volitional error only insofar as the concept of duty comes into play. Without it, the choice between general satisfaction over time and momentary pleasures is simply a choice guided analytically by what the interplay of faculties presents to the agent as producing the greatest happiness. In the case of the gout sufferer, were it not for the representation of the duty to maintain a position conducive to fulfilling duty, he would not have acted irrationally.

4.1.5. Synthesis of the negative model

As we shall see, this model has the historical and interpretative advantage of accounting for all parts of Kant's work relevant to the issue. It can adequately explain the critical revisions undertaken by Kant in KpV and KU without abandoning major theses stated in GMS, such as the

radical distinction between the analyticity of the HI and the syntheticity of the CI. Philosophically, its account of non-moral actions closely resembles the way Hume presented his theory of instrumental rationality. This is entirely accurate, and yet it does not reduce Kant's theory of action to a kind of Humeanism. This model merely presents his theory of prudential actions as ultimately determined by the sensible aspect of the agent. I recall here that this by no means implies that reason is absent, but only that it does not override inclination in this domain. It is also not true that instrumental rationality is absent; rather, such rationality exists only insofar as it is subordinated to moral rationality. Moreover, when guided merely by HIs, the errors agents may commit are by no means expressions of irrational actions. They are merely intellectual errors that can be analyzed *post factum*. As long as it is subordinated merely to a technical imperative, happiness is pursued mechanically. *Therefore, non-moral actions (including those that concern happiness) are not susceptible to volitional errors, which are the only ones that truly give rise to irrationality. Although historically accurate, this model does not allow us to construct a Kantian theory of instrumental irrationality.*

4.2. The positive prudential irrationality model

Whereas the negative model interprets Kant's theory of prudential action as a species of Humeanism, the positive model asserts the very opposite: instrumental reason – including prudential actions – is normative because it is grounded in unconditional reasoning. In this positive model, it is practical reason, rather than inclination, that ultimately governs prudential action. Korsgaard is perhaps the foremost defender of this positive model. In her essay *The Normativity of Instrumental Reason*, she argues that (i) instrumental reason alone cannot fully account for how we deliberate about our ends and, therefore, (ii) it is insufficient on its own to ground a theory of instrumental irrationality. To address this, (iii) unconditional reason must play a significant role in the process of setting and acting upon ends.

4.2.1. Establishing the argument

Korsgaard's argument begins with the thesis that both instrumental and moral reason share a common source: "the autonomy or self-government of the rational agent" (Korsgaard, 2008, p. 23). To defend this view – just as Kant did – she must criticize two major philosophical traditions: empiricism and rationalism. To illustrate her criticism of empiricism, which she takes Hume to exemplify most clearly, let us turn to Howard's case (Korsgaard, 2008, p. 39). Suppose Howard, a man in his thirties, wants to live past the age of fifty and must take a series of injections in order to do so. However, Howard has a deep aversion to injections and therefore refuses the treatment. Let us further assume that Howard makes no intellectual error: he fully understands the consequences of his omission. According to Hume's account – as well as Timmerman's version of Kant – we cannot say that Howard is acting irrationally: he is simply acting on what he most strongly desires. His omission is merely a means to an end that he wills, namely, avoiding injections. As we have seen – and as Korsgaard (2008, p. 40) rightly notes – under the negative model, he cannot even be said to violate the principle of instrumental reason. As Korsgaard summarizes:

> The problem is coming from the fact that Hume identifies a person's *end* as what he *wants most*, and the criterion of what the person wants *most* appears to be what he actually *does*. The person's ends are taken to be revealed in his conduct (KORSGAARD, 2008, p. 40).

In this distinctly Schopenhauerian fashion, this account of prudential action can never be evaluated in normative terms. By itself, the prudential (instrumental) principle of adopting the best means cannot generate normativity. Normativity arises only insofar as we take the agent to be capable of freely committing to or abandoning ends. Rationality – and thus irrationality – emerges only when the instrumental principle is supported and grounded in a principle for choosing ends. That is why "Hume has no resources for distinguishing the activity of the person herself from the operation of beliefs, desires, and other forces in her" (Korsgaard, 2008, p. 45). These sensible elements may be sufficient, under the negative model, to explain human action entirely – but certainly not to evaluate it normatively.

In its revised form, the instrumental principle has, according to Korsgaard (2008, p. 46), the following structure: "if we will an end, then we ought to will the means to that end"[24]. Here, an important comment must be made. Korsgaard claims that the distinction between will (or desire) and wish is not available within the Humean empiricist framework. While this may well be the case for Hume, it is certainly not the case for the way the negative model was previously presented. As exemplified by the case of Switzerland, we can distinguish a mere wish from a true desire – but only through *post-factum* analysis. If Korsgaard is right, however, the positive model enables a distinction between desire and wish *ex ante*, i.e., prior to the action's actual occurrence. If volition is attributed to the capacity to freely adopt ends (even prudential ones), then we are no longer dealing with a will that merely results from the conflicting interplay of multiple wishes. This means that we can, at the very least, regard the agent's action as the product of a free will – and, therefore, that the agent is internally capable of distinguishing between what she merely wishes and what she truly wills. Within her own mental state, the agent can identify *ex ante* what she genuinely wills and what she merely wishes.

In this formulation, the 'ought' necessarily presupposes the possibility of transgressing this very ought. Beyond being a rational demand, it is a rational demand that can be freely violated – that is, willingly left unfulfilled by the agent. Korsgaard (2008, p. 51) provides one way to represent this structure syllogistically

> Whoever wills the end wills the means insofar as he is rational.
> I will the end.
> Therefore I will the means insofar as I am rational.
> Therefore I *ought* to will the means.

However, as she rightly observes, it is impossible to provide a non-trivial explanation for why an agent understands as necessary the adoption of means to their ends, since rationality (even in its volitional sense) is already presupposed in the premise. Thus, we cannot properly grasp the agent's *motivational foundation* for committing to the means for his end. This constitutes a

[24] This formulation of the instrumental principle is textually grounded especially in the following passage: "How an imperative of skill is possible requires no special discussion. Whoever wills the end, also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power." (GMS AA 04: 417). This passage has generated numerous readings of the hypothetical imperative as a normative or prescriptive principle, rather than a descriptive one. Timmerman's (2022, p. 62) interpretation of this passage, who presents the HI as descriptive, is offered in the following terms: "It remains true that imperatives apply only to those rational agents that can violate them; but it does not follow that they must be able to violate them knowingly or willingly. The parenthetical remark is perfectly compatible with there being a multiplicity of technical rules that we find hard to follow because we have trouble finding out what they say." A less structural and more historical interpretation would simply state that Kant revised this in the KpV and in the KU. Surprisingly, Korsgaard (2008, p. 48) calls the parenthetical part a caveat and notes that this would raise further problems for Kant's argument.

pivotal moment in her argument, as Korsgaard must partially abandon a distinction central to Kant – namely, the analytic-synthetic distinction.

4.2.2. Exegetical and historical limits of the argument

This brings us directly to the exegetical and historical limits of Korsgaard's argument, when Kant's own work is taken as the standard. As previously stated, Korsgaard must partially abandon Kant's analytic-synthetic distinction in order to reinterpret the hypothetical imperative. According to her, there is a "historical explanation for what has gone wrong here" (Korsgaard, 2008, p. 51). She argues that Kant falls into a kind of realist rationalism (*à la* Leibniz and Clarke) when he considers that the moral law would be analytic for us if our will were perfect. If the moral law is expressed through a series of truths or facts external to our agency, then the internalist criterion for why we ought to act rationally can never be fulfilled. And if that criterion is not satisfied, no coherent theory of rationality can be sustained[25]. In her new and final formulation of the instrumental principle, Korsgaard presents it as the commitment an agent makes to the end they set for themselves. As Korsgaard (2008, p. 58) explains:

> [F]or the instrumental principle to provide you with a reason, you must think that the fact that you will an end is a reason for the end. It's not exactly that there has to be a further reason; it's just that you must take the act of your own will to be normative for you.

This formulation, however, differs very little from the third formula of the CI provided by Kant, as Korsgaard (2008, p. 58) herself acknowledges. The analyticity of the principle is certainly called into question here[26]. What gradually takes place in her argument is the incorporation of the HI into the CI. As a result, the HI – and its normativity – becomes a special case of the CI. This is certainly a philosophically rich and fruitful construction. As Korsgaard points out (2008, p. 60 fn), one of the conclusions of her essay is that there can be no instrumental norms unless there are also unconditional ones.

However, this interpretation raises historical and exegetical concerns. The reduction of all practical rationality to the CI is not, in itself, a serious problem. The greater challenge lies in demonstrating why we cannot simply suppose that Kant did not regard the instrumental principle as rationally guided – that is, subject to normative assessment. In other words, why should the internalist requirement regarding the instrumental principle concern us? To support her case, Korsgaard (2008, p. 60-61, 67) suggests that there is a contradiction in those who fail to act on the instrumental principle: if you do not act on this principle, you cannot claim to have a unified will. Yet, in Kantian terms, there is no clear reason to suppose that such unity must be grounded in the instrumental principle. One could argue, instead, that the unity of the will derives solely from pure practical reason – that is, from acting according to the CI alone. This could be a potential response by those defending the negative model.

[25] This concern with the internalist criterion already appears in *The Sources of Normativity* (Korsgaard, 1996, p. 14; 46).

[26] She attempts, to some extent, to preserve the analyticity of the principle by asserting that "to will an end just is to cause or realize the end, hence to will to take the means to the end," and thus that "the instrumental principle is constitutive of an act of the will. If you do not follow it, you are not willing the end at all." This, however, can only be seen as a metaphor for how Kant described the analyticity of practical principles. In Korsgaard's position, the analyticity of the instrumental principle derives from a double exclusionary condition: either you will the end, or you abandon the instrumental principle. Yet I could not rationally abandon the principle, since it is constitutive of the will. For Kant, however, this formulation would count as a synthetic one. If I can – albeit at the cost of being irrational – freely abandon the principle, then my relation to it is synthetic. As Kant maintains, the analyticity of a principle derives from its irresistible motivational force upon the will. An analytic principle, by definition, can never be transgressed.

Furthermore, Korsgaard states that "Kant's account" of "the principle of prudence as a hypothetical imperative (...) is in need of revision" (2008, p. 29 fn). She also claims that it is necessary to offer a kind of "deduction" for HIs, analogous to the one Kant provides for the CI (Korsgaard, 2008, p. 29). The problem is that Kant explicitly states that all hypothetical imperatives are grounded in experience and do not require any special deduction (GMS, AA 04: 419-420). At a certain point – and this is of major significance for a historical interpretation – Korsgaard appears to abandon the analytic-synthetic distinction altogether. She concludes: "I am inclined to think that my argument shows the distinction to be less important than Kant thought" (2008, p. 62-63 fn).

4.2.3. Expanding the model

Building upon Korsgaard's considerations – yet also going beyond them – I would now like to show how the positive model could be expanded. Following Korsgaard (2008, p. 67-68), the first premise we must accept in order to develop this expanded model is that the hypothetical imperative is merely "an aspect of the categorical imperative," and thus that "there is only one principle of practical reason, and it is the categorical imperative."

The advantage of this model is that it can certainly provide a better account of Kant's considerations regarding the indirect duty to promote one's own happiness. To support this, we take the concept of *Willkür* (choice or elective will) as central and divide duties relating to the agent into two levels: (i) a direct or superior level and (ii) an indirect or subordinate level. The first level concerns maxims that are universalizable in the Kantian sense – that is, maxims that could be willed by all rational beings in any circumstance. From this level derive the perfect and imperfect duties of both right and virtue. At the indirect and subordinate level, there exists a single duty to pursue one's own happiness, whose maxims are not in themselves universalizable. This level is entirely conditioned by the superior level. Therefore, the agent may legitimately pursue their own happiness, provided that doing so does not violate the higher-level duties. In other words, this subordinate duty is only expressed when it does not contradict the direct duties. Finally, this expanded positive model allows for a more substantive distinction between the ideal of happiness (practical good) and immediate pleasure (agreeable).

At this point, a highly relevant issue arises. Once we accept the premise that the CI is the sole principle of practical reason, two possible paths emerge: (i) we may consider that every action is moral (at some level) and therefore there is no space for rationality outside morality (even within actions concerning one's own happiness); or (ii) we may consider that not all actions falling under the CI's domain are moral. If we adopt (i), then we are led to the conclusion that non-moral irrational actions do not exist, since every action is, in some sense, moral. In this case, prudentially irrational actions would still exist, but they would be considered as belonging to the domain of morality rather than to a non-moral domain. On the other hand, if we adopt (ii), then the instrumental principle that guides, for instance, the pursuit of happiness could be seen as a principle grounded in an unconditional principle that nevertheless lies beyond the sphere of morality. In that case, there would be norms of reason that are non-moral (such as, perhaps, the pursuit of one's own happiness).

Nevertheless, I believe that both (i) and (ii) would be highly controversial from a Kantian perspective and would require abandoning key components of his philosophy. Both positions

fail to capture the strict correlation Kant establishes between free action and morality. Kant is particularly concerned with demonstrating that there is a necessary correlation between morally relevant actions and free actions, as we have previously elaborated. It would be far more plausible, from a Kantian standpoint – contrary to what (i) maintains – to recognize the existence of a broad set of non-moral actions, namely, those determined mechanically by nature and which do not contradict the moral law. Choosing whether to buy a house or an apartment is hardly a non-action, but it is also difficult to regard it as a moral action. Moreover, to defend (ii), one would need to argue that every action is, in fact, a free action. This would entail a philosophical structure quite distinct from Kant's, since freedom is understood by him as the *ratio essendi* only of the moral law and, therefore, only of morally relevant actions (KpV AA 05:04). In such a model, *Willkür* would need to possess a spontaneity so extensive that it would surpass even the bounds of morality itself.

5. Concluding remarks

We now enter particularly challenging terrain for the development of a robust theory of irrationality grounded in Kant's philosophy. If we align with Timmerman's position, instrumental irrationality does not exist, and all forms of irrationality must be understood as fundamentally moral. If we adopt Korsgaard's view, we face two problematic alternatives: either every action is regarded as moral – resulting in an inflation of the moral domain – or the essential link between free will and the moral law is undermined. In both scenarios, the synthetic-analytic distinction collapses. The central challenge for those attempting to construct a Kantian account of instrumental irrationality is therefore to explain how instrumental rationality can be possible without reducing it to the moral domain, while simultaneously preserving the intrinsic connection between the moral law and freedom. Put differently, the task is to formulate a theory that satisfies three major conditions: (i) reason exercises a decisive and authoritative role in determining and acting upon instrumental ends; (ii) this form of reasoning is not reducible to morality, and thus does not merely dissolve hypothetical imperatives into components of the categorical imperative; and (iii) this free end-directed agency occurs without presupposing a spontaneity of reason that transcends the boundaries of morality itself. Condition (i) is rejected by Timmerman, whereas conditions (ii) and (iii) remain unfulfilled within Korsgaard's account.

Bibliographic References

HILL, T E. Jr. 2013 Kantian autonomy and contemporary ideas of autonomy. In: SENSEN, O. *Kant on Moral Autonomy*. Cambridge University Press.

KANT, I. 1996. An Answer to the Question: What Is Enlightenment? Trans. James Schmidt. In: SCHMIDT, J. (Ed.). *What is Enlightenment?*: eighteenth-century answers and twentieth-century questions. London: University of California Press.

KANT, I. 2006. *Anthropology from a Pragmatic Point of View*. Trans. Robert B. Louden. Cambridge: Cambridge University Press.

KANT, I. 1987. *Critique of Power of Judgment*. Trans. Werner S. Pluhar. Indianopolis: Hackett Publishing Company.

KANT, I. 1997. *Critique of Practical Reason*. Trans. Mary J. Gregor. Cambridge: Cambridge University Press.

KANT, I. 1998. *Critique of Pure Reason*. Trans. Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

KANT, I. 2002. *Groundwork for Metaphysics of Morals*. Trans. Allen W. Wood. New York: Yale University Press.

KANT, I. 1900-. *Kants gesammelte Schriften*. Berlin & New York: Walter de Gruyter.

KANT, I. 1991. *The Metaphysics of Morals*. Trans. Mary J. Gregor. New York: Cambridge University Press.

KANT, I. 1996. What does it mean to orient oneself in thinking? Trans. Allen W. Wood. In: WOOD, A. W.; DI GIOVANNI, G. (Orgs.). *Religion and Rational Theology*. Cambridge: Cambridge University Press.

KLEIN, J. T. 2023. Enlightenment as the normative principle of social rationality. *Studia Kantiana*, Curitiba, v. 21, n. 1, p. 99–117. DOI: 10.5380/sk.v21i1.91982. Disponível em: https://revistas.ufpr.br/studiakantiana/article/view/91982. Acesso em: 31 jul. 2025.

KOHL, M. 2018. Kant's Critique of Instrumental Reason. *Pacific Philosophical Quarterly*, v. 99, p. 489–516.

KORSGAARD, C. M. 2008. *The Constitution of Agency*: Essays on Practical Reason and Moral Psychology. Oxford: Oxford University Press.

KORSGAARD, C. M. et al. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

MARTÍNEZ, L. 2023. The Kantian view of dark representations and their function in practical life, according to the anthropological notes of the Critical Period. *Studia Kantiana*, Curitiba, v. 21, n. 1, p. 49–59. DOI: 10.5380/sk.v21i1.91540. Disponível em: https://revistas.ufpr.br/studiakantiana/article/view/91540. Acesso em: 31 jul. 2025.

MERLE, J.-C. 2023. Action irrationality, systemic practical irrationality, and the remedy in Kant. *Studia Kantiana*, Curitiba, v. 21, n. 1, p. 9–18. DOI: 10.5380/sk.v21i1.91472. Disponível em: https://revistas.ufpr.br/studiakantiana/article/view/91472. Acesso em: 31 jul. 2025.

SCHOPENHAUER, A. 1960. *Essay on the Freedom of the Will*. Trans. Konstantin Kolenda. New York: The Liberal Arts Press.

TIMMERMAN, J. 2022. *Kant's Will at Crossroads*: An Essay on the Failings of Practical Rationality. New York: Oxford University Press.