Contents

Research Articles

Report

Organon F 31 (4) 2024: 338–366
https://doi.org/10.31577/orgf.2024.31401

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Kant on Moral Autocracy, Moral Faith and Happiness

Neşe Aksoy*

Received: 20 May 2024 / Accepted: 2 August 2024

Abstract: Kant's account of moral autocracy as the strength of will to impose inner self-constraint is usually interpreted as being strengthened and promoted by the pursuit of the obligatory ends of perfection and holiness. In this context, moral autocracy is seen as something that can be achieved through one's self-activity guided and encouraged by the ideals of perfection and holiness. In this paper, I argue that, in addition to the moral ends of perfection and holiness, moral autocracy also requires moral belief or faith as a guide to achieve a lifestyle that is peculiar to the highest good, i.e. a life conduct that is well-pleasing to God. In view of this, I argue that moral autocracy in conjunction with moral belief or faith leads to a morally pure and refined way of living directed towards the purity of the moral law. Finally, I conclude that the morally refined and elevated form of lifestyle achieved by morally autocratic activity opens the ground for happiness that does not only refer to the satisfaction of one's needs and desires but is also inclusive of the elements of enduring contentment, well-being, and bliss, which leads me to interpret Kant's account of happiness as encapsulating a eudaimonistic dimension in it.

* Bulgarian Academy of Sciences
https://orcid.org/0000-0002-3815-4350
The Institute of Philosophy and Sociology Bulgarian, Academy of Sciences, 13A Moskovska Str., 1000 Sofia, Bulgaria
aksoynesee@gmail.com

Keywords: Moral autocracy; moral faith; perfection; happiness; holiness.

Introduction

Kant's notion of moral autocracy[1] is usually understood as self-discipline and self-governance that can be achieved through the individual striving of agents to realize the moral goals of perfection and holiness. In this context, scholars such as Denis (2013, 2019), Stratton Lake (2005), Gregor (1963) and Allison (1990) aptly point out that striving for the realization of perfection and holiness plays a crucial role in strengthening one's autocratic activity in order to achieve conformity with the purity of the moral law. Denis (2013) points out that moral perfection plays an essential role in strengthening autocratic striving and gives agents direction, integrity, and meaning to their lives (Denis 2013, 179). Gregor (1963) also claims that while the duties of holiness and perfection are imperfect duties that entail some latitude in their realization, the pursuit of perfection and holiness is essential to the cultivation of one's resolute activity and virtuous disposition (Gregory 1963, 170). Allison (1990) also sees the cultivation of moral autocracy and strength of will to be related to the ideal of holiness and maintains that the resolute struggle of agents against inclinations and evil cannot be conceived independently of the pursuit of holiness as an indispensable ideal (Allison 1990, 171). While these accounts provide a comprehensive reading of moral autocracy to be based on the pursuit of perfection and holiness, what they essentially have in common and somewhat limitedly is that they view moral autocracy merely as a matter of ethics or morality by excluding religion from its purview. Indeed, this has a firm basis in Kant's notion of the highest good (*summum bonum*), which entails that religion should be included in the framework of ethics or morality under the

[1] I have used the translations of Cambridge editions of the works of Immanuel Kant that are abbreviated as follows: CPR: Critique of Pure Reason; G: Groundwork of the Metaphysics of Morals; CPrR: Critique of Practical Reason; MM: The Metaphysics of Morals; Rel: Religion within the Boundaries of Mere Reason; LE: Lectures on Ethics; LA: Lectures on Anthropology; Refl: Notes and Fragments, LPDR: Lectures on the Philosophical Doctrine of Religion, and O: Orientation.

questions "What may I hope for?" (*CPR*, A 805/B 833) and "What then is the result of our right conduct?" (*Rel*, 6:5). Kant's inclusion of moral belief or faith at the stage of the distribution of happiness thus suggests that religion does not have a practical significance in the context of the cultivation of morally autocratic activity to achieve the purity of the moral law, which is strictly associated with the principle that "ought implies can." However, as I will argue against the standard reading of moral autocracy, the morally autocratic or virtuous pursuit of perfection and holiness cannot be fully cultivated without appealing to moral belief or faith as an active disposition of agents. To this end, I will argue that while Kant seems to include moral belief or faith in the realm of ethics only in the context of the distribution of happiness for one's moral worthiness or virtue, he more fundamentally attributes an active role to moral belief or faith in the process of the cultivation of virtuous activity and moral struggle by constituting a model or moral ideal that encourages and guides agents to their moral ends (without violating the principle of "ought implies can").

To elucidate my argument, I follow three steps. First, I explain why I think that moral belief plays an indispensable role in the process of cultivating virtue and self-discipline by claiming that moral belief encourages and guides agents to impose inner constraint on themselves in order to achieve self-mastery and autonomy of their moral character. Second, I argue that moral belief or faith is an active disposition of agents that helps them to orient themselves towards their moral goals and enhance their virtuous activity and disposition, which, I will claim, opens up the possibility of a refined and morally pure way of life. Finally, the discussion of the moral purity and refinement that can be achieved through moral autocracy and the commitment to moral faith leads me to the claim that Kant's concept of happiness contains eudaimonistic elements, such that happiness for Kant is more than the maximal satisfaction of one's empirical needs and desires, which contains the elements of moral self-satisfaction, well-being, and bliss. This account will ultimately show that Kant's notion of the highest good involves a form of lifestyle that is gained by the autocratic struggle of agents seeking purity, perfection, and holiness and it cannot be conceived independently of moral satisfaction with a lifestyle that is characterized by purely refined and elevated elements.

1. Virtue, autocracy and obligatory ends

In the Introduction to the *Doctrine of Virtue* (*Tugendlehre*) Kant sets out by distinguishing between two types of duty, namely duties of right and duties of virtue, on the basis of the fact that the former relies on external constraint, while the latter depends upon internal self-constraint (*MM*, 6:383). This distinction is based on the fact that the duties of virtue are end duties, which cannot be coerced by external factors but can only be practised by the inner coercion of agents to make something an end by free choice (*MM*, 6:381). According to Kant, duties of virtue are exercised in the form of autocratic activity (*autarchia*) of imposing self-constraint on the self, since finite rational agents are always under the threat of inclinations and evil that leads them to violate the duty (*MM*, 6:383). Kant characterizes such autocratic activity as "the capacity to master one's inclinations when they rebel against the moral law" (*MM*, 6:383) and "the capacity to master oneself, to possess oneself, to be sufficient to oneself" (*LE*, 27:656). As the lecture notes on ethics make it clear, autocractic activity, which relies on one's inner strength and resolution, ultimately gives rise to self-mastery and self-governance in one's moral character (*LE*, 27:361). As Kant suggests, moral character, when not limited by moral discipline, creates nothing more than a "plaything" that has no stable and enduring resistance against inclinations and evil, but is dominated by them (*LE*, 27:362). However, when the character is governed by moral self-discipline and autocracy, it gains moral self-mastery and self-control by its ability to control sensibility under its rational principles (*LE*, 27:361). How, then, can we understand moral self-discipline and autocracy as a capacity that must be cultivated to establish the possibility of moral character that would elevate us to self-mastery and autonomy?

Kant's account of moral character (*Character*) relies on the exercise of agent's capacity of freedom. In lectures on anthropology, Kant is reported to have said, "Proper character is character of freedom" (*LA*, 25:1385). In the *Religion* he likewise argues that moral *Gesinnung* (moral disposition) or the fundamental maxim that defines the moral character of agent is freely chosen and therefore imputable (*Rel*, 6:25). By defining the moral *Gesinnung* as "the subjective ground of the adoption of the maxims" (*Rel*, 6:25),

Kant argues that moral evil or good can be imputed to agents when they choose certain maxims over the others on the basis of free-choice. Kant's characterization of moral character as a freely chosen maxim or way of thinking (*Denkungsart*) (*LA*, 25:649), which is an underlying and enduring disposition of agents, implies that it is not innate but must be cultivated and improved by labor and moral striving (*Rel*, 6:25; *LA,* 25:654).[2] On the other hand, the freely chosen and acquired aspect of moral character does not suggest that agents are loosely free to determine an evil or good character by their personal choice, but that they are obligated to cultivate a positive capacity of free choice by adhering to the moral law to create the possibility of forming a morally good character. Indeed, this has a firm basis in Kant's famous distinction between *Willkür* and *Wille* as the two reciprocal capacities of human volition. Thus, in the *Metaphysics of Morals* Kant claims that *Willkür* signifies a positive capacity and ability when it conforms to the *Wille*, whereas it signifies a misuse of one's freedom and incapacity when it deviates from the moral law (*MM*, 6:227). Likewise, Kant's account in the *Groundwork* suggests that free will, if not bound by "the lawful causality of immutable laws", is an absurdity (*Unding*) (*G*, 4:446).[3] In underscoring the idea that freedom of choice, if not based on the *Wille*, is a negative capacity, Kant affirms that deviation from the moral law reflects the misuse of one's freedom.[4] In parallel to his two-sense characterization of *Willkür* (the positive and negative capacity of *Willkür*), one can surmise

[2] For a comprehensive discussion on the dynamic structure of moral character that is not "a fixed and innate ground of moral change" but a progression towards a moral direction and end, see Coble (2003).

[3] For a helpful reading on Kant's notion of free will from the perspective of moral necessitarianism, see Ware (2023, 1–28).

[4] This account has been defended by Silber (1960), who has argued that the true and rational use of Willkür is to orient itself to the principle of the Wille (Silber 1960, cvi). Silber holds the rational agents have an original predisposition of the will to personality and they can realize it only by subordinating their Willkür to the Wille (Silber 1960, civ, cv). Thus, Silber suggests that the heteronomous choice of the Willkür is a misuse of one's freedom. Henry Allison, on the other hand, argues that the human Willkür has the capacity to deviate from the moral law or obey the Wille, but this capacity of the deviation of the moral law merely refers to a misuse of freedom (Allison 1990, 135–36).

that Kant develops a similar approach in his theory of virtue in constructing a morally good character. Accordingly, while he identifies virtue as an autocratic and courageous activity (*fortitudo moralis*), denoting the willkürlich and meritorious capacity of agents to impose self-constraint on themselves (*MM*, 6:380), he affirms that the meritorious quality of virtue does in no way entail that agents are loosely free to pursue it or not. In this context, while he recognizes virtue as an ideal that can never be fully attained, he firmly affirms that the constant striving for approximation to virtue is a real duty (*MM*, 6:406). He further asserts that, despite its meritorious quality, the pursuit of virtue has an inner worth in itself when it is sought for its own sake (*MM*, 6:409). In light of the Kantian notion of autonomy, it is easy to see that, despite the meritorious and willkürlich aspect of virtue, the true exercise of virtue, which improves the autonomous activity of agents, consists in striving for it in a rigorous and resolute manner. Within the framework of his notion of moral character, Kant emphasizes the vitality of such virtuous striving, by stating: "Character is that which marks a resolution in principles in the human being" (*LA*, 25:1169). In this sense, Kant holds that moral character has an inner value in itself to the extent that agents freely choose to strive for virtuous disposition by their all powers and resolution that signifies the true use of their freedom.

Kant's advocacy of autocratic and resolute activity of virtue as the fundamental element in determining moral character is closely related to the duties of seeking one's own moral perfection and holiness. Kant defines holiness as "the purity of one's disposition to duty", which refers to acting out of pure respect for the moral law (*MM*, 6:446). Moral perfection, on the other hand, is defined as "fulfilling all one's duties and attaining completely one's moral end with regard to oneself" (*MM*, 6:446). The duties of holiness and moral perfection, then, appear to be one's task to cultivate virtue to attain the final end of his vocation, that is, the complete fulfilment of his moral duties and ends. In his lecture course on ethics, Kant is recorded as affirming that the duties of virtue are not strictly obligatory but to be cultivated by the free choice of agents that desire to pursue the end of moral perfection and holiness as the ultimate end of their vocation (*LE*, 25:579). However, he notes that to determine the extent to which agents desire to seek moral perfection and holiness is up to them, since it is a wide duty

that is not strictly obligatory but to be pursued to the degree possible (*LE*, 27:582). In these lecture course notes, then, moral perfection and holiness are seen as meritorious duties that depend on the execution of the strength of will and autocratic activity of agents to the degree that they can. However, as the *Metaphysics of Morals* makes clear, Kant describes moral perfection and holiness as highly exceptional in comparison to the other duties of virtue. According to his account, both moral perfection and holiness are unfathomable ideals that can never be attained by agents but must strictly be promoted in an ongoing progress (*MM*, 6:447). In this context, moral perfection (perfection and holiness) is seen as "a narrow and perfect duty in terms of its quality" but "a wide and imperfect one in terms of its degree because of the frailty (*fragilitas*) of human nature" (*MM*, 6:447). In Kant's view, then, moral perfection and holiness are the duties that must be strictly pursued by agents but there is latitude in their fulfilment because human agents have a frailty by nature that prevents them from committing themselves fully to these duties.

The fact that Kant refers to the duties of moral perfection and holiness as necessary and obligatory ends in determining one's moral purpose and final vocation suggests clearly that they play an indispensable role in strengthening the moral resolution and strength of will of agents in their pursuit of morally good character. As such, scholars such as Denis (2013, 2019), Stratton Lake (2005)[5], Gregor (1963), and Allison (1990) contend that agents' virtuous disposition to self-impose autocracy in order to realize the full potential of their autonomy and moral character cannot be conceived independently of the pursuit of the obligatory ends, as they serve as a guide and ideal for agents in their pursuit of the adherence to the moral law.

Lara Denis (2013), for instance, argues that obligatory or objective ends strengthen inner freedom and enhance agents' motivation to battle evil and

[5] Phillip Stratton Lake (2005) argues that when we strive to make ourselves virtuous or struggle to make the moral law as the sole incentive of our actions, this requires not only respect for the moral law as the sufficient ground of our action, but also that we make our perfection as an end (Stratton Lake 2005, 105–6). Stratton thus suggests that we cannot cultivate virtuous disposition as resolution and strength (striving to become virtuous) without appealing to perfection as an end.

inclinations (Denis 2013, 178). In her view, unconditionally necessary ends give direction, integrity and meaning to our lives toward the realization of the highest good as the highest end of practical reason (Denis 2013, 179). Denis further writes that moral perfection is an end that aids agents to realize their full potential of autonomy and personality and it is a necessary end that must be embraced in order to enhance the dignity of character and moral agency, despite the human frailties that are faced in realizing it (Denis 1999, 24).

Mary Gregor (1963) also argues that while the duties of holiness and perfection are imperfect duties that entail certain latitude in their performance, the pursuit of perfection and holiness are vital in the cultivation of virtue (Gregor 1963, 170). According to Gregor, Kant's notion of the moral life as the cultivation of virtue cannot be achieved independently of the end of moral perfection, since moral perfection should be viewed as the concrete capacity that must be developed through the maxims of agents in the process of forming a moral character (Gregor 1963, 171).

Henry Allison (1990) sees the cultivation of moral autocracy and strength of will in connection with the ideal of holiness, claiming that the resolute struggle of agents to battle inclinations and evil cannot be conceived independently of seeking holiness as an indispensable ideal (Allison 1990, 171). Allison firmly affirms that to fail to adopt the maxim of the pursuit of holiness literally means to remain outside of the moral struggle (or autocracy) against evil, which deprives one's character of moral worth (Allison 1990, 179). Thus, Allison's account aptly suggests that moral perfection and holiness play an essential role in strengthening one's resolution to control inclinations and evil and that the duties of moral perfection and holiness have an impact on the cultivation of one's character, moral life and orienting agents as ethical beings (Allison 1990, 178).

The accounts of Denis (1999, 2013), Gregor (1963) and Allison (1990) offer a comprehensive reading of the obligatory ends of perfection and holiness as the necessary ideals that help agents improve their moral strength and incentive to pursue the purity of the moral law. They rightly point out that virtue as an autocratic activity of agents would remain inefficient unless it is provided with a practical guidance and orientation by the moral ends and ideals of perfection and holiness. What these accounts essentailly

have in common is that they regard moral autocracy or moral strength as a capacity to be developed within the framework of morality or ethics, and exclude religion from its purview. Without a doubt, this is the standard reading of moral autocracy in the literature, which is consistent with the notion that Kant includes religion and theology in the realm of ethics or morality merely within the context of the highest good (*summum bonum*), under the questions "What may I hope?" (*CPR*, A 805/B 833) and "What is then the result of this right conduct of ours?" (*Rel*, 6:5). According to the usual reading, then, religion or theology has no practical significance within the scope of the cultivation or acquisition of virtue (as it would suggest the violation of the principle of "ought implies can"),[6] but enters the scene as part of the distribution of happiness through divine help in rewarding one's virtue and moral worthiness. In his discussion of the highest good (*summum bonum*), Kant seems to advocate the idea that virtuous activity has to be fulfilled for its own sake with human powers ("ought implies can") and the appeal to moral belief in divine help takes on practical significance especially for the distribution of happiness (*CPrR*, 5:124). However, a close examination of Kant's theory of the highest good tells us that moral faith plays a practical role not only in obtaining divine help for the distribution of happiness, but also in strengthening agents' virtuous activity with respect to the moral good. In parallel, Allen Wood (1970) points out

[6] Regarding this basic problem within the context of Kant's notion of highest good, Silber (1959) aims to offer a solution by his two-sense theory of the highest good, namely the immanent and transcendent conceptions of the highest good (Silber 1959, 469–86). As such, Silber argues that moral agents strive for moral worthiness within the "ought implies can" principle in an immanent sense, but there is a transcendent aspect to it that constitutes an ideal for practical reason. As he further suggests, the transcendent ideal of the highest good (including the notions of holiness, and moral belief) plays a regulative role in realizing one's duty to the highest good on an immanent level. Likewise, Pasternack (2017) argues that virtue is a capacity to be cultivated and realized by human capacity ("ought implies can") and the divine assistance of happiness is an aid that is not incompatible with the autonomy of agents (Pasternack 2017, 461–64). According to his account, divine assistance for the distribution of happiness does not clash with the autonomy of agents because agents are merely provided with the element of happiness as an encouragement that does not interrupt their individual struggle and resolution to achieve this end by their own powers.

that the endorsement of moral belief or hope merely in the context of the divine help for receiving happiness in proportion to one's moral worthiness is a deeply mistaken idea (Wood 1970, 166). In a helpful passage, Wood underscores that Kant's inclusion of God "in the ends of morality" as the distributor of happiness leads to the misunderstanding that moral belief must be cultivated for the purpose of obtaining happiness, stating:

> Moral faith does not consist so much of an expectation of future happiness as an acceptance of present sufferings. Moral faith does not promise me a world better than the actual one, but consists in the courage to trust that this world, as it is, a world in which there is suffering and apparent moral failure, is itself a morally good world. In moral faith, I do not flee this world to a better one; rather, I *choose* this world, I refuse to despair of it, I make it by my choice the world in which I will rationally act in pursuit of my final moral end. Trust in God is thus not characterized by the gleeful anticipation of my "future rewards," but rather by "humility and modesty, combined with resignation" (Wood 1970, 169).

In this excerpt, Wood makes clear that the primary purpose of moral belief or hope cannot be understood as the expectation, belief or hope for happiness as a reward of our virtuous conduct, but rather that it is to cultivate a firm trust in the real possibility of the existence of God such that it encourages agents to overcome moral despair and motivates them to struggle rigorously to make their life and world morally better. Parallel to Wood's assertion that moral belief or faith plays a profoundly significant role in encouraging agents to pursue and strive for the moral good meticulously and overcome moral despair, my contention is that moral belief or faith is an active and dynamic disposition to encourage agents to strive for virtue and autonomy, not simply because of the encouragement it provides via the distribution of happiness, but primarily because it gives encouragement to agents by directing them to seek virtue or moral worthiness for moral improvement as an end in itself. In contrast to the usual reading of moral autocracy that is conceived as a capacity to be cultivated by human powers through the model and ideal of obligatory ends, the next section will thus focus on the fact that the autocratic activity of agents towards perfection and holiness in the process of gaining

moral worthiness (or the highest good) also requires moral belief or faith as an active disposition of themselves.

2. Moral faith, virtue and moral perfection

Kant's theory of moral belief and faith (*moralischer Glaube*) is formulated as a moral postulate that is practically necesary for the real possibility of the highest good (*summum bonum*). In the second *Critique* Kant defines it as a subjective need (*Bedürfnis*) that is "not commanded as a duty" but is based on a voluntary judgment of agents that is "conducive to the moral purposes" (*CPR*, 5:146; O, 8:141.). Likewise, in the first *Critique* and the *Orientation* article Kant defines moral belief and faith (*Glaube*) as a form of conviction that is "objectively insufficient but subjectively sufficient" (*CPR,* A 822/B 850), and that stands between the epistemic assents of opinion (*Meinung*) and knowledge (*Wissen*). Kant's notion of moral belief or faith is predominatly examined as an indispensable theological element of the highest good, especially in terms of the role it plays in grounding the practical intellligibility and consistency of morality.[7] Yet, the practical significance of moral belief or faith as an active element of one's moral disposition in cultivating virtue or moral worthiness is severely neglected in the

[7] In this regard, Wood (1970) stresses that moral belief has a pivotal role in establishing the highest good and thereby avoiding the state of *absurdum practicum* or the practical absurdity of moral activity (Wood 1970, 26–27 and 29–30). Förster (2000) has drawn attention to this aspect of Kant's discussion of the moral-practical idea of God as a binding element in human morality, by remarking that God is analytically connected to the categorical imperative (Förster 2000, 142) and our freedom as a person demands that we relate ourselves to the binding nature of the categorical imperative that is ultimately grounded on the moral-practical idea of God (Förster 2000, 141). Palmquist (2000) likewise holds that moral postulates are indispensable for morality, as their absence would lead morality to the state of "abyss of meaninglessness" (Palmquist 2000, 75). Along similar lines, Zeldin (1971) endorses the theological approach, holding that moral belief is strictly necessary for morality, as it prevents the command of the moral law from becoming a mere "law of thought" (Zeldin 1971, 47).

literature.[8] Kant's account of moral belief or faith, however, would remain very incomplete if its active content, which helps agents to orient their virtuous activity and struggle toward moral perfection and holiness, is overlooked. In parallel, in this section, I want to focus on moral belief or faith as an integral element of agents' moral disposition that encourages and promotes their autocratic striving for virtue or moral worthiness in an authentic and real experience that helps them to cultivate morally pure and refined life conduct.

The fact that Kant ascribes practical significance to moral belief and faith as an active disposition of human volition that aids agents to improve their virtuous struggle for moral perfection and holiness has a firm basis in Kant's discussion of moral belief and faith throughout his corpus. In the first *Critique*, Kant articulates this radical aspect of moral belief and faith by contrasting it with doctrinal belief in God, asserting that the moral certainty that moral belief and faith entails is highly different from the relatively unstable content of the doctrinal belief (theoretical and hypothetical belief in the existence of God) (*CPR*, A 828/B 856). Kant firmly affirms that moral belief or faith entails an "absolutely necessary" and "inescapably fixed" conviction that there must be a God and a future life, provided that agents abide by the moral law (*CPR*, A 828/B 856). This tells us that moral belief or faith has a stable aspect, so that it is impossible to renounce it, as this would mean overturning all our moral principles (*CPR*, A 828/B 856). This is to say that moral belief or faith in the existence of God and a future world is inescapably vital for the orientation and foundation of one's moral activity, the absence of which would result in the lack of meaning, purpose or intelligibility of one's moral disposition and life (*CPR*, A 828/B 856). If we read this in light of Kant's famous claim in the Lectures on Philosophical Doctrine of Religion, in which he argues that the absence of assent to moral belief or faith leads to the state of *absurdum practicum* (*LPDR*, 28:1083), we can easily see that moral theology, or the assumption of God as part of morality, leads agents to become "better human beings" and that its denial leads to the state of *absurdum practicum* and reduces agents into "scoundrels" (*LPDR,* 28:1083).

[8] For a helpful reading that endorses the idea that moral belief or faith is an active disposition of one's moral disposition, see Wood (1970, 160–62).

At other times, Kant appears to reiterate the idea that moral belief or faith plays an active role in grounding the moral activity of agents. The account of moral belief in the second *Critique* is based on the idea that moral belief is a subjective need (*Bedürfnis*) that arises from the moral law for the ultimate realization of the highest good (*CPrR*, 5:146). Kant further notes that moral belief or faith is practically necessary for promoting the morally commanded ends (*CPrR*, 5:146). Accordingly, he emphasizes the indispensable significance of moral belief or faith in human life, noting that it arises from one's moral disposition and that it is therefore never possible to renounce it altogether (*CPrR*, 5:146). Thus, Kant writes that respect for the moral law leads agents to presuppose the existence of God, which ultimately improves their moral disposition and moral activity. As Kant points out, without such a genuine experience of the true respect for the moral law and the moral presupposition of God, human conduct would simply resemble "a puppet show" in which "everything would gesticulate well but there would be no life in the figures" (*CPrR*, 5:147). For Kant, then, the moral law when accompanied with moral belief or faith leads to a genuine and authentic experience in the moral life that elevates the virtuous disposition.

In the third *Critique* Kant spells out the practical significance of moral belief or faith as an active element of one's moral disposition by defining it as "habitus or reason's moral way of thinking in the affirmation of that which is inaccessible for theoretical cognition" (*CPJ*, 5:471). In this sense, he emphasizes the active role of moral faith in God through his notion of moral teleology, which suggests that the supersensible agency of God enters the scene of teleological paradigm of nature as the ground that connects the causality of the human agent as noumenon (freedom) to his final ends (*CPJ,* 5:448). By comparing physical teleology with moral teleology, Kant claims that the moral idea of God, which belongs to moral teleology, grounds the inner lawfulness of the causality of nature and the causality of the moral agents as noumenon. This, then, suggests that the moral idea of God as the ultimate ground for connecting human agents to their moral ends, plays an active role in enabling agents to ground their practical vocation as moral beings. In this sense, Kant's endorsement of the moral idea of God as the ultimate ground of synthesis of the teleological purposes of nature (in a

heuristic sense) shows that it has an immanent role in directing human life towards its ultimate purposes.

The main textual evidence for this reading, that moral faith has an active role in grounding one's virtuous activity, lies in Kant's account of moral faith in the *Religion within the Boundaries of Mere Reason*, where he describes moral faith as "an active disposition toward achieving a good-life conduct" (*Rel*, 6:116). In the third part of the *Religion*, Kant fundamentally proposes that pure religious faith is part of a person's moral disposition (or *Gesinnung*)[9] that guides him into a morally good way of life. This is primarily based on the idea that moral faith is an active disposition that guides and promotes the enhancement of one's virtuous disposition as part of the realisation of a good life conduct. He writes:

> The living faith in the prototype of a humanity well-pleasing to God (the Son of God) refers, in itself, to a moral idea of reason, insofar as the latter serves for us not only as guideline but as incentive as well; it is therefore, all the same whether I start out from it (as rational faith) or from the principle of a good life conduct (*Rel*, 6:119).

In this passage, Kant portrays moral faith as equal to the pursuit of a good life conduct that is well-pleasing to God within an authentic experience.[10] He thus emphasizes that moral belief is more than an abstract idea, but that has an immanent role in determining one's moral experience. Accordingly, Kant further notes that, "the disposition of virtue has something to do with *actual*, which is in itself well-pleasing to God and conforms to what is best for the world" (*Rel*, 6:173). In this remark, Kant also suggests that the main purpose of virtuous disposition is to struggle to establish harmony with the moral idea of God in an actual experience. In this context, what

[9] For a parallel reading, see Munzel (1999, 205–6); Wood (1970, 156); and Palmquist (2015, 247–48).

[10] Akin to my approach, Palmquist (2015) argues that moral belief as a pursuit of "lifestyle that is satisfactory to God" is not a psychological attitude or metaphysical disposition but a volition that is embdeded in one's *Gesinnung* (Palmquist 2015, 247–48). Likewise, Wood (1970) holds the view that moral belief or faith or "personal trust in God" is an immanent element of human volition and moral *Gesinnung* (Wood 1970, 160–62).

is remarkable to note is that moral faith has an active aspect in aiding agents to improve their virtuous activity toward the purity of the moral law, namely, perfection and holiness, when agents create a moral world for themselves that conforms to the ideals of holiness, beneficence and justice of God (*LPDR*, 28:1076). This is essentially evident in Kant's notion of moral theology, which entails that the moral idea of God has a determinate character in itself insofar as it is understood as an ideal that is to be pursued in one's moral activity (*LPDR*, 28:1073). While Kant assumes that the moral idea of God is a mystery in itself, it has no mystery if we conceive it as a moral ideal and model for our moral purposes (*Rel*, 6:137-140). In parallel, he suggests that the cultivation of moral faith does not help us to know the real essence of God in itself, which is a mystery, but has practical implication, in the context of our morality, that we should aspire to the moral idea of God as the highest form of perfection (*Rel*, 6:139). Accordingly, in this picture Kant paints, moral faith in God and immortality seems to play an immanent role in our moral experience insofar as we cultivate it in our moral struggle for perfection and holiness.

The fact that holiness and perfection encompasses the pursuit of a life conduct that is well-pleasing to God suggests that it opens up a lifestyle that is characterized by morally pure and refined elements that are reinforced by the autocratic activity. Thus, in the Lectures on Philosophical Doctrine of Religion, Kant writes: "*Holiness* is the absolute or unlimited moral perfection of the will… Thus the human being can *never* be *holy, but of course* [he can be] *virtuous.* For virtue consists precisely in *self-overcoming*" (*LPDR*, 28:1076). In this passage, we can easily see that the pursuit of holiness refers to the state of constant struggle to control the natural inclinations and evil in one's heteronomous nature and conform to the purity of the moral law. In this context, Kant identifies virtue as a never-ending process of achieving victory over one's immoral inclinations and evil and pursuing the state of holiness as conformity to the purity of the moral law. The pursuit of holiness (or self-overcoming or self-conquest) essentially means that agents gain proper rational control over their inclinations and evil through their radical struggle towards perfection and holiness.[11] The moral struggle towards perfection and

[11] Regarding Kant's notion of moral self-overcoming or self-conquest, Baxley (2010) has critically argued that the notion of self-overcoming involves the pursuit of a

holiness thus involves a radical transformation in agents, which gives rise to a morally pure and refined way of living that conforms to the purity of the moral law. In a similar fashion, Kant's account of "change of heart" as the radical transformation of agents from an evil or bad disposition to a good disposition involves a radical change in one's virtuous disposition that affects one's moral *Gesinnung* or way of thinking (*Denkungsart*) (*Rel*, 6:47). By this, Kant means that the radical transformation to the purity and holiness of the moral law cannot be understood as a habitual or gradual development towards the realization of the moral good (*virtus phenomenon*), but involves an endless struggle to be well-pleasing to God through the revolutionary change in one's virtuous disposition (*virtus noumenon* or the intelligible character of virtue). Kant's formulation of *virtus noumenon* is thus something highly transformative that affects the moral *Gesinnung* or character of agents, leading them to rise to the sublimity of perfection and holiness in a constant struggle. The fact that Kant refers to agents who undergo a revolutionary transformation from the bad disposition to the good under the terms of "new man" or "change of heart" suggests that virtuous activity from evil toward the predisposition of the good entails a noumenal and atemporal change in a person's disposition or character[12] that elevates him to a morally sublime and pure state (*Rel*, 6:47).

"disembodied God" that suggests the extirpation and suppression of sensible inclinations and desires altogether (Baxley 2010, 65). Baxley affirms that this notion of self-overcoming leads to the standard idea that Kant's theory of virtue entails constant tension between sensibility and reason and constant battle with the self (Baxley 2010, 67). However, as she notes, moral self-overcoming, in Kant's view, can be better understood as a rational self-governance that does not necessarily mean to fight sensibility or our sensible self, but more fundamentally to cultivate a rational mastery over our sensible inclinationas and desires in accordance with our moral principles (Baxley 2010, 75–79). In doing so, Baxley contends that Kant's notion of virtue as the pursuit of perfection and holiness has an affirmative aspect in its scope, as it explicitly denies a "harsh and dehumanizing" form of self-discipline or autocracy, which suggests that the deprivation of all enjoyment and pleasure of life through self-discipline is not a proper form of virtue (Baxley 2010, 76).

[12] On the timeless and atemporal aspect of moral character (or *Gesinnung*), see Allison (1990, 140–141); Wood (1984, 97–99); Hill (2014, 91–93).

Moreover, Kant's account of moral perfection and holiness has a radical aspect to it, for it entails that agents emulate a model or ideal of human being that autocratically struggles to perfect himself, which leads them to create a morally pure and refined life conduct. Accordingly, Kant argues in the *Religion* that the pure conformity to the moral law or the duty to rise to moral perfection and holiness, involves emulating "the prototype of human being that is well-pleasing to God", which is characterized as an ideal and model of human being that the practical reason has within itself (*Rel,* 6:62). According to this account, the pursuit of holiness and moral perfection is based on the ideal of human being that pursues moral perfection and holiness through an autocratic struggle to overcome evil and unruly inclinations and comply with the purity and holiness of the moral law. Kant depicts the autocratic agent struggling to conform to the purity and holiness of the moral law as

> willing not only execute in person all human duties, and at the same time to spread goodness about him as far wide as possible through teaching and example... And surrounded by obstacles and yet—in the midst of the greatest possible temptations- victorious" (*Rel*, 6:62).

Kant's notion of the ideal autocratic agent struggling for moral perfection and holiness provides a guide and model for ordinary human agents to strengthen their virtuous disposition towards a refined and morally pure state of moral existence. Moral autocracy, then, entails that agents cultivate their moral character and lifestyle to an elevated state by exposing themselves to a rigorous struggle to overcome their sensible and heteronomous nature and conform to the purity of the moral law. At first glance, the pursuit of perfection and holiness suggests an ascetic lifestyle characterized by rigorous striving and struggling for moral improvement. However, Kant's notion of virtue is not compatible with the notion of ascetism in a robust sense, as Kant clearly affirms that ascetism without joyful and cheerful elements is an inaccurate type of ascetism, which is known to represent "monkish ascetism" (*MM*, 6:485).[13] Thus, Gregor (1963) argues that Kant's

[13] In this sense, in the *Doctrine of Virtue* Kant argues for a moderate form of moral ascetism that is characterized with cheerfulness and valiance as an outcome of one's

notion of moral perfection does not suggest a strict and harsh form of self-discipline aiming to curb inclinations or sensible desires altogether (Gregor 1963, 171). Instead, Gregor argues that Kant's theory of virtue is opposed to moral purism per se, for it entails that inclinations and evil should only be marshalled and curbed to the degree that one can avoid vices and evil (Gregor 1963, 172). In this regard, Gregor argues that moral perfection does not refer to the "the habit of ascetism" (Gregor 1963, 172). Akin to Gregor, my contention is that Kant's notion of moral autocracy does not entail ascetism in an ordinary sense. However, unlike Gregor, I believe that the autocratic pursuit of the purity of the moral law inevitably contains ascetic elements in the moderate sense, to suggest that moral lifestyle that is well-pleasing to God aims to overcome a robust and uncultivated form of lifestyle by establishing refinement and simplicity in itself.

In sum, we can now see that morally autocractic activity when accompanied with moral belief or faith as an active disposition in guiding agents towards a better virtuous dispositon results in an authentic lifestyle characterized by refined and morally pure elements. It has thus become clear that the pursuit of moral perfection and holiness with the aid of moral faith opens up a life conduct that raises one's moral state and existence towards a more sublime and refined form of experience. This means that the pursuit of moral perfection and holiness requires agents to overcome the sensible inclinations and desires that limit their sphere of life, which is burdened by robust and unruly elements. On the contrary, the autocratic pursuit of perfection and holiness enables them to cultivate an authentic experience of moral purity that elevates them to the sublimity of their moral personality and existence. On this basis, in the next section, I will focus on the way in which agents' virtuous struggle of for perfection and holiness through their commitment to moral belief leads to a lifestyle that contains morally pure

moral struggle for becoming virtuous and morally worthy. Comparing it with monkish ascetism, Kant affirms that virtuous lifestyle does not necessarily mean the eradication of inclinations or desires through the strict punishments and brutal practises over the self, but rather battling the unruly and non-moral inclinations to the degree that we gain mastery over them which "makes one valiant and cheerful in the consciousness of one's restored freedom" (*MM*, 6:485).

and refined elements, which will lead me to interpret Kant's theory of happiness *(Glückseligkeit)* as containing eudaimonistic elements related to such life conduct.

3. Virtue, moral faith and happiness

My considerations so far have shown that moral autocracy to impose inner self-constraint to comply with the purity and holiness of the moral law requires the element of moral belief and faith in order to give encouragement and order to our moral activity within an authentic experience. Clearly, this way of life conduct, which strives for the perfection and holiness of the moral law through the ideal image of God as a moral model (or moral worthiness), is characterized by moral refinement, purity and sublimity. According to the classical theories of virtue, such a virtuous lifestyle is accompanied by a form of happiness that includes a refined form of satisfaction with one's disposition and state. Nevertheless, on a usual reading, Kant's theory of happiness *(Glückseligkeit)* as a reward for virtuous activity or moral worthiness does not favour such a lifesyle that can be associated with a eudaimonistic kind of happiness. Instead, Kant's notion of the highest good encompasses the idea that morally worthy agents can hope for the possibility of a happiness that grants them with "satisfaction of one's all inclinations" (*CPR*, A 806/B 834) or the state in which "everything goes according to their wish and will" (*CPrR*, 5:124). Such definitions of happiness, which emphasize the maximal satisfaction of one's needs and desires, lead some scholars to interpret Kant's notion of happiness as contradictory to the eudaimonistic concept of happiness, since it contains a pleasure-based content.[14] Contrary to such accounts, in this section I will argue that Kant's

[14] Wood (2000) argues that Kantian notion of human nature is starkly different from the eudaimonistic notion of human nature in the sense that happiness cannot be simply explained through the activity of virtue. Instead, Wood holds that the Kantian notion of human nature has a natural-social aspect as well as its moral aspect, which leads him to suggest that happiness has as a pleasure and desire-directed aspect which cannot be identified in the classical eudaimonist accounts (Wood 2000, 265–67). See also Louden (2015, 118–20) for a non-eudaimonist reading of Kant's notion of happiness.

notion of happiness is not antithetical to eudaimonism,[15] since, as I will suggest, the full picture of happiness is not confined to satisfaction of one's empirical needs and wishes, but also includes the eudaimonistic elements of enduring well-being, contentment and bliss to be pursued as an ideal. This account will ultimately indicate that within the framework of the highest good, in which agents strive for purity and refinement through the pursuit of perfection and holiness, they could achieve a form of happiness that is much deeper and refined than the mere satisfaction of their empirical needs and desires.

Despite Kant's sharp criticism to ancient eudaimonism for violating the rule of autonomy of reason by subordinating morality to pleasure and happiness[16], there are some elements in his account of happiness that bears

[15] For similar accounts that identify eudaimonistic elements in Kant's theory of happiness see Holberg (2018); Grenberg (2022); Elzondo (2023); Engstrom (1996).

[16] In the *Metaphysics of Morals* Kant makes it clear that eudaemonism, which assumes pleasure or happiness as the guiding principle for one's moral actions leads to the problem of heteronomy that is a fatal mistake in morality (*MM*, 6:378). Kant's denouncement of eudaemonism as a tragically mistaken approach in ethics, therefore, relies on the idea that the eudaimonists ground morality on the principle of happiness as the ultimate end of humanity. Nevertheless, as Irwin (1996) suggests, Kant's assumption that eudaemonism reduces morality to the pursuit of happiness is problematic in itself (Irwin 1996, 81). Irwin argues that the eudaimonists regard happiness as the end of morality not in the sense that morality should be subordinated to pleasure but rather in the sense that morality should be purposeful towards the realization of the highest good (*summum bonum*) (Irwin 1996, 83). In this context, Irwin suggests that happiness, as it is understood by the ancient eudaimonists (specifically Aristotle and the Stoics), does not necessarily refer to pleasure, but in fact refers to an enduring state of contentment and peace of mind that is achieved through virtue (*summum bonum*). Irwin's point suggests that eudaimonistic approach in ethics is not a failure due to the problem of heteronomy, since it grounds morality not on the principle of pleasure or happiness, but rather on the purpose of achieving the state of eudaimonia that relies on virtuous activity (*summum bonum*). In fact, this eudaimonistic approach is not entirely excluded from Kant's deontological ethics, since Kant never denounces the idea that happiness can be pursued by agents on a rational and moral ground, thereby suggesting that morality is not completely free from the pursuit of happiness insofar it rests upon moral worthiness.

affinity with eaudaimonism, especially with regard to the eudaimonistic elements he attributes to the content of happiness. As is well known, according to the standard account of eudaimonism, happiness means a pure and elevated state of contentment that is largely free from unruly inclinations and desires. Thus Aristotle describes eudaimonia as a pure and stable form of pleasure and contentment achieved through virtue (*NE*, 1177a). In contrast to the empirical and sensible content of Kantian happiness, Aristotle's eudaimonia refers to the state of the highest good (*summum bonum*), which is characterized by enduring contentment and pleasure and is achieved through a morally pure and elevated lifestyle. In contrast, Kant argues that contentment with one's person and freedom cannot be regarded as a form of happiness, but as something that is akin to *beatitudo*, which denotes self-sufficieny that only belongs to God (*CprR*, 5:118). In this context, Kant seems to recognize self-contentment as a form of moral satisfaction that lies between happiness and *beatitudo*.[17] Nevertheless, this distinction between self-contentment and happiness does not necessarily mean that Kant excludes self-contentment with one's own person and freedom from the purview of his concept of happiness altogether.[18] This is evidently clear from his following remarks:

Another pivotal point that Kant is critical of eudaemonism is that the ancient eudaimonists (Epicurus and Zeno) recognize the highest good as the supreme principle of morality (*LE*, 27:248–50). According to Kant, this assumption violates the principle of autonomy of reason that is detrimental to ethics. Nevertheless, Kant's criticism of ancient eudaemonism for violating the rule of autonomy of reason is not easily applicable to the Aristotelian eudaemonism, since within his discussion of the problem of heteronomy in the Epicurean and Stoic notion of the highest good and happiness, Kant does not mention Aristotle's notion of eudaimonia as a fallacious doctrine (Louden 2015, 112–14). Therefore, one could suggest that Kant's notion of happiness is not in direct opposition to Aristotelian notion of eudaimonia.

[17] This point is also made by Walschots (2017), who holds that self-contentment is a form of moral satisfaction that is a "middle place" between happiness and bliss (Walschots 2017, 287).

[18] Despite his non-eudaimonist reading of Kant's notion of happiness, Wood (2000) does not exclude contentment from the scope of happiness by holding that complete happiness has three different aspects, namely "pleasure, contentment with one's state and desire-satisfaction" (Wood 2000, 267).

> Contentment with our entire existence is happiness; among human beings, this also requires physical causes, i.e., welfare (*Refl*, 18:460). Happiness is really not the greatest sum of gratifications but the pleasure of being satisfied by the consciousness of one's own power, at least this is the essential formal condition of happiness, although still other material conditions (as with experience) are required (*Refl*, 19:277).[19]

In these remarks, Kant is concerned with the idea that contentment with one's freedom and person is a form of happiness insofar as it is accompanied with physical contentment or welfare with one's state. In the first place, he assumes that happiness involves physical and moral contentment with one's whole existence. The second remark suggests that happiness as contentment with one's own power denotes a form of satisfaction that does not necessarily mean the maximal gratification of one's needs and desires, but rather a moral satisfaction that is accompanied by the physical satisfaction of one's state in a moderate manner. Likewise, Kant's famous assertion in the *Groundwork* that happiness is an indirect duty to be pursued as "a form of contentment with one's condition", insofar as it denotes the satisfaction of one's needs and desires to avoid transgression of the moral law, clearly implies that happiness for Kant does not simply signify maximal satisfaction of one's needs and desires, but refers to the satisfaction of one's needs and desires to the extent that it sustains the physical and moral well-being of agents (*G*, 4:399). I guess Kant's inclusion of contentment within his notion of happiness through the moderate form of physical satisfaction is consistent with his account of the moral purity and elevation that the pursuit of the perfection and holiness leads agents into. Such a reading of happiness reveals that Kant's account of happiness does not involve a robust form of physical satisfaction of one's needs and desires, but, more fundamentally, a refined form of satisfaction with one's state that arises through the morally autocratic pursuit of perfection and holiness.

Associating Kant's notion of happiness with a pure and refined way of life conduct guided and controlled with moral principles and laws leads to the idea that happiness refers to a state of existence that is characterized

[19] See also *LE,* 27:648–50.

by enduring contentment and well-being.[20] Accordingly, in the Lecture notes on Philosophical Doctrine of Religion, Kant characterizes happiness as an uninterrupted inner pleasure and well-being that helps us to overcome the hardships and troubles of life:

> Human good fortune is *not a possession, but a progression toward happiness.* Yet full self-contentment, the consoling consciousness of rectitude, is a good which can never be stolen from us, whatever the quality of our external state may be. And in fact all earthly happiness is far outweighed by the thought that as morally good human beings we have made ourselves worthy of an uninterrupted future happiness. Of course this inner pleasure in our own person can never compensate for the loss of an externally happy state, but it can still uplift us even in the most troubled life when it is combined with the prospect for the future (*LPDR*, 28:1090).

This passage clearly shows that Kant's concept of happiness has a eudaimonistic aspect, for it states that happiness refers to an inner pleasure that has an enduring and stable quality in the face of the troubles and misfortunes of life. Kant believes that while such inner contentment cannot make up for the lack of earthly happiness, it is a primary aspect of happiness that raises one's spirit as he faces and combats the hardships of life. Clearly, this eudaimonistic aspect of Kant's notion of happiness has a strong continuity with his notion of an autocratic lifestyle characterized by morally pure and refined elements. Therefore, it seems plausible that the moral pursuit of purity and refinement through virtuous endeavour opens up the possibility of happiness, which enables them to withstand the hardships of life with inner strength and power.

Another eudaimonistic aspect that we can attribute to Kant's notion of happiness is that he does not exclude the element of divine blessedness and bliss as a form of moral satisfaction as an ideal to be pursued through the pursuit of perfection anf holiness. In the *Nicomachean Ethics*, Aristotle's account of happiness as a life of wisdom, contemplation and noble actions

[20] For a eudaimonist reading of well-being by identifying it as the worth of one's condition, see Elzondo (2023, 8–9).

encapsulates the idea that humans can attain a state of contentment akin to divine bliss or a blissful life (*NE*, 1179a). He fundamentally suggests that such bliss and contentment is complete and self-sufficient to the extent that man finds enjoyment in the morally pure and contemplative elements that enable him to lead a "godlike activity." On the contrary, Kant is highly critical of the idea that happiness refers to such a divine form of contentment and bliss. In the lectures on ethics, he states bluntly:

> Well-pleasedness with one's own existence, when this existence is dependent, is called *happiness. Thus happiness is contentment with my own dependent existence.* But a complete well-pleasedness with one's independent existence is called *acquiescentia in-semetipso* or self-sufficiency *(beatitudo).* This blessedness of a being consists therefore in a *well-pleasedness* with one's own existence *apart from any need,* and thus it belongs solely to God alone; for he alone is independent (*LPDR*, 28:1060).

In this passage, Kant makes it clear that blessedness (*beatitudo*) refers to a state of complete independece from physical causes and inclinations, which he identifies as a state that is unique to God. In a similar vein, in the second *Critique*, Kant states that blessedness refers to a "self-sufficiency that can be ascribed to the supreme being" (*CPrR*, 5:119). Admittedly, Kant's notion of bliss has no place in the standard reading of happiness, which assumes that virtue or moral worthiness is accompanied by happiness, which denotes the "satisfaction of all one's inclinations or needs" as a physical or empirical form of satisfaction and fulfilment. Therefore, the standard reading of happiness condemns the idea that happiness has a eudiamonistic aspect because it involves the element of eternal or divine pleasure or contentment. Nonetheless, Kant's notion of happiness does not entirely exclude the prospect of divine pleasure and bliss as an ideal that can be attained through the pursuit of virtuous activity for perfection and holiness. In the *Religion*, this aspect of happiness becomes clear in Kant's notion of "Kingdom of God", which is characterized as an ethical state to be achieved with the cooperation between human efforts and the divine aid that would bring salvation and bliss to human agents. Kant characterizes the ethical state as a sublime and pure state to be achieved with the aid of divine power, which will help human agents to achieve moral perfection with the help of divine

bliss (*Rel*, 6:183-185). In this sense, he envisions the ethical state as a family-like community in which God treats human beings like a father who helps his children to achieve a union of hearts (*Rel*, 6:102). This metaphor of the loving relationship between God and human agents is also echoed in Kant's notion of divine blessedness (or godliness), where he argues that human agents, as children of God, are obliged to respect his law by free choice, which gives them pleasure (*Rel*, 6:182). Accordingly, Kant suggests that respect and love for God as the supersensible being behind the possibility of the highest good gives agents pleasure and bliss because they associate themselves with the law and goodness of God, which helps them cultivate their moral perfection and holiness (*Rel*, 6:185).

By characterizing moral faith as "a receptivity to (worthiness of) eternal happiness" (*Rel*, 6:115), Kant further affirms that the pursuit of holiness and perfection through practical faith brings about the hope for a blessedness and bliss:

> In the *practical foith in this Son of God* (so far as he is represented as having taken up human nature) the human being can thus hope to become pleasing to God (and thereby blessed); that is, only a human being conscious of such a moral disposition in himself as enables him to *believe* and self-assuredly trust that he, under similar temptations and afflictions (so far as these are made the touchstone of that idea), would steadfastly cling to the prototype of humanity and follow this prototype's example in loyal emulation, only such a human being, and he alone, is entitled to consider himself not an unworthy object of divine pleasure (*Rel,* 6:62).

In this excerpt, Kant makes it clear that the moral struggle to fulfill moral duties to perfection and holiness and moral faith by emulating the ideal image of humanity (a human being that seeks to be well-pleasing to God through its autocratic self-struggle) will ultimately lead to the possibility of divine grace, which will cherish their moral activity with divine blessedness and bliss. This clearly suggests that divine bliss is an unattainable ideal for ordinary human agents. As such, Kant notably associates it with an ideal image of human being making superhuman efforts to realize such divine pleasure and blissful state. Nevertheless, this ideal image of human being

that seeks divine bliss and blessedness through moral striving and struggle is not a purely abstract notion. Instead, Kant's formulation of such a prototype of humanity as an ideal to be constantly pursued by the ordinary agents has a firm basis within the scope of morality and ethics as a moral duty to be fulfilled. Likewise, one should avoid the assumption that such divine pleasure or bliss is an afterlife issue. In his critique of the Christian doctrine of the highest good, which entails that happiness (or *beatitudo*) is an afterlife issue (*CPrR*, 5:129; *Rel*, 6:135), Kant conversely claims that the "Kingdom of God" is "a beautiful ideal of the moral world-epoch brought about by the introduction of the true universal religion and *foreseen* in faith in its completion - one which we do not *see directly;* in the manner of an empirical completion but *have a glimpse* in the continuous advance and approximation toward the highest possible good on earth" (*Rel*, 6:136). By framing the "Kingdom of God" as a duty to be approached in an ongoing struggle, Kant makes it clear that moral striving of moral worthiness with the aid of moral faith would ultimately bring divine pleasure and bliss to agents through their pursuit of holiness and perfection. Thus, it seems that while Kant severely criticizes the notion of *beatitudo* or bliss as a possibility that can be attained by human agents, at a deeper level, he retains the notion of the pursuit of divine bliss or pleasure as an object of hope within the framework of morality or ethics insofar as agents pursue a lifestyle that is pleasing to God.

Conclusion

In conclusion, throughout this paper, I have argued that morally autocratic activity based on the pursuit of perfection and holiness, requires moral belief or faith as an active disposition of agents to aid them to cultivate virtuous disposition towards a life conduct that is morally pure and refined. I have further elaborated that moral autocracy involves self-conquest (self-overcoming), which demands agents to control their inclinations and evil and seek the purity and holiness of the moral law. In the course of such virtuous activity towards moral purity and refinement, moral belief or faith, then, emerges to play a crucial role in guiding and encouraging agents to lead a lifestyle that is well-pleasing to God, i.e., life conduct that is peculiar to the

highest good. I have pointed out that such life conduct that is grounded on the active disposition of moral belief inevitably involves overcoming a robust and uncultivated form of living and raising one's moral character and lifestyle into an elevated and sublime state. Finally, I have concluded that the deeper and broader reading of happiness in Kant's account has a eudaimonistic aspect that adds to the maximal satisfaction of one's needs and desires and includes a form of moral self-satisfaction and bliss to the extent that agents achieve perfection and holiness through their autocractic self-struggle.

References

Allison, Henry E. 1990. *Kant' Theory of Freedom*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139172295

Aristotle. 2014. *Nicomachean Ethics*, edited and translated by Roger Crisp, Cambridge: Cambridge University Press.

Baxley, Anne M. 2010. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511779466

Cicovacki, Predrag. 2001. *Kant's Legacy: Essays in Honor of Lewis White Beck* (Vol. 2). Rochester, New York: University Rochester Press. https://doi.org/10.1017/9781580466011

Coble, Kelly. 2003. "Kant's Dynamic Theory of Character." *Kantian Review*, *7*: 38–71. https://doi.org/10.1017/S1369415400001734

Denis, Lara. 1999. "Kant on the Perfection of Others." *The Southern Journal of Philosophy*, *37*(1): 21–42. https://doi.org/10.1111/j.2041-6962.1999.tb00855.x

Elizondo, E. Sonny. 2023. "Kantian Eudaimonism." *Journal of the American Philosophical Association*, 9(4): 655–69. https://doi.org/10.1017/apa.2022.30

Engstrom, Stephen. 1996. "Happiness and the Highest Good in Aristotle and Kant." In *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, edited by Stephen Engstrom and Jennifer Whiting, 102–38. Cambridge: Cambridge University Press.

Förster, Eckart. 2000. *Kant'' Final Synthesis: An Essay on the Opus Postumum*. Cambridge, Massachusetts: Harvard University Press.

Gregor, Mary J. 1963. *Laws of Freedom: A Study of Kant's Method of Applying the Categorical Imperative in the* Metpaphysik der Sitten. Oxford: Basil Blackwell.

Grenberg, Jeanine M. 2022. *Kant's Deontological Eudaemonism: The Dutiful Pursuit of Virtue and Happiness*. Oxford: Oxford University Press.

Holberg, Erica A. 2018. "Kant's Quasi-Eudaimonism." *The Southern Journal of Philosophy*, 56(3): 317–41. https://doi.org/10.1111/sjp.12296

Hills, Allison. 2014. “Gesinnung: Responsibility, Moral Worth, and Character.” In *Kant’s Religion within the Boundaries of Mere Reason: A Critical Guide*, edited by Gordon E. Michalson. 79–97. Cambridge: Cambridge University Press.

Irwin, Terence. 1996. “Kant’s Criticisms of Eudaimonism.” In *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, edited by Stephen Engstrom and Jennifer Whiting, 63–101. Cambridge: Cambridge University Press.

Kant, Immanuel. 1996. *Lectures on the Philosophical Doctrine of Religion*. In *Religion and Rational Theology*, translated by Allen W. Wood and George di Giovanni. Cambridge: Cambridge University Press.

Kant, Immanuel. 1997. *Lectures on Ethics*, edited by d. Peter Heath and J. B. Schneewind, translated by Peter Heath. Cambridge: Cambridge University Press.

Kant, Immanuel. 1998. *Critique of Pure Reason*, translated and edited by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

Kant, Immanuel. 2002. *Critique of the Power of Judgment*, edited by Paul Guyer, translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press.

Kant, Immanuel. 2005. *Notes and Fragments*, edited by Paul Guyer, translated Curtis Bowman, Paul Guyer, and Frederick Raucher. Cambridge: Cambridge University Press.

Kant, Immanuel. 2012. *Groundwork of the Metaphysics of Morals*, edited by Mary Gregor and Jens Timmermann. Cambridge: Cambridge University Press.

Kant, Immanuel. 2015. *Critique of Practical Reason*, translated and edited by Mary Gregor. Cambridge: Cambridge University Press.

Kant, Immanuel. 2017. *The Metaphysics of Morals*, edited by Lara Denis, translated by Mary Gregor. Cambridge: Cambridge University Press.

Kant, Immanuel. 2018. *Religion Within the Boundaries of Mere Reason*, edited Allen Wood and George di Giovanni. Cambridge: Cambridge University Press.

Kant, Immanuel. 1960. *Religion Within the Limits of Reason Alone*, translated and edited by Theodore M. Greene and Hoyt H. Hudson. New York: Harper & Row.

Kant, Immanuel. 2012. *Lectures on Anthropology*, edited by Allen W. Wood and Robert B. Louden. Cambridge: Cambridge University Press.

Kant, Immanuel. 2018. "What Does it Mean to Orient Oneself in Thinking?" In *Religion Within the Boundaries of Mere Reason and Other Writings*, edited by Allen Wood and George di Giovanni. Cambridge: Cambridge University Press.

Louden, Robert B. 2015. “ ‘The End of All Human Action’/’The Final Object of All My Conduct’: Aristotle and Kant on the Highest Good.” In *The Highest Good in Aristotle and Kant*, edited by Joachin Aufderheide and Ralf M. Bader,

pp. 112–128. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198714019.003.0006

Munzel, G. Felicitas. 1999. *Kant's Conception of Moral Character: The "Critical Link of Morality, Anthropology, and Reflective Judgment*. Chicago: University of Chicago Press.

Palmquist, Stephen R. 2015. "What is Kantian Gesinnung? On the Priority of Volition over Metaphysics and Psychology in Religion within the Bounds of Bare Reason." *Kantian Review*, 20(2): 235–64. https://doi.org/10.1017/S1369415415000035

Pasternack, Lawrence 2017. "Restoring Kant's Conception of the Highest Good." *Journal of the History of Philosophy*, 55(3): 435–68. https://doi.org/10.1353/hph.2017.0049

Silber, John R. 1963. "The Importance of the Highest Good in Kant's Ethics." *Ethics*, 73(3): 179–97. https://www.jstor.org/stable/40753543

Stratton-Lake, Philip. 2005. *Kant, Duty and Moral Worth*. New York and London: Routledge.

Walschots, Michael. 2017. "Kant on Moral Satisfaction." *Kantian Review*, 22(2): 281–303. https://doi.org/10.1017/S136941541700005X

Ware, Owen. 2023. *Kant on Freedom. Elements in the Philosophy of Immanuel Kant*. Cambridge: Cambridge University Press.

Wood, Allen W. 1970. *Kant's Moral Religion*. Ithaca, New York: Cornell University Press.

Wood, Allen W. 2001. "Kant versus Eudaimonism." *Kant's Legacy: Essays in Honor of Lewis White Beck*, edited by Predrag Cicovacki, 261–82. Rochester, New York: University of Rochester Press.

Wood, Allen W. 1984. "Kant's Compatibilism." In *Self and Nature in Kant's Philosophy*. 73–101. Ithaca, New York: Cornell University Press.

Zeldin, Mary-Barbara. 1971. "The Summum Bonum, the Moral Law, and the Existence of God." *Kant Studien*, 62(1–4), 43–54. https://doi.org/10.1515/kant.1971.62.1-4.43

Organon F 31 (4) 2024: 367–387
https://doi.org/10.31577/orgf.2024.31402

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Compulsion, Ignorance, and Involuntary Action: An Aristotelian Analysis

Huiyuhl Yi*

Received: 5 May 2023 / Revised 5 August 2024 / Accepted: 7 August 2024

Abstract: Some remarks in the *Eudemian Ethics* and the *Nichomachean Ethics* indicate that the voluntariness of actions is significantly related to compulsion and ignorance. According to a plausible interpretation, these remarks suggest that if an agent performs an action under compulsion or due to ignorance of some relevant facts, then she does so involuntarily. An objection to this interpretation with regard to compulsion is that an agent can voluntarily do what she is compelled to do. With regard to ignorance, one might object that it is necessary to clarify the proper range of relevant facts when considering whether an action performed out of ignorance is involuntary. In this paper, I develop two principles that align with the view that compulsion and ignorance are sufficient conditions for involuntary actions, while accommodating potential counterexamples and complications.

Keywords: Aristotle; compulsion; *Eudemian Ethics*; *Nichomachean Ethics*; ignorance; involuntary action.

* Ulsan National Institute of Science and Technology
http://orcid.org/0000-0002-2221-6406
School of Liberal Arts, 50 Unist-gil, Ulju-gun, Ulsan, 44919, Republic of Korea
huiyuhl@unist.ac.kr

1. Introduction

There are various conditions that can absolve us from the presumptive moral responsibility for our actions. Aristotle famously identifies two such exculpatory conditions—compulsion and ignorance—which he links to the concept of voluntariness. For instance, in Book III of the *Nichomachean Ethics* (hereafter NE), Aristotle makes the following remarks: "Those things, then, are thought involuntary, which take place by force or by reason of ignorance" (NE, 1109b35–1110a1)[1] and "that which is done by force or by reason of ignorance is involuntary" (NE, 1111a21). Also, in Book II of the *Eudemian Ethics* (hereafter EE), he writes: "Anything one does without ignorance that is up to oneself not to do is necessarily voluntary, and the voluntary is this. Everything that one does in ignorance and due to ignorance is involuntary" (EE, 1225b8–11).[2] These remarks suggest that compulsion and ignorance are each a sufficient condition for involuntary action. We can formulate the ideas in the form of two principles, which may be called *the compulsion-voluntariness principle* and *the ignorance-voluntariness principle*, respectively, or:

(CV) If S Φ-s under compulsion, then S Φ-s involuntarily, and

(IV) If S Φ-s due to ignorance of relevant facts about Φ-ing, then S Φ-s involuntarily,

in which 'S' stands for the agent and 'Φ' stands for the action under a certain description in the relevant context. As subsequent analysis will reveal, critical considerations may cast doubt on the truth of these principles. In the following discussion, I examine potential counterexamples and observations that challenge them. Additionally, I formulate revised principles to address the complications that arise. By offering these revised principles,

[1] I use David Ross's (2009) translation of the *Nichomachean Ethics* in this paper.

[2] I use the translation of Brad Inwood and Raphael Woolf (2013). Robert Heinaman (1988, 253) analyzes the notion of voluntary action in EE as follows: A Φ-ed voluntarily if and only if (i) A Φ-ed, (ii) It was in A's power not to Φ, (iii) A Φ-ed with knowledge, (iv) A Φ-ed through himself. According to Heinaman, (ii) and (iv) are crucial in understanding Aristotle's views on compulsion. Part of this paper can be seen as a complement to the clarification of (iv).

my aim is to articulate the relationship between compulsion and involuntary action, as well as the relationship between ignorance and involuntary action, while striving to accommodate an Aristotelian perspective on these notions.

2. Compulsion and involuntary action

I begin by considering the principle governing compulsion and involuntary action, namely, (CV). There may be apparent counterexamples to this principle. Consider the following scenarios:

Poisoning (I)
George could not stand the arrogance of his boss and planned to poison him. He was strongly determined to murder his boss. But just before putting his plan into action, George was stopped by Elaine, who also wanted to eliminate him but did not want to get her hands dirty. Elaine threatened George so that he would poison the boss. Afterward, George gladly, and voluntarily, poisoned him, although his action was performed under compulsion.

Robbery
Jerry is a bank robber. On his way to rob a bank, he was stopped by Kramer, the sniper, who had a personal vendetta against the bank owner but did not want to risk getting caught. Kramer threatened Jerry so that he would rob the bank, telling him that doing otherwise would cost his life. Although Jerry was a bit surprised and intimidated, he gladly, and voluntarily, robbed the bank.

In both scenarios, each agent voluntarily did what he did although he was threatened to do so. Therefore, apparently, these cases work as counterexamples to (CV). It might be objected that the mere fact that an agent is doing something gladly does not imply necessarily that she thereby does it voluntarily. However, in providing these examples, I do not mean to suggest that voluntary actions are always accompanied by pleasure, even though Aristotle may be interpreted as asserting that voluntary actions should be pleasant, while involuntary actions tend to be painful (NE, 1110b12–22).

Why the action in each case can be characterized as voluntary is because its driving force stems from the agent. According to Aristotle, "the voluntary would seem to be that of which the moving principle is in the agent himself, he being aware of the particular circumstances of the action" (NE, 1111a22–24).[3] Many philosophers throughout history have acknowledged that voluntary actions must originate from the agent. For instance, in the *Summa Theologica*, Thomas Aquinas argues that an action is voluntary when it is "according to the inclination of the will," asserting that "it is impossible for a thing to be absolutely coerced or violent, and voluntary" (I Q82 A1), as exemplified by a man who is "dragged by force" (I.II Q6 A4). Similarly, in *An Essay Concerning Human Understanding*, John Locke emphasizes that voluntary action is the product of volition. Since volition is the "particular determination of the mind" (II.XXI.30, 250), he asserts that the "forbearance or performance of action, consequent to such order or command of the mind is called [v]oluntary," whereas "whatsoever action is performed without such a thought of the mind is called [i]nvoluntary" (II.XXI.5, 236). If a voluntary action originates from the agent, as suggested by the preceding remarks, then it follows that the agent in each of the previous examples performs the action in question voluntarily.

However, it may be argued that these examples take the notion of compulsion too broadly. There, given that "the moving principle" resides in each of the agent when they acted, the threat does not play a role in the execution of the act. Therefore, it is reasonable to claim that George and Jerry were not really forced to do what they did. In a similar line of reasoning, Harry Frankfurt (1969, 832–33) claims that in a circumstance in which an agent is committed to perform a wrongful action while being threatened to do it, if she is not motivated by the threat in performing the action, then it is most plausible to say that she is not really coerced to do it. Here, the crucial idea is that in order for someone to be properly coerced

[3] In a similar vein, in his analysis of relevant passages from EE and NE on this point, John M. Cooper (2013, 276–77) observes that "[t]he voluntary is whatever action has its originating source within an agent," and that "what it is to be voluntary is to be an action that has that sort of source, viz., a causal one internal to the agent."

or compelled to do something, the driving force of the action must not originate from the agent herself.

In this sense, those who defend (CV) may argue that cases like Poisoning (I) and Robbery misinterpret the notion of compulsion as used by Aristotle. As an example of an action performed under compulsion, Aristotle mentions a case of a person who "were to be carried somewhere by a wind, or by men who had him in their power" (NE, 1110a2–3). He also states that actions are involuntary "when the cause is in the external circumstances and the agent contributes nothing" (NE, 1110b2).[4] In the previous scenarios, the agents' actions stemmed from themselves, and they did contribute to what they did. The objection goes that, from Aristotle's viewpoint, they did not act under compulsion.

In response, I would like to note that the notion of compulsory action held by this objection is inconsistent. If the cause of the alleged "action" is in the external environment and not in the agent herself, how could she be said to perform an action at all? This is a legitimate question because, in most natural understandings of the term 'act' or 'action', when an agent acts, she should be able to control the immediate result of the action, such as the change in her bodily position, at her own disposal. Thus, when someone is carried away by a strong gust of wind, she cannot be said to act in any appropriate sense of the term; such an incident may be something compulsory, but it cannot be an action. Likewise, if a person is taken somewhere by those who have overpowered him, he does not seem to exercise his own agency.[5] This consideration leads us to think that if we follow the notion of

[4] Jozef Müller (2015) has argued that on the Aristotelian conception of voluntary action, the cause of the action must be based on the conditions under which the agent is the efficient cause of her action *qua* the individual that she is. On this view, a voluntary action differs from an action performed in virtue of one's nature. By contrast, Giulio Di Basilio (2021, 12–16) argues that, according to Aristotle, the impulse that underpins the voluntariness of an action must be not only internal, but also *natural*, to the agent.

[5] Cooper (2013, 279–82) offers alternative interpretations of the two examples given by Aristotle. In his view, the first example involves a person who ends up going somewhere he did not wish to go, just like "a sailor in a boat who is blown by the wind to a landing in some harbour [he was] not sailing to." In such a case, the sailor was indeed *acting* in confrontation with the wind (e.g., erecting the boat,

compulsion suggested by the objection—i.e., if 'compulsory' implies that the cause of the given action is in the external circumstances (or "the moving principle is outside" (NE, 1110a2)) and the agent contributes nothing—then it makes no sense to talk of "compulsory action." What is compulsory in that sense cannot be an action. If this is correct, then in performing a compulsory action, an agent should be able to exercise some sort of control over what she does although she is forced or compelled to do so.

The notion of compulsory actions may be understood as analogous to what Aristotle described as "mixed actions."[6] According to Aristotle, mixed actions involve "things that are done from fear of greater evils or for some noble object" (NE, 1110a4–5). Typically, refusal to perform such actions results in severe consequences, such as "beating or imprisonment or death" (EE, 1225a5). On the one hand, these actions seem involuntary, since "no one would choose any such act in itself" (NE, 1110a19). On the other hand, they appear voluntary in the sense that they "are chosen at the time when they are done, and the end of an action is relative to the occasion" (NE, 1110a12–13).

According to these remarks, individuals who perform mixed actions can be seen as being compelled to act in a certain manner by external forces or

keeping the rudder, and so on), but involuntarily arrives at an undesired place. Likewise, the second example may involve a walking person who is "pushed… and restrained from going off in other directions" and ends up somewhere he didn't intend to go to by those who have him under their control. The problem of Cooper's suggestion, though, is that what happens at the end of the process in each example is not the result of the agent exercising his own agency. For instance, though it is correct to say that the sailor performs many actions during the voyage in the proper sense of the term, it is not these actions that transport him to the destination. Rather, it is the wind that causes his migration. For this reason, it seems correct to say that he was *made* to reach the place because of external circumstances, as opposed to exercising his agency involuntarily. A similar remark applies to the second example.

[6] For an extensive discussion of the nature of mixed actions, see (Nielsen 2007). According to Nielsen, 'mixed action' does not fall under a category of the ontological structure of actions considered by Aristotle. She argues that it is a provisional label for coerced acts with the purpose of enlightening Aristotle's genuine view on compulsion.

circumstances, even though it is not impossible for them to choose otherwise. For instance, consider a situation in which someone is coerced into performing a disgraceful act to save family members held hostage by a tyrant. Similarly, imagine a scenario in which a ship's captain must jettison the cargo during a storm to ensure the safety of the crew and himself (NE, 1110a5–9). In both instances, the actions of the agents can be characterized as mixed. However, it is crucial to note that compulsory actions are not synonymous with mixed actions described as such. In his discussion of mixed actions, Aristotle refers to cases where the agent is forced to act against their will. For example, if an agent coerced into performing a disgraceful act to save their family happens to have desired to perform the act, it is still accurate to say that they act under compulsion. Nevertheless, Aristotle would not consider this an appropriate example of a mixed action.

Regarding the preceding discussion, one might suggest that we interpret the Aristotelian notion of compulsory action in terms of reasons rather than causes. On this interpretation, the presence of an internal reason for an action is sufficient to deem that action voluntary. Therefore, compulsory actions rule out internal reasons in the sense that one is compelled to act only if one lacks an internal reason for doing so. To say that the cause of a compulsory action is external to the agent simply means that the reason for the action is external to the agent.[7] In this line of thinking, Poisoning (I) and Robbery fail to exemplify compulsory actions because the agent in each scenario acts for a reason internal to the agent. Similarly, a dishonorable action performed under threat does not constitute a compulsory action if the agent had a desire to perform it, as this desire indicates that the agent had an internal reason for acting.

I find this alternative interpretation misguided because it does not align with ordinary cases of compulsory actions. In typical situations where a person is compelled to act, they do so to avoid informed undesirable consequences, which may be overwhelming or devastating. Consequently, it is natural for the agent to develop a strong desire to avoid such consequences. This desire must belong to what Bernard Williams terms the "subjective

[7] I am grateful to an anonymous reviewer from *Organon F* for guiding me to consider this alternative interpretation.

motivational set" (1979, 18). If this is correct, then in carrying out a compulsory action in most ordinary contexts, the agent would have an internal reason for acting as they do.

If we allow for the possibility that a compulsory action can be performed even when the agent has an internal reason to do so, then (CV) is once again vulnerable to the aforementioned counterexamples. In the Poisoning (I) scenario, when George is coerced by Elaine to poison his boss, following her directions would align perfectly with his purpose, indicating that he had an internal reason to poison his boss. The same applies to Jerry in the Robbery scenario. Although it is not impossible for Jerry to refuse to rob the bank despite Kramer's instructions, following the directions would be most advantageous for him. This again demonstrates that Jerry had an internal reason to rob the bank. Generally, it is entirely possible for a threatened person to willingly and voluntarily do what they are coerced into doing. In such cases, (CV) would be shown to be false.

However, rephrasing compulsory actions in terms of reasons rather than causes may illuminate how we should revise (CV). As noted, the problem with (CV) arises from the fact that the agent's action is not motivated by the coercive force in the previous examples. Rather, it is the internal reason of each agent that explains their behavior. Given that an agent has an internal reason to act in a certain way, the mere fact that they are under compulsion does not necessarily imply that they act *because of* the compulsion. Compulsory actions should be those in which one is motivated to act by the compulsion itself. As such, it is tempting to add a clause specifying that the agent's action should originate from the compulsion. Accordingly, to avoid the difficulty raised by the previous examples, I suggest that (CV) be revised to:

(CV*) If (i) S Φ-s under compulsion, and (ii) S would not Φ if S was not under compulsion, then S Φ-s involuntarily.

Thanks to the addition of clause (ii), (CV*) now successfully explains why the agents' actions in the preceding examples are not involuntary. Given that George had an internal reason to poison his boss, he would have poisoned his boss anyway even if he had not been threatened. Hence, his action of poisoning the boss is not an instance of an involuntary action. The same is true for Jerry: provided that he had an internal reason to rob the bank,

he would have done it anyway even without the threat from Kramer. The fact that the two agents would have acted in the way that they actually did even without the threats shows that the threats presented to them did not play a part in the performance of their actions. The added clause places a constraint which enforces this point.

Another theoretical advantage of (CV*) is that it avoids the aforementioned problem inherent in the Aristotelian understanding of compulsory action. According to Aristotle, the impelling cause of a compulsory action must originate from the external environment rather than from the agent. This raises a difficulty in comprehending how something classified as compulsory can be considered an action at all. Notably, (CV*) does not reference the cause of the agent's action, yet it aligns with Aristotle's view that the driving force of a voluntary action must reside within the agent herself. In this regard as well, (CV*) represents an improvement.

3. Ignorance and involuntary action

I now turn to the Aristotelian principle regarding ignorance and involuntary actions. While some remarks in EE and NE seem to imply something like (IV), Aristotle's distinction between involuntary and non-voluntary actions (NE, 1110b17–24) indicates that he has a more sophisticated principle in mind. In differentiating the two kinds of actions, he states: "Everything that is done by reason of ignorance is *not* voluntary; it is only what produces pain and regret that is *in*voluntary" (NE, 1110b18–19). This remark suggests that some non-voluntary actions are not involuntary; a non-voluntary action is only involuntary if the agent's performance of it leads to regret. This reveals a problem with (IV): if a pertinent sort of ignorance is a sufficient condition for involuntary action as stated in (IV), then an action stemming from ignorance can be involuntary even without involving a regret, contra Aristotle.[8]

[8] One might object here that regret does not work as a necessary condition for involuntariness but simply as its symptom. However, even if that were the case, it remains true that some non-voluntary actions are accompanied by regret while some non-voluntary actions are not. Since Aristotle clearly meant to equate involuntary

To address this complication, let us note that an action accompanied by regret can typically be characterized as one that the agent would not have performed if she had known some relevant facts. For instance, Mr. and Mrs. Perry would not have pushed their son Neil too harshly to send him to an Ivy League medical school if they had known that it would have ultimately caused him to take his own life. In this sense, they genuinely regret what they did: if they had known better, they would never have acted in the ways they actually did. If it is appropriate to understand the nature of regret-inducing actions in the manner previously described, then, to accommodate the notion of involuntary (as opposed to non-voluntary) action, we can revise (IV) as follows:

(IV*) If (i) S Φ-s due to ignorance of relevant facts about Φ-ing, and (ii) S would not Φ without such ignorance, then S Φ-s involuntarily.

The newly added clause (ii) is meant to capture Aristotle's characterization of involuntary actions with regard to ignorance, which, unlike non-voluntary actions that are not also involuntary, involves the regret of the agent.[9]

Although (IV*) seems more promising in accommodating an Aristotelian notion of ignorance and involuntary actions, it may be argued that (IV*) is open to counterexamples. People commonly make decisions while experiencing epistemic limitations. For example, they may not fully understand the consequences of their actions or be aware of all available courses of action at the time of decision-making. In such circumstances, it is hard to say that their actions are involuntary. To illustrate this point, consider the following scenarios:

---

actions with the former as opposed to the latter, it is important to distinguish involuntary actions from non-voluntary actions that are not involuntary.

[9] In formulating (IV*), I characterize the role of regret in purely epistemic terms because this principle specifies how ignorance is related to volitional action. However, I do not mean to deny that there are other aspects of regret that can be highlighted with regard to voluntariness of an action. For instance, it may be argued that regret, as a kind of reactive attitude, can be used to self-assess the agent's actions. Here, I follow Audrey L. Anton (2020) in taking regret as a self-reactive attitude.

Poisoning (II)
Someone slipped a deadly poison into my water bottle, which killed me after I drank the water. Since I was not suicidal, I drank it without knowing it had been poisoned. Had I known it had been poisoned, I would not have drunk it.

Chess
While playing chess against Borgov, Beth made a particular move (moving the Knight to the F3 square) without realizing that there was another move (moving the Queen to the C7 square) that would have secured her the championship. Unfortunately, her actual move turned out to be a mistake, resulting in her defeat in the game.[10]

Gas Station
Jill went to a gas station to fill up her car's gas tank. Unbeknownst to her, another gas station around the corner offered cheaper prices than the one she went to. Jill would not have filled up with gas there if she had known this fact.

These examples may seem straightforward. In each case, the agent acted out of ignorance, and she would not have done what she did if she had known better. Nevertheless, in each case, the action in question was performed in accordance with the agent's autonomous decision-making process. Hence, it may be argued that the agent acted voluntarily. For example, in the story of Poisoning (II), I willingly picked up the bottle and poured the water into my mouth, making my action of drinking seem voluntary. Similarly, in the story of Chess, Beth made a conscious decision to make her move and acted based on her own judgment. Thus, her action appears to be voluntary. Additionally, in the story of Gas Station, Jill freely stopped by the gas station and grabbed the gas pump handle of her own accord, indicating that she filled up with gas there voluntarily. These instances might be viewed as counterexamples to (IV*).

[10] I am indebted to the reviewer from *Organon F* for this example. This scenario demonstrates how one can seem to act voluntarily even without awareness of the available actions. I am grateful to the reviewer for inviting me to address it.

However, I think we can develop a strategy to save (IV*) from this type of examples. The main tactic is to argue that the facts cited in the preceding examples—i.e., that the water had been poisoned, that the move Beth was making would cost her the match, and that the price at a nearby gas station was cheaper—are not really relevant to the voluntariness of the actions as they are currently described. Note that (i) in the antecedent of (IV*) only refers to a *relevant* kind of fact with respect to the ignorance that gives rise to the agent's action. Hence, if an agent's action is guided by ignorance of a certain fact, but that fact is *not* relevant to the action in question, then (IV*) is not applicable to this sort of case.

This approach naturally prompts an inquiry into the scope of relevant facts: what constitutes the appropriate range of relevant facts in applying (IV*)? What criteria can be used to discern which facts are relevant and which are not? In this context, Aristotle identifies several potential candidates for relevant facts concerning ignorance and voluntariness:

> A man may be ignorant, then, of who he is, what he is doing, what or whom he is acting on, and sometimes also what (e.g., what instrument) he is doing it with, and to what end (e.g., he may think his act will conduce to someone's safety), and how he is doing it (e.g., whether gently or violently). […] The ignorance may relate […] to any of these things, and the man who was ignorant of any of these is thought to have acted involuntarily, […] especially if he was ignorant on the most important points; and these are thought to be the circumstances of the action and its end (NE, 1111a3–19).

The last statement highlights that the most important kinds of relevant facts pertaining to ignorance and voluntariness in performing an action have to do with the nature and the purpose of the action. More specifically, Aristotle seems to think that there are two paradigm examples of relevant facts with respect to ignorance: facts about what the agent is doing, and facts about what she is doing it for (or why she is doing it). In what follows, I will show that none of the preceding examples undermines (IV*), while focusing on whether they cite a relevant fact with respect to what the agent is doing or why the agent is doing it.

It seems plain that when a person performs an action under a certain description (say, when S Φ-s), the fact that she is Φ-ing is a relevant fact with regard to her action of Φ-ing. So, if an agent does not know what she is doing in performing an action, this should be a strong ground to think that she is not doing it voluntarily. In other words, we may plausibly hold that if S is ignorant of the fact that S is Φ-ing, then S does not Φ voluntarily. Then, it is crucial to ask whether each of the agents in the preceding examples knew what they were doing. In my view, the answer depends on how we describe their actions. An action can be described in more than one way.[11] For example, we could redescribe the story of Poisoning (II) in terms of *killing oneself* as opposed to *drinking water*. The result would be the following: I was (inadvertently) killing myself when I drank the water because I was ignorant of the fact that the water had been poisoned; had I known this fact, I would not have killed myself; therefore, my action of killing myself was involuntary. In this new description of the story, I did not know what I was doing (in the sense that I was unaware that I was killing myself). Here, the fact that the water had been poisoned is entirely relevant with respect to what I was doing, since my action of self-killing was only possible owing to this fact. After all, for a person to be killed by consuming some material, it is essential that the material is deadly. However, if we describe my action in terms of drinking water as originally stated, then I surely knew what I was doing (given that I was aware that I was drinking water). The fact that the water was poisoned is hardly relevant here, because whether or not one can successfully drink water has nothing

[11] Here, I follow the standard view regarding the individuation of action, according to which a single "basic action," which involves the immediate outcome of the agent's bodily movements, can be instantiated in many different actions-under-a-description. In this view, it is possible that an action under a certain description is voluntary, while the same action under a different description is not. This may differ from Aristotle's own view, according to which multiple distinct actions may occur at a time, as opposed to a single action occurring under many different descriptions. See Cooper (2013, 277–78). If we follow Aristotle's view in assessing (IV*), then each occurrence of 'Φ' should be taken to denote an action as opposed to an action-under-a-description.

to do with its toxicity.[12] In this case, some other facts—e.g., the fact that water is drinkable or that it was indeed water that was contained in my bottle—are relevant to my action of drinking water. But I was not ignorant of these facts (or so we can stipulate).

The preceding observation reveals that Poisoning (II) does not constitute a counterexample to (IV*) in considering ignorance with respect to what the agent is doing. If we characterize the story while describing my action in terms of drinking water as originally stated, then the consequent of (IV*) would indeed be instantiated as being false because I was voluntarily drinking the water. However, the antecedent would *not* be instantiated as being true because the fact of which I was ignorant (i.e., the fact that the water was poisoned) is not relevant to my action of drinking water. By contrast, I was not ignorant of the *relevant* facts with respect to what I was doing–e.g., I surely knew that water is a drinkable substance. On the other hand, if we characterize the story by describing my action in terms of another verb such as killing myself, then the antecedent of (IV*) would be instantiated as being true, given that I killed myself due to ignorance of a relevant fact (that the water was poisoned), and I would not have done it had I known this fact; however, the consequent of (IV*) would *not* be instantiated as being false, because I did not voluntarily kill myself. Therefore, neither description of the poisoning story refutes (IV*) as we consider whether the agent was ignorant of a relevant fact with respect to what he was doing.

Could this example refute (IV*) if we focus on the ignorance of relevant facts with respect to the purpose of the action (or why the agent did what she did)? Suppose that we redescribe the story in a way that underscores

[12] It does not make a difference if the agent is faced with partial ignorance of the situation. Suppose the agent knew that what she was about to drink was water, but also knew, without knowing that this particular cup of water is poisoned, that in general water can be poisoned, and that one should take a risk in drinking a cup of water. (We can imagine that the agent was suffering from extreme thirst and was offered several cups of water after being informed that one of them was poisoned.) In this case, we can say that the agent was voluntarily taking a risk. Furthermore, the fact that the water in the chosen cup was poisoned appears to be relevant to her taking a risk. However, it is not *because of* ignorance of this fact that she took a risk. Rather, she took a risk because she was dying from thirst. Hence, this example does not falsify (IV*).

the purpose of the action. The result would be as follows: I was quenching my thirst by drinking water because I was ignorant of the fact that the water had been poisoned; had I known this fact, I would not have done it. In this new version of the story, my action of quenching my thirst still seems voluntary. Therefore, the consequent of (IV*) would be instantiated as being false. However, the antecedent of (IV*) would *not* be instantiated as being true, since the fact of which I was ignorant (namely, that the water was poisoned) is not relevant to my action of quenching my thirst. In general, the toxicity of material is not pertinent to its quenching capacity. There are certain facts that are relevant to my action of quenching my thirst—e.g., the fact that water tends to quench one's thirst. However, I was not ignorant of this fact, and that is precisely why I drank the water. Once again, (IV*) is not refuted by this version of the story.

So far, I have argued that Poisoning (II) does not undermine (IV*) with respect to what the agent is doing and why the agent is doing it. This strategy equally applies to the Chess story. Let me first focus on relevant facts in terms of the nature of the action (or what the agent is doing). I would like to note that, although Beth was ignorant of the fact that moving the Queen would win her the game, this fact is barely relevant in terms of what she was doing. After all, whether there are alternative moves (such as moving the Queen) has no bearing on her successfully making her actual move (namely, moving the Knight). There are other facts that *are* relevant to her making the actual move, such as the fact that the piece she was about to pick up was indeed the Knight or the fact that the Knight was in the position to be moved to the F3 square. But these are not the facts of which Beth was ignorant, given that she understood the rules of chess and knew she was moving the Kight to the aforementioned position at the time she was doing it.

The preceding observation reveals that the original description of Chess does not falsify (IV*). It is true that the consequent of (IV*) would be instantiated as being false—Beth did not make the move involuntarily. However, the antecedent would *not* be instantiated as being true, because Beth was not ignorant of any facts *relevant* to her moving the Knight. What she did not know—the fact that there existed another move that would secure her victory—is not relevant to her making the actual move.

Let us now see whether the Chess story can threaten (IV*) when we focus on the ignorance of some relevant facts with respect to the purpose of the action (or why the agent is doing the action). First, to redescribe what the agent is doing in terms of the purpose of the action, let us assume that Beth aimed to win the game by making her actual move. Then, to be germane to this aim, we can restate the story as follows: Beth was (inadvertently) incurring a defeat because she was ignorant of the fact that the move she was about to make was a mistake and there was a winning move available to her. Had she known this fact, she would not have incurred a defeat by making the actual move; therefore, her action of incurring a defeat was involuntary. In describing the story this way, we can say that Beth did not know what she was doing (in that she was unaware that she was incurring a defeat). The fact that she could win the game by making the alternative move is perfectly relevant to her action of incurring a defeat because, under the circumstances, she lost the game by failing to make the alternative move (or so we can stipulate).

However, even this version of the Chess story does not undermine (IV*). Here, it is true that the antecedent of (IV*) would be instantiated as being true since she incurred a defeat due to ignorance of a relevant fact. Had she known better (i.e., had she known that the move she was about to make would cost her the game or that she could win the game by making the alternative move), she would not have done it. However, in this case, the consequent of (IV*) would *not* be instantiated as being false because Beth did not incur a defeat voluntarily.

The same conclusion follows as to the Gas Station story. To focus on relevant facts in terms of the nature of the action (or what the agent is doing), I would like to note that, although Jill was ignorant of the fact that the price at a nearby gas station was cheaper, this fact is not relevant to the nature of her action in the sense that the price of gas has no bearing on whether or not she could successfully fill up with gas. On the other hand, Jill was not ignorant of some other facts that *are* relevant to her filling up with gas, such as the fact that what was contained in the fuel tank is indeed gasoline or the fact that what was in the tank is not water, given that she understood that she was filling up with gas at the time she was doing it.

It is possible to imagine defeasible circumstances that lead to skepticism about Jill knowing what she was doing. For instance, suppose some of the fuel tanks in the station contained water instead of gasoline, although the particular tank Jill selected happened to contain gasoline. In this case, it is arguable that Jill did not know that she was filling up with gas at the time she was doing it because she did not *know* that her tank contained gasoline. However, it does not follow that (IV*) is undermined by this skeptical scenario. The antecedent of (IV*) requires that there be a causal connection between the agent's action and the ignorance of the relevant fact in the sense that the agent does what she does *because of* the ignorance.[13] However, in this skeptical scenario, Jill's ignorance did not play a causal role for her action at all: it is not because she did not know that some of the fuel tanks contained water that she filled up with gas. Hence, the antecedent of (IV*) would be instantiated as being false in considering this skeptical version of the scenario.

Let us resume our discussion of the original description of Gas Station to see whether it causes a problem for (IV*). Here, it is true that the consequent of (IV*) would be instantiated as being false—Jill did not fill up with gas involuntarily. However, the antecedent would *not* be instantiated as being true, because Jill was not ignorant of any facts *relevant* to her filling up with gas, such as the fact that the fuel tank contained gasoline as opposed to some other material. The fact that the price at a nearby gas station was cheaper, which she did not know, is not relevant to her filling up with gas.

Finally, I want to explore if the Gas Station scenario might challenge (IV*) concerning the purpose of the action (or why the agent is performing the action). Assume Jill intended to save money when buying gas. With this goal in mind, the story can be reframed as follows: Jill was (unintentionally) causing herself a financial loss because she did not realize that a nearby gas station had cheaper prices; had she known this, she would not have caused the loss (by buying gas there). Thus, her action of causing herself a loss was involuntary. By framing the story this way, we can argue

[13] Susan Sauvé Meyer (2011, 176–79) emphasizes that, according to the Aristotelian conception of voluntariness, an agent, in acting involuntarily, must not only act *in ignorance* but also act *because of ignorance*.

that Jill was unaware of what she was doing (since she did not know that she was causing herself a loss). Furthermore, the fact that the nearby gas station had a lower price is highly relevant to her causing a loss, as the loss resulted from the price difference between the two gas stations.

Nevertheless, this version of the Gas Station story does not invalidate (IV*). In this scenario, the antecedent of (IV*) would indeed be satisfied because she caused herself a loss due to ignorance of a relevant fact; if she had known about the lower price at a nearby gas station, she would not have done it. However, in this case, the consequent of (IV*) would *not* be falsified, because Jill did not voluntarily cause the loss to herself.

So far, I have focused on the ignorance of relevant facts in terms of what the agent is doing and why the agent is doing it to show that the preceding apparent counterexamples do not undermine (IV*). What about the other candidates for relevant facts listed in the previously quoted passage in NE? In my view, they are either not unknown to the agent or not in fact pertinent to the examples. For instance, in accordance with the quoted passage, we may include what Aristotle would regard as legitimate relevant facts in the Gas Station story as follows:

> The fact that it was Jill who was filling up with gas [who the agent is]
> The fact that Jill was filling up with gas into the gas tank of her car [what or whom the agent is acting on]
> The fact that Jill was using a particular gas pump nozzle of the fuel tank [what the agent is acting with]
> The fact that Jill was gripping the nozzle firmly in filling up with gas [how the agent performs the act]

We can reasonably suppose that Jill was not ignorant of the first three facts. As for the fourth, Jill might be unaware of this fact; however, how hard she was gripping the nozzle is largely irrelevant to whether she could successfully carry out her action of filling up with gas.[14] Therefore, none of the

[14] There may be other candidates regarding how the agent performs the act that could be considered relevant to the successful execution of Jill's act, such as her mood or level of distraction while performing the task. However, if we describe her act as simply filling up with gas, her lack of awareness of her psychological state or conditions is irrelevant to her performing the action. In contrast, if we describe the

listed facts causes a problem for (IV*). A similar observation can be made regarding the other two examples.

The preceding discussion demonstrates the difficulty in describing the scenarios from the previous examples such that the antecedent of (IV*) is instantiated as true while the consequent of (IV*) is instantiated as false. Consequently, none of these examples provides a counterexample to (IV*) when we consider possible candidates for relevant facts, including those pertaining to the nature of the action (i.e., what the agent is doing) or the purpose of the action (i.e., why the agent is doing it).

4. Conclusion

I have examined the Aristotelian notions of compulsion and ignorance, which understand them as sufficient conditions for involuntary actions, and represented them in the form of two principles: (CV) and (IV). Cases such as Poisoning (I) and Robbery present potential counterexamples to (CV), based on the observation that in such cases, the agent can willingly and voluntarily perform an action they are coerced or compelled to do. To resolve this difficulty, I noted that in these cases, the driving force of the action does not stem from the threat or compulsion itself and proposed a revised principle (CV*), which includes a clause specifying that the agent's action must originate from the compulsion. I also argued that (CV*) successfully addresses a potential issue within the Aristotelian notion of a compulsory action, which requires that the action must stem from the external environment and not from the agent. Since (CV*) makes no reference to the cause of the agent's action, the objection that what is compulsory in the sense envisaged by Aristotle cannot be an action is not applicable to (CV*).

Regarding ignorance and involuntariness, I first observed that (IV) fails to accommodate Aristotle's distinction between non-voluntary and involuntary actions, wherein involuntary actions involve cases where the agent experiences regret. I then proposed a revised principle (IV*) based on the

situation as her spilling some gas, then her psychological condition becomes relevant. However, in that case, we cannot claim that she spilled the gas voluntarily. Therefore, in either scenario, these considerations do not pose a problem for (IV*).

observation that cases where an agent regrets their actions are typically cases where they would not have acted as they did if they had known better. Subsequently, I discussed whether the scenarios of Poisoning (II), Chess, and Gas Station can undermine (IV*). My contention is that (IV*) can be defended against such apparent counterexamples by restricting the range of relevant facts incorporated in (IV*) to particular types of facts, such as those pertaining to the nature or purpose of the action under the given description. Since (CV*) and (IV*) reflect the remarks in EE and NE while being immune to the complications previously discussed, I offer them as a plausible Aristotelian account of compulsion and ignorance related to the issue of involuntary actions.

Acknowledgements

I would like to express my gratitude to an anonymous reviewer from *Organon F* for valuable feedback. I am also thankful to Matthew Hanser for his insightful discussions on an earlier draft of this paper.

References

Anton, Audrey L. 2020. "Regret as a Reactive Attitude: The Conditions of Responsibility and Revision." In *The Moral Psychology of Regret*, edited by A. Gotlib, 45–75. London: Rowman and Littlefield.

Aquinas, Thomas. 1920 [1273]. *The Summa Theologica of St. Thomas Aquinas.* Second and Revised Edition, translated by Fathers of the English Dominican Province. London: Burns Oats and Washbourne. Online edition: https://newadvent.org/summa/

Aristotle. 2009. *The Nicomachean Ethics*, translated by D. Ross, Revised with an Introduction and Notes by L. Brown. Oxford: Oxford University Press.

Aristotle. 2013. *Eudemian Ethics*, translated and edited by B. Inwood and R. Woolf. Cambridge: Cambridge University Press.

Cooper, John M. 2013. "Aristotelian Responsibility." In *Oxford Studies in Ancient Philosophy* (Vol. XLV), edited by B. Inwood, 265–312. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679430.003.0008

Di Basilio, Giulio. 2021. "Aristotle on the Voluntariness of Self-control and the Lack of Self-control." *British Journal for the History of Philosophy* 29 (1): 4–23. https://doi.org/10.1080/09608788.2020.1748567

Frankfurt, Harry. 1969. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66 (2): 829–39. https://doi.org/10.2307/2023833

Heinaman, Robert. 1988. "Compulsion and Voluntary Action in the *Eudemian Ethics*." *Noûs*, 22 (2): 253–281. https://doi.org/10.2307/2215862

Locke, John. 1975 [1690]. *An Essay Concerning Human Understanding*, edited by P. H. Nidditch. Oxford: Oxford University Press.

Meyer, Susan Sauvé. 2011. *Aristotle on Moral Responsibility: Character and Cause*. Oxford: Oxford University Press.

Müller, Jozef. 2015. "Agency and Responsibility in Aristotle's *Eudemian Ethics*." *Phronesis* 60 (2): 206–51. https://doi.org/10.1163/15685284-12341284

Nielsen, Karen. 2007. "Dirtying Aristotle's Hand? Aristotle's Analysis of 'Mixed Acts' in the *Nicomachean Ethics* III, 1." *Phronesis* 52 (3): 270–300. https://doi.org/10.1163/156852807X208017

Williams, Bernard. 1979. "Internal and External Reasons," In *Rational Action: Studies in Philosophy and Social Sciences*, edited by R. Harrison, 17–28. Cambridge: Cambridge University Press.

Organon F 31 (4) 2024: 388–398
https://doi.org/10.31577/orgf.2024.21403
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

What Is so Bad about Permanent Coincidence without Identity?

Harold W. Noonan*

Received: 17 July 2024 / 10 September 2024 / Accepted: 18 October 2024

Abstract: 'What is so bad about permanent coincidence without identity?' (Mackie 2008: 163). This is the very question at the heart of the debate between pluralists and monists about constitution (Baker 1997, Fine 2003, Gibbard 1975, Johnston 1992, Lewis 1986, Thomson 1983). My answer to Mackie's question is that it contradicts a supervenience principle we all believe we know to be true. I approach this by considering three possibilities and the supervenience principles with which they conflict. One is somewhat politically controversial; the others are described by Wittgenstein (1967) and Dummett (1979). I focus on the possibility described by Dummett and the supervenience principle with which it conflicts. Our reaction to that possibility shows that we believe that supervenience principle to be true. But I argue that (as is obvious), it is inconsistent with permanent coincidence without identity. That is what is so bad about permanent coincidence without identity.

Keywords: Distinctions; supervenience; macrophysical; microphysical; pluralism.

* University of Nottingham
https://orcid.org/0000-0001-5538-5444
Department of Philosophy, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom.
Harold.noonan@nottingham.ac.uk

'What is so bad about permanent coincidence without identity?' (Mackie 2008, 1963). This is the question at the heart of the debate between pluralists and monists about constitution (Baker 1997, Fine 2003, Gibbard 1975, Johnston 1992, Lewis 1986, Thomson 1983). My answer to Mackie's question is that it contradicts a supervenience principle we all believe we know to be true. I approach this by considering three possibilities and the supervenience principles with which they conflict. One is somewhat politically controversial, the others are described by Wittgenstein (1967) and Dummett (1979). I focus on the possibility described by Dummett and the supervenience principle with which it conflicts. Our reaction to that possibility shows that we believe that supervenience principle to be true. But I argue that (as is obvious) it is inconsistent with permanent coincidence without identity. That is what is so bad about permanent coincidence without identity.

Imagine two small, no longer existent, material objects, one located in America and the other in Australia. They were always the same size, weight and colour. They were always composed of the same type of stuff. Chemists could find no difference at the level of chemical investigation, nor physicists at a more fundamental level. They were always composed of exactly the same type of fundamental particles, arranged in exactly the same way. For short, they were always intrinsically *microphysically indistinguishable* (i.e., indistinguishable with respect to the satisfaction of such predicates as 'contains an atom of carbon' which refer to and quantify over only microphysical entities – molecules, atoms and sub-atomic particles, the properties and relations these possess and the relations among them).

Yet they differed. Although this never happened, if the American one had been put under sufficient pressure it would have been destroyed, but the Australian one would have survived exactly the same pressure.

How can this be? Perhaps it might be that they differed in relational respects even though they were intrinsically microphysically indistinguishable.

Believers in homeopathy believe it is possible for two identical vials of liquid, one prepared in the proper fashion by succussion (shaking and dilution beyond (far beyond) Avogadro's limit) and one just taken from the tap, to differ in their (medical) properties, though no microphysical

examination of their intrinsic properties will reveal any differences.[1] They can differ in properties, their causal powers, just in virtue of their different histories of preparation, which they 'remember'.

Perhaps, then, the American object and the Australian object differed in how they would have responded to the pressure because they were created differently; like the two vials of liquid they had different histories though they were never intrinsically microphysically different.

Wittgenstein in *Zettel* (1967: section 208) describes an imaginary case that fits this model. Suppose seeds from two different types of plant are indistinguishable under the most careful microphysical investigation. Yet the seeds from one type of plant will develop differently from the seeds from the other – each will develop, if allowed to, into a plant of the type it has come from. So, two seeds which are not allowed to develop and so are microphysically indistinguishable throughout their existence, will differ in that one would develop differently from the other if allowed to grow. The explanation of this is that they have different histories.[2] Wittgenstein does not deny that the world could be this way.

---

[1] Homeopaths measure dilution on a 'C scale', diluting a substance by a factor of 100 at each stage. So a 6C dilution has the original substance diluted by a factor of $100^{-6}=10^{-12}$. Much higher dilutions are common and more dilute substances are considered by homeopaths to be stronger and deeper-acting. A popular homeopathic treatment for the flu is *Oscillococcinum*, a 200C dilution of Muscovy duck liver. The ingredients of a one-gram tube are: Active ingredient: *Anas Barbariae Hepatis et Cordis Extractum* (extract of Muscovy duck liver and heart) 200C 1×10^{-400} g, (less than the mass of a proton (1.67×10^{-24} g)). Inactive ingredient: 0.85 g sucrose, 0.15 g lactose (100% sugar). When Boiron (the company that makes Oscillococcinum) was asked if it was safe, they replied: 'Of course. There's nothing in it.'

[2] Noonan (2015) notes this case and presents an argument against the pluralist similar to the one developed below. But Noonan (2015) does not emphasise the crucial points highlighted below (see fn. 9 and following). Pluralists are committed to denying the supervenience of the macrophysical on the microphysical even when the microphysical is described using *all* the resources the pluralist allows himself in describing the macrophysical realm, appealing to relational as well as non-relational properties, modal and dispositional as well as categorical properties, sortal and 'sortalish' Bennett (2004, 341) properties as well as non-sortal properties. Secondly, pluralists must explain – whilst retaining the assumption that these predicates do indeed stand *for properties* of things in the world – how such predicates as 'would survive

But let us suppose that the American and Australian objects did not have such a difference in their histories – at least, if their backgrounds are described in the most explicit detail at the microphysical level no such difference appears. At this level of description not only they themselves but their surroundings and the circumstances of their origination were indistinguishable. In short, they were microphysically indistinguishable in all respects, extrinsically as well as intrinsically, and yet if the American one had been put under sufficient pressure it would have been destroyed, but the Australian one would have survived exactly the same pressure. (Of course, this won't be true. One is in America the other in Australia. So the particles composing one are a certain distance from the particles composing the Washington Monument, say, whilst the particles composing the other are not that distance from the Washington Monument or anything similar. But this relational difference is no help in resolving our puzzle so we can imagine it away and suppose that to whatever distance we care to go their environments are microphysically indistinguishable.)

Can we make sense of this? I think we can. Consider the following possibility.[3] There is a substance as sweet as sugar which is found, by giving minute samples to human tasters, to be a mixture of two types of substance, one twice as sweet as sugar, the other tasteless. However, no way can be found to identify the type of one of the minute samples without appeal to human tasters. The failure of intensive investigations to find differences supports the hypothesis that this inability is *absolute* – no way could be found except by the use of human tasters to distinguish the two types; it is not simply that the difference cannot be found by us or will not be found by us – there is no difference. Hence two samples in fact never tasted might be indistinguishable at the microphysical level in all respects, intrinsic and extrinsic, past, present and future, dispositional and modally as well as categoritically, and yet differ in that one *would have* tasted very sweet if tasted,

being crushed' stand for properties which do not supervene on the microphysical although, for example, the predicate 'would taste sweet if tasted' (see Dummett's example following) stands for a property which does.

[3] See Dummett (1979, 14). Dummett's interest is in the primary/secondary quality distinction, the relevance of his example to the topic of this paper has not hitherto been noted as far as I am aware.

the other tasteless. Of course, they won't be extrinsically indistinguishable, because one say might be near the Washington Monument and the other near no such structure, but again this difference is no help in resolving the puzzle, so we can imagine it away.

Of course, we don't believe (now) that this is how the world works, but it is not contradictory to suppose it works this way.[4] So perhaps we can suppose our American and Australian object were so related: if the American one had been put under sufficient pressure it would have been destroyed, but the Australian one would have survived exactly the same pressure. So, of course, one could never deduce from the most detailed microphysical description of them, their surroundings, and their history, which would have been destroyed and which would have survived.

This sounds like magic. In fact, it sounds more like magic than the homeopathic hypothesis which after all, its defenders claim to be science.[5] For since, according to practitioners of homeopathy, what makes the difference in the powers of two microphysically intrinsically indistinguishable vials of liquid is their method of preparation, homeopathy is testable in randomized, placebo-controlled trials. In fact, it has been tested.[6] So we might say, appealing to the criterion of falsifiability as a criterion of the scientific, that the suggestion that the American and Australian objects might have differed at the macrophysical level in the way described (would have responded differently to identical pressures), yet in fact never differed at all ever in any way microphysically, intrinsically or extrinsically, is, unlike the homeopathic hypothesis, not a scientific but a merely magical hypothesis. At any rate, we are sure that the world is not like this.

---

[4] And Dummett's speculation is not flat out inconsistent with every formulation of the supervenience of the macrophysical on the microphysical. It is not inconsistent, for example, with weak global supervenience (Sider 1999), which can only be counter-exemplified by a *pair* of worlds.

[5] The British Homeopathic Society, for example. NHS England withdrew funding for homeopathic medication in 2017, the British Homeopathic Society mounted a legal challenge which it lost in 2018. There were similar developments elsewhere. In France, funding was withdrawn in 2021.

[6] That is why governments are withdrawing funding.

I hope I have created some puzzlement by describing the case of the American and Australian objects. How could there have been two such objects so similar in all respects (microphysically) and yet so different (macrophysically)? At least, even if there could have been, it seems this could only have been so because the world could have been radically different from the way we know it actually is.

But, course, pluralists can say, nothing could be further from the truth. The explanation of the macrophysical difference can be quite simple. The American object was a statue of the infant Goliath (call it 'American Goliath'), and the Australian object was a lump of clay (call it 'Australian Lumpl'), at all times coincident with a statue qualitatively indistinguishable from American Goliath.[7] So, of course, they were microphysically indistinguishable in general terms[8] throughout their existence not only intrinsically but also in their surroundings and the circumstances of their origination. And of course, the American object, the statue, would have been destroyed if it had been put under pressure, e.g., rolled into a ball, and the Australian object, the lump of clay, would not. Statues and lumps of clay have different persistence conditions.

This, of course, is what we all *say*. But if we think that the two differed because it is correct to say of the one 'it would have been destroyed if rolled into a ball' and not correct to say this of the other, we must say that spatially separate material objects might be *macrophysically* distinguishable in their general unrealized capacities (e.g., for resisting destruction) though indistinguishable *microphysically* in all general respects at all times intrinsically and extrinsically and, of course, also *microphysically* indistinguishable dispositionally, modally and in all sortal respects, that is, microphysically indistinguishable in all respects pluralists think that macrophysical

[7] See Gibbard (1975) for the original story of (permanently coincident) Goliath and Lumpl. The puzzle is not restricted to artefacts. Kripke gives a well-known unpublished example of a (rootless) flowering plant, which in fact never flowers, and its permanently coincident stem.

[8] Of course, they were composed of *numerically* distinct particles and externally related to numerically distinct particles, so they were distinguishable microphysically in *non*-general terms, unlike American Goliath and American Lumpl.

objects, whether or not coincident, can differ.[9] Pluralists,[10] who think that in the Gibbard case, the all-time-coincident statue Goliath and lump of clay, Lumpl, are numerically distinct material objects differing in, e.g., their general macrophysically specifiable modal properties, are committed to this (and not just in cases involving artefacts, of course; recall Kripke's plant, fn. 7). If coincident material objects differ in general respects in this way macrophysically though indistinguishable microphysically in all general respects, so, of course, do many indistinguishable spatially separate ones. If coincident American Goliath and American Lumpl do in fact so differ then there are in fact microphysically indistinguishable *spatially separate* objects which so differ: the two statues are modally indistinguishable, ditto the two lumps, but American Goliath differs modally at the macroscopic level from American Lumpl according to the pluralist; so then must American Goliath and Australian Lumpl. So the puzzle I began with is resolved by the pluralist.

But, of course, this only highlights the challenge to the pluralist brought out by noting that it is not only a failure of the supervenience of the macrophysical on the microphysical in the special case of *coincident* objects he is committed to. The challenge is to explain how the case of American

[9] I.e., indistinguishable in respect of the satisfaction of all dispositional, modal and sortal predicates which refer to and quantify only over microphysical entities, both those within the macrophysical objects and those outside them, the properties these possess and the relations between them (e.g., 'contains an atom which could combine with two atoms of oxygen to form a molecule of carbon dioxide').

[10] See References. Pluralists are often said to face a 'grounding problem'. What grounds the non-identity of the all-time coincident statue and lump? That is not my question. Nor is my question how they can differ in sortal properties without differing non-sortally, nor how they can differ in modal or dispositional ways without differing non-modally or non-dispositionally. My observation is that if the pluralist appeals to a difference in macrophysical dispositions to explain the differences of the coincident statue and lump he must accept also that there is a vast number of cases of spatially separate (and obviously numerically distinct) objects differing merely in general macrophysical respects, without any basis for the difference in microphysical respects. The distinction I am emphasising is between macrophysical and microphysical properties, whether intrinsic or extrinsic, categorical or dispositional, modal or non-modal, sortal or non-sortal.

Goliath and Australian Lumpl differs from the case of Dummett's two lumps of 'sugar', macrophysically distinguishable, but composed of microphysically indistinguishable samples. Or else, to allow that for all he knows the world does contain cases of this latter type too. I assume the latter is a step too far, even for the pluralist, and I have no idea how the pluralist can explain the difference between the cases. The pluralist must allow that the concepts he applies at the macrophysical level – of sorts, identity conditions, essences etc., are also applicable at the microphysical level. But then if the statue and the piece of clay are in all ways, including these ways, microphysical indistinguishable, as the lumps of 'sugar' are, what can he say to explain his claim that the statue and the clay are macrophysically different whilst denying that the 'sugar' lumps can be? It will not help if the pluralist commits to pluralism all the way down. Even if coincident with any atom there is a plenitude of other particles differing only in sortalish ways, since all of these are contained within the putatively distinct but microphysically indistinguishable statue and piece of clay, it leaves it still a mystery how they can be macrophysically distinguishable if Dummett's two lumps are not. What the pluralist needs to do, to maintain a distinction between the case of the statue and the clay and the case of Dummett's two 'sugars' lumps, is to explain the relevant difference between the predicates 'would have been destroyed if squashed into a ball' and 'would have tasted sweet on the tongue of a normal human being' whilst retaining the pluralist assumption that each predicate denotes a property possessed by macrophysical objects. To put the point another way: the pluralist thinks that there is a property denoted by the predicate 'would have been destroyed if squashed into a ball' and another by the predicate 'would have tasted sweet on the tongue of a normal human being'. He thinks the first can vary between microphysically indiscernible objects but the second cannot. So he needs to explain why these two properties differ in this way.[11] He needs not merely

[11] The monist will explain the intuition that the statue would have ceased to exist if squashed without any appeal to a property the predicate denotes. The statue would cease to exist if squashed he will say because it will not continue to exist with a different shape. It is a statue. No statue undergoes a change of shape. That is a de dicto necessary truth. It is never the case in any possible world that there is a statue which has different shapes at different times. So since the statue is a statue it will

to label the predicates differently using the familiar terminology of the pluralist, e.g., as 'sortalish' versus 'non-sortalish'. He must make clear how this difference explains how the properties he assumes the predicates denote can vary with respect to supervenience on the microphysical. I can see no way the pluralist can do this. That is why pluralism must be rejected. It entails a magical hypothesis inconsistent with how the world works. It entails something we (now) *know* to be false. To put the point another way: you are completely confident you will never encounter cases of Dummett's type. But how can you be if you are a pluralist? For according to you, cases of non-coinciding, microphysically indistinguishable but macrophysically distinguishable objects are everywhere. So this is my answer to Mackie's question: 'What is meant to be so bad about permanent coincidence without identity?' (2008, 163): it entails that the world is full of cases of non-coinciding objects microphysically indistinguishable in all respects but macrophysically distinguishable – which we are certain we know to be false. We are certain that there can be no real differences at the macro level without differences at the micro level.

The rejection of pluralism requires us to acknowledge that when we say of the American object 'it would have been destroyed if rolled into a ball' we cannot be ascribing to it the same property we are denying of the Australian object when we say 'it would not have been destroyed if rolled into a ball' since there is no real difference of properties between them. In line with this Lewis (1986) says that modal predication is inconstant – the reference of a token of a modal predicate depends on the subject term to which it is attached (as the reference 'is so-called because of his size' differs when attached to 'Barbarelli' and 'Giorgione').[12] Hence, though both the American and Australian objects are both statues and lumps of clay, when we say

not continue to exist if squashed. Our intuition here that if we squash the clay the statue will cease to exist is explained by our knowledge of a de dicto necessary truth (see Lewis 1986, 193 on rivers and restaurants).

[12] There are other options consistent with monism. (i) Either the statue or piece of clay, or both, does not exist (Van Inwagen 1995). (ii) One of them, perhaps the statue, is not a material object, but say, a mathematical object, maybe a function from times to pieces of clay, so that they are not microphysically indiscernible. (iii) Though there *are* two coincident material objects in America where American

of the American object, picked out as a statue, 'it would have been destroyed' this is ascribing to it a different property from the one denied of the Australian object when it is said of it, picked out as a lump of clay, 'it would not have been destroyed'. So what we say can be explained without supposing any real difference in properties between them.

The inconstancy of modal predication is a linguistic hypothesis. But given how we are certain that the world is; that is, that spatially separate material things cannot differ in general macrophysical respects without differing in some general way, at some time, intrinsically or extrinsically, dispositionally, modally or sortally, in microphysical respects, a linguistic hypothesis is a necessary one (unless we choose from the options in fn. 12). How this inconstancy is to be explained is another matter. Lewis emphasises that his own (counterpart theoretic) proposal is just one among many. But, however we choose to explain it, pluralism should be rejected since it entails a magical hypothesis inconsistent with how we believe we know the world is.[13]

Goliath is (and in Australia where Australian Lumpl is), they do not differ in their macrophysically specifiable modal properties (maybe because there aren't any (Quine 1976a, Sider 2008)), or are only weakly discriminable, i.e., though there is a formula with two free variables satisfied by the two objects taken in either order, but not by either object taken twice, there is no formula in one free variable satisfied by one and not the other, so they are like Black's two spheres (Quine 1976b). I leave these suggestions aside. They are alternative ways of rejecting pluralism. But what matters is that they are ways of rejecting pluralism.

[13] A final option, not really distinct (i.e., only semantically distinct) from the Lewisean one (see Lewis 1986: sec. 4.3),) is to adopt Lewis' realism about possible worlds and to identify material things including the piece of clay and the statue with transworld individuals unified by counterpart relations – the piece-of-clay counterpart relation for the first and the statue counterpart relation for the second. Then they will be microphysically distinguishable because their microphysical parts will also be transworld individuals, so there will be no counterexample to the supervenience principle. Dummett's two lumps of 'sugar' will also be transworld individuals. But both will consist of world stages unified by the lump-of-'sugar' counterpart relation.

References

Baker, L. Rudder. 1997. "Why Constitution is not Identity." *Journal of Philosophy*, 94: 599–621. https://doi.org/10.2307/2564596

Bennett, Karen. 2004 "Spatio-temporal coincidence and the grounding problem." *Philosophical Studies*, 118 (3): 339–71. https://doi.org/10.1023/B:PHIL.0000026471.20355.54

Dummett, Michael. 1979. "Common Sense and Physics." In *Perception and Identity*, edited by G.F. Macdonald, London: MacMillan, 1–40. https://doi.org/10.1007/978-1-349-04862-5_1

Fine, Kit. 2003. "The Non-Identity of a Material Thing and its Matter." *Mind*, 112: 195–234. https://doi.org/10.1093/mind/112.446.195

Gibbard, Allan. 1975. "Contingent Identity." *Journal of Philosophical Logic*, 4: 187–221. https://doi.org/10.1007/BF00693273

Johnston, Mark. 1992. "Constitution is not Identity." *Mind*, 101: 89–105. https://doi.org/10.1093/mind/101.401.89

Lewis, David. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.

Mackie, Penelope. 2008. "Coincidence and Identity." *Royal Institute of Philosophy Supplement*, 62: 151–176. https://doi.org/10.1017/S1358246108000623

Noonan, W. Harrold. 2015. "Against Strong Pluralism." *Philosophia*, 43 (4): 1081–87. https://doi.org/10.1007/s11406-015-9640-4

Quine, W. V. O. 1976a. "Three Grades of Modal Involvement." In *The Ways of Paradox, and Other Essays*. 158–76. Cambridge, Massachusetts: Harvard University Press.

Quine, W.V.O. 1976b. "Grades of Discriminability." *The Journal of Philosophy*, 73: 113–16. https://doi.org/10.2307/2025739

Sider, Teodore. 1999. "Global Supervenience and Identity across Times and Worlds." *Philosophy and Phenomenological Research*, 59: 913–37. https://doi.org/10.2307/2653562

Sider, Teodore., 2008. "Yet Another Paper on the Supervenience Argument against Coincident Entities." *Philosophy and Phenomenological Research*, 77: 613–24. https://doi.org/10.1111/j.1933-1592.2008.00212.x

Thomson, Judith Jarvis. 1998. "The Statue and the Clay." *Noûs*, 32: 148–73. https://doi.org/10.1111/0029-4624.00094

Van Inwagen, Peter. 1995. *Material Beings*. Ithaca, New York: Cornell University Press. https://doi.org/10.7591/9781501713033

Wittgenstein, Ludwig. 1967. *Zettel*, translated by G.E.M. Anscombe, edited by G.E.M. Anscombe and G.H. von Wright. Oxford: Blackwell.

Organon F 31 (4) 2024: 399–410
https://doi.org/10.31577/orgf.2024.31404

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Is Extensible Markup Language Perspectivist?

Timothy Tambassi*

Received: 18 July 2024 / Accepted: 15 November 2024

Abstract: Both Extensible Markup Language [XML] and Formal Ontologies [ISOs] have something to do with partitioning. XML partitions data using elements. FOs partition domains using representational primitives. On the basis of the partitioning of FOs, the philosophical debate has outlined an epistemological view about FOs, namely perspectivism. For perspectivism, partitioning a domain means making a mental distinction between those entities on which we focus and those that fall outside our interest. This partitioning provides a perspective on the domain. In Tambassi (2023) it has been argued that perspectivism is an underlying assumption of FOs. In this paper, I investigate whether the same is true of XML: that is, whether and how XML is perspectivist. I begin by defining FOs and presenting the main claims of perspectivism in order to discuss how these claims apply to FOs. Then I move on to XML, showing both the perspectivism of XML and how the claims apply to XML. The purpose is therefore purely speculative. Discussing whether XML is committed to perspectivism may help to clarify some of the theoretical assumptions of this markup metalanguage. More generally, the idea is that since the creators of markup languages develop these languages under the guidance of some theoretical assumptions, for

* Ca' Foscari University of Venice
https://orcid.org/0000-0001-6460-5920
Ca' Foscari University of Venice, Department of Philosophy and Cultural Heritage, Malcanton Marcorà - Dorsoduro 3484/D, Calle Contarini, 30123 Venezia, Italy.
timothy.tambassi@gmail.com

the sake of methodological accuracy these assumptions should be subjected to critical analysis rather than remain implicit and unexamined.

Keywords: Philosophy of language; perspectivism; XML; partitions; markup languages; epistemology.

[1] If someone were to argue that Extensible Markup Language [XML] and Formal Ontologies [FOs] have little in common, they would have many strings to their bow. The most important one, in my view, is this. As its name suggests, XML is a markup language – or rather, a metalanguage that allows users to define their own customized markup languages (Attenborough 2003). FOs are neither languages nor metalanguages; they are artifacts specified by ontological languages (Gruber 2009). And XML is not even one of those languages. As for the "little" that XML and FOs have in common, there is one similarity that caught my attention. Both XML and FOs have something to do with partitioning. XML partitions data using elements. FOs partition domains of interest using representational primitives. Precisely on the basis of this partitioning performed by FOs, the philosophical debate has outlined an epistemological view about FOs, namely perspectivism.[1] For this kind of perspectivism – which does not coincide with perspectivism in the philosophy of science[2] – partitioning a domain means making a mental distinction between those entities on which we focus and those that fall outside our (domain of) interest. According to this view, this partitioning provides a perspective on the domain. Moreover, according to perspectivism, whatever domain we consider, there can in principle be multiple, equally valid, and overlapping perspectives on the same domain.

[2] Now, in Tambassi (2023) it has been argued that perspectivism is not just one of the philosophical views that populate the debate on FOs, but an underlying assumption of FOs. In other words, FOs are perspectivist. In this paper I investigate whether the same is true of XML. I begin by defining FOs and presenting the main claims of perspectivism. The idea is

[1] See Munn and Smith eds. (2008); Keen et al. (2012); Tambassi (2022, 2023).

[2] On perspectivism in philosophy of science, see Chakravartty (2010); Agazzi (2014, 2016); Massimi (2022); Massimi and McCoy eds. (2020).

not to prove the perspectivism of FOs, but rather to show how these claims apply to FOs. This is also to avoid any overlap with Tambassi (2023). I then move on to XML, showing both the perspectivism of XML and how the claims (of perspectivism) apply to XML. The argument is based on a parallel between FOs and XML. More specifically, the facets of perspectivism about FOs that I present in Sects. [5–9] correspond to the facets of perspectivism about XML that I present in Sects. [11–15]. This is not intended to exhaust the ways in which perspectivism relates to FOs and XML, nor the debate about FO and XML partitions. The only aim is to clarify whether and how XML is perspectivist. On the grounds that XML and FOs have little in common, it is not even ruled out that perspectivism might apply differently to FOs and XML.

[3] The purpose of this article is therefore purely speculative. I believe that discussing whether XML is perspectivist may help to clarify some of the theoretical assumptions of this markup metalanguage. More generally, the idea is that since the creators of markup (meta)languages develop those languages under the guidance of some theoretical assumptions, for the sake of methodological accuracy these assumptions should be subjected to critical analysis rather than remain implicit and unexamined. The focus on XML is not accidental. First, XML is still widely used, and there are many other markup languages based on XML. This means that the present critical analysis is, at least in principle, extendable to other markup languages. Second, XML not only supports the exchange of data, but is also both human- and machine-readable. In other words, XML – like FOs – supports communication between humans, between humans and machines, and between machines (Goy and Magro 2015). And while supporting this communication is certainly not the prerogative of XML and FOs alone, we cannot even rule out the possibility that determining whether XML is perspectivist may also shed new light on some of the theoretical assumptions behind such communication.

[4] According to Gruber (2009),[3] FOs are sets of representational primitives used to model a domain of knowledge. Primitives are instances,

[3] For further competing definitions of FO, see Gruber (1993); Guarino and Giaretta (1995); Borst (1997); Studer et al. (1998); Uschold and Jasper (1999); Noy and McGuinness (2003); Tambassi and Magro (2015).

classes, relations, and properties.[4] Instances are the basic units of FOs. Classes, which may contain sub-classes and/or be sub-classes of other classes, are sets of instances that share common features. Relations represent the ways in which both classes and instances interact with each other. Properties describe the various features of a class and of its instances. Perspectivism is an epistemological view about FOs that makes three distinct claims: 1) there are different ways of representing a domain (of interest); 2) there can be multiple, equally valid, and overlapping perspectives on a domain; 3) a perspective partitions a domain, drawing a mental division between the entities we focus on and those that are outside our interest. In the next five sections, I explore how these claims relate to and apply to FOs.

[5] The first claim, one might say, is not exactly theoretical: it simply helps us to define what perspectivism refers to, namely representation(s) and domain(s) of interest. Regarding the domain of interest, the first claim does not impose any restriction. Accordingly, we can represent any domain of interest. Regarding representations, the first claim does not exclude any way of representing a domain of interest (Munn and Smith eds. 2008). Therefore, even though perspectivism is about FOs, FOs are not the only way to represent a domain. This means that, at least in principle, perspectivism does not apply only to FOs, which is the theoretical basis for our investigation of perspectivism about XML.

[6] The second and third claims concern the partitioning of the domain, which in terms of FO development involves at least two stages: outlining the entities that populate the domain *and* systematizing these entities within representational primitives (Noy and McGuinness 2003). Although this division is intended to preserve the possibility that different systematizations may follow from the same outline, we should also note that the two stages are interrelated. Indeed, the second stage is based on the first. However, outlining the entities without systematizing them would mean not including them in a FO, which would obscure why the partition is specifically concerned with FOs. For this reason, even though the process of partitioning the domain seems to be more related to the first stage, I consider the second stage to be part of the same process.

[4] For competing lists of representational primitives, see Noy and McGuinness (2003); Jaziri and Gargouri (2010); Laurini (2017).

[**7**] Regarding the first stage (outlining the entities that populate the domain), the third claim provides the main guidance on how to outline the entities. (The question of whether such a stage also deals with the definition of the domain is discussed in Tambassi 2023). This guidance does not impose any constraint on how to (cognitively) partition the domain and, thus, on the entities resulting from the partition. In other words, according to the third claim, no partition or entity is excluded in principle. This is also confirmed by the second claim, which further specifies that we can also partition the same domain in multiple ways. To summarize: each partition results in a list of entities (in principle a different list for each partition); the partitions of a domain can be multiple; there are no restrictions on domains, partitions, and entities.

[**8**] Regarding the second stage (systematizing the entities within representational primitives), both the second and the third claim do not impose any restriction. Suppose, then, that the partition of a domain results in a list of entities that includes "love". How do we systematize such an entity? Under one of the four representational primitives, of course. Therefore, we could regard "love" as a class with different loves as instances or sub-classes, as the relation "is in love with" between two instances, as the property "being in love" of an instance, and so on. Much depends on how we model the domain (Gruber 2009) – that is, on the perspective (on the domain) that we adopt. Not restricting the systematization of entities is perspectivism's response to the (potential) multiplicity of perspectives on the same domain that we might adopt and then represent (Tambassi 2023). And to those who argue that perspectivism must face a multiplicity of perspectives because it assumes such a multiplicity, a perspectivist could reply that since perspectives are cognitive partitions, their multiplicity can hardly be denied.

[**9**] Perspectivism about FOs cannot go beyond the limits set by the IT advances of FOs. (Of course, this does not exclude the possibility that FOs may go beyond these limits in the future, also because of perspectivism.) To say the opposite would be to say that perspectivism is not about FOs. Now, regardless of the language in which FOs are expressed, the current debate generally recognizes four representational primitives: instances, classes, relations, and properties. Each representational primitive can, in

principle, include any entity. Systematization thus results in entities of these representational primitives. Returning to our example, this means that, depending on how we systematize "love", it will still result in one of these representational primitives. Otherwise, "love" would not be an entity of the FO, which would contradict the results of the domain partition. Systematization, some critics might say, proves that perspectivism is inconsistent: for it both rejects any constraint on systematization and accepts constraints on the representational primitives within which entities are systematized. But I am not sure that the inconsistency holds. First, the constraints on representational primitives seem to refer to the limits of FOs, not to those of perspectivism. (At most, the limits would refer to perspectivism in a transitive way, that is, as far as it deals with FOs; but perspectivism in itself, as the simple sum of the three claims of Sect. [4], does not have these limits.) Second, the fact that representational primitives are currently four, but potentially include any entity, should, at least in principle, limit, if not eliminate, any constraint.[5]

[10] According to Dykes (2005), the purpose of XML is (to serialize and) to represent data (more precisely, documents). The building blocks of any XML document are elements, which consist of content, tags (markups), and

[5] Because of this, and because the advances in FOs may modify the current representational primitives (and thus bring new limits and possibilities to the systematization of entities), Cumpa (2019, 149) considers FOs to be "variantist", meaning that their structure «can change due to certain practical preferences». Although I agree with Cumpa about the variantism of FOs, I think we need to be careful here not to equate variantism about representational primitives with variantism about the entities they contain, and perhaps also variantism with perspectivism. With respect to representational primitives, variantism concerns the fact that the list of representational primitives might vary due to the advances in FOs. Therefore, future systematizations might lead to entities of other representational primitives. With respect to entities, variantism concerns the fact that they can vary from FO to FO. FOs can thus model different domains of inquiry, support different perspectives on those domains, and so on. All this should also explain how variantism and perspectivism differ. The former is ontological, in that it is about representational primitives and their entities as such; the latter is epistemological, in that it is about the perspectives (on any domain) that we can represent by means of entities belonging to representational primitives.

attributes, but can also be empty or contain other elements. Content indicates data (sometimes named and represented by entities). Tags have the primary function of labeling data within the document. A tag may also have any number of attributes (name-value pairs) that provide further information about the element for which they are declared. Discussing whether XML is perspectivist means understanding the three claims of perspectivism (see Sect. [4]) as not just about FOs, and showing whether and how XML is committed to these claims.

[11] Now, the first claim of perspectivism is that there are different ways of representing a domain. The claim thus refers to representation(s) and domain(s). Regarding representation(s), Sect. [5] shows that FOs are only one way of representation and that the first claim does not exclude that XML could be another. I would add here that XML does not claim to be the only way. Rather, as its name suggests, XML is extensible: that is, we can create other markup languages from XML (Canducci 2022), and thus new ways of representing. This means that, in terms of representations, XML makes the first claim of perspectivism. What can be represented? Perspectivism does not exclude any domain, whereas XML is about any arbitrary data (Dykes 2005). Therefore, perspectivism applies to XML if and only if "(any arbitrary) data" (henceforth, "data") is a proper or improper subset of "any domain". However, considering "data" as a proper subset of "any domains" (i.e., if x is an element of "data", then x is also an element of "any domain", but not *vice versa*) would imply that XML imposes a constraint on perspectivism: the domain(s) must be domain(s) of data. Conversely, considering "data" as an improper subset of "any domain" (i.e., "data" and "any domain" have the same elements) would remove such a constraint, implying that XML adopts the first claim of perspectivism. That said, I think that both alternatives approach the issue in the wrong way. "Data" and "any domain" are here the two sides of the same coin: for data can refer to, or be data of, any domain. For this reason, not only does the first claim apply to XML, but XML also assumes the first claim of perspectivism.

[12] Perspectivism is specifically about partitioning domains. There are the second and third claims to witness this. Now, Sect. [6] says that FO partitioning involves two consecutive stages: outlining the entities that

populate the domain and systematizing these entities within representational primitives. I think that both stages apply to XML as well, although there is a difference regarding the second stage. Indeed, XML has no representational primitives within which to systemize entities; rather, the XML systematization is done through elements consisting of content, tags, and attributes (name-value pairs). As for the first stage, the only element of confusion can be represented by "entity", which has a technical meaning in XML (see Sect. [10]) and a non-technical meaning in the two stages of the partition of the domain and in the third claim of perspectivism. However, since the non-technical meaning also concerns XML partitioning, I will only refer to the non-technical meaning of "entity" from now on. This should avoid any confusion.

[13] Also with regard to the first stage (outlining the entities that populate the domain), the most I can add to what is said in Sect. [7] is this. The fact that XML can describe any arbitrary data should be the reason why XML has, at least in principle, no restrictions on entities, partitions, and domains, as well as on multiple partitions of the same domain. (And insofar as the building blocks of XML are elements, those elements should have no restrictions either.) However, to show specifically that both the second and the third claims apply to XML, and vice versa, suppose that the domain consists of the starting grid of a Formula 1 race. We can certainly represent such a domain by an XML document that focuses on entities such as drivers and drivers' cars, and not on entities such as (best) lap times. But nothing prevents us from representing the domain by another XML document that focuses on entities such as drivers and (best) lap times, leaving aside entities such as drivers' cars. These two XML documents provide two different perspectives on a domain. This means that we can provide a perspective on a domain by using XML, but also that XML can provide a perspective on a domain, because XML assumes the third claim of perspectivism. Moreover, to the extent that the XML documents provide two examples of multiple perspectives on the same domain (resulting in different lists of entities), we could also say that we can provide different perspectives on the same domain by using XML, but also that XML can provide different perspectives on the same domain, because XML assumes the second claim of perspectivism.

[14] The second stage (systematizing these entities within representational primitives) concerns the systematization, on which neither the second nor the third claim imposes any restriction. Now, since XML assumes both, and the third claim imposes no constraint on the entities, any XML document can in principle include and (thus) systematize any entity. But how can entities be systematized? Sect. [12] states that XML systematization is done by means of elements consisting of content, tags, and attributes (name-value pairs). Sect. [10] adds that each element can also be empty or contain other elements. I further state here that the entities resulting from the first stage are ultimately systematized as contents and/or attribute values of an XML document. In other words, regardless of the entities resulting from the partition, these entities result in contents and/or attribute values. What about tags and attribute names? Tags label contents, and attribute names do the same for attribute values. This means that if we were to systematize the partition of a domain resulting in a list of entities that includes "love" by means of an XML document, then "love" would be a content or an attribute value within such a document – regardless of the tag or the attribute name we label "love". The choice of systematization of the elements, as well as their contents, tags, and attributes, will depend on how we model the domain – that is, on the perspective (on the domain) that we adopt. Therefore, we can simply conclude this Sect. by re-reading the last two sentences of Sect. [8]. (Not restricting the systematization of entities is the response of perspectivism to the (potential) multiplicity of perspectives on the same domain that we might adopt and then represent. And to those who argue that perspectivism must face a multiplicity of perspectives because it assumes such a multiplicity, a perspectivist could reply that, since perspectives are cognitive partitions, their multiplicity can hardly be denied.)

[15] Like perspectivism about FOs, perspectivism about XML cannot go beyond the limits set by the standards defined by XML. To say the opposite would be to say that perspectivism is not about XML. Unlike FOs, however, XML cannot be separated from the language in which XML is expressed, since XML is itself a language, whose building blocks are elements consisting of content, tags, and attributes (name-value pairs). Contents and attribute values can, in principle, include any entity. And any

tags and attribute names can be associated with those contents and attribute values (respectively). Depending on how we systematize "love", the systematization will thus be one of contents and attribute values. Otherwise, we would contradict the results of the domain partition. As with perspectivism about FOs, this might prove the inconsistency of perspectivism about XML: it rejects any constraint on systematization and accepts constraints on where to systematize entities. For the very same reason stated in Sect. [9], I think there is no inconsistency. The constraint refers to the limits of XML, not to those of perspectivism. Moreover, since contents and attribute values are potentially inclusive of any entity, any constraint should be limited, if not eliminated.

[16] If the argument of Sects. [11–15] is correct, then I should have given the reason(s) to believe that XML is perspectivist. If so, then communication between humans, between humans and machines, and between machines based on XML would embrace the claims of perspectivism. Moreover, since many markup languages are based on XML (see Sect. [3]), I suggest that these languages might themselves be perspectivist. And the same can be said for whatever uses XML – unless that "whatever" and those languages provide some defenses against (the claims of) perspectivism. Finally, insofar as perspectivism applies differently to FOs and XML (simply because FOs and XML are different), we cannot even rule out the possibility that perspectivism also applies differently to the languages based on XML, or to whatever uses XML. Nor can we rule out the possibility of a combination of the various applications, resulting in mixed forms of perspectivism.

Funding

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

Agazzi, Evandro. 2014. *Scientific Objectivity and Its Contexts*. Cham: Springer. https://doi.org/10.1007/978-3-319-04660-0

Agazzi, Evandro. 2016. "Scientific Realism Within Perspectivism and Perspectivism within Scientific Realism." *Axiomathes*, 26: 349–65. https://doi.org/10.1007/s10516-016-9304-4

Attenborough, Mary. 2003. *Mathematics for Electrical Engineering and Computing*. Oxford and Boston: Newnes.

Borst, WN. 1997. "Construction of Engineering Ontologies." Ph.D. Thesis. University of Twente.

Canducci, Massimo. 2022. *XML per tutti*. Milan: Apogeo.

Chakravartty, Anjan. 2010. "Perspectivism, Inconsistent Models, and Contrastive Explanation." *Studies in History and Philosophy of Science*, 41: 405–12. https://doi.org/10.1016/j.shpsa.2010.10.007

Cumpa, Javier. 2019. "Structure and Completeness: A Defense of Factualism in Categorial Ontology." *Acta Analytica*, 34(2): 145–53. https://doi.org/10.1007/s12136-019-00380-6

Dykes, Lucinda. 2005. *XML for Dummies*. Hoboken, New Jersey: Wiley.

Goy, Anna, and Diego Magro. 2015. "What Are Ontologies Useful for? Encyclopedia of Information Science and Technology." In *Encyclopedia of Information Science and Technology*, edited by Mehdi Khosrow-Pour, 7456–64. Hershey, Pennsylvania: IGI Global. https://doi.org/10.4018/978-1-4666-5888-2.ch734

Gruber, Thomas R. 1993. "A Translation Approach to Portable Ontology Specifications." *Knowledge Acquisition*, 5(2): 199–220. https://doi.org/10.1006/knac.1993.1008

Gruber, Thomas R. 2009. "Ontology." In *Encyclopedia of Database Systems*, edited by Ling Liu and M. Tamer Özsu, 1963–65. Cham: Springer. https://doi.org/10.1007/978-0-387-39940-9_1318

Guarino, Nicola, and Pierdaniele Giaretta. 1995. "Ontologies and Knowledge Bases: Towards a Terminological Clarification." In *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, edited by N.J. Mars, 25–32. Amsterdam: IOS Press.

Jaziri, Wassim, and Faiez Gargouri. 2010. "Ontology Theory, Management and Design: An Overview and Future Directions." In *Ontology Theory, Management and Design: Advanced Tools and Models*, edited by Faiez Gargouri and Wassim Jaziri, 27–77. Hershey, Pennsylvania: Information Science Reference. https://doi.org/10.4018/978-1-61520-859-3.ch002

Keen, Emily, Chris Keen, and Simon Milton. 2012. "The Relevance of Perspectivism to the Task of Modularisation in Ontology Development." In *ACIS 2012 Proceedings*. 100. https://aisel.aisnet.org/acis2012/100

Laurini, Robert. 2017. *Geographic Knowledge Infrastructure: Applications to Territorial Intelligence and Smart Cities*. Amsterdam: ISTE-Elsevier.

Massimi, Michela. 2022. "Perspectival Ontology: Between Situated Knowledge and Multiculturalism." *The Monist*, 105 (2): 214–28. https://doi.org/10.1093/monist/onab032

Massimi, Michela, and CD McCoy, eds. 2020. *Understanding Perspectivism: Scientific Challenges and Methodological Prospects*. London and New York: Routledge.

Munn, Katherine, and Barry Smith, eds. 2008. *Applied Ontology: An Introduction*. Berlin: Ontos-Verlag.

Noy, NF, and DL McGuinness. 2003. *Ontology Development 101: A Guide to Creating Zour First Ontology*. Stanford, California: Stanford University.

Studer, Rudi, VR Benjamins, and Dieter Fensel. 1998. "Knowledge Engineering: Principles and Methods." *IEEE Transactions on Data and Knowledge Engineering*, 25 (1–2): 161–97. https://doi.org/10.1016/S0169-023X(97)00056-6

Tambassi, Timothy. 2022. "Ontological Perspectivism and Geographical Categorizations." *Philosophia*, 50: 307–20. https://doi.org/10.1007/s11406-021-00371-1

Tambassi, Timothy. 2023. "On Perspectivism of Information System Ontologies." *Foundations of Science*, 29: 571–85. https://doi.org/10.1007/s10699-023-09900-5

Tambassi, Timothy, and Diego Magro. 2015. "Ontologie informatiche della geografia. Una sistematizzazione del dibattito contemporaneo." *Rivista di estetica*, 58: 191–205. https://doi.org/10.4000/estetica.2752

Uschold, Mike, and Robert Jasper. 1999. "A Framework for Understanding and Classifying Ontology Applications." In *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5) Stockholm, Sweden, August 2, 1999*, edited by VR Benjamins, Balakrishnan Chandrasekaran, Asunción Gomez-Perez, Nicola Guarino, and Mike Uschold. https://ceur-ws.org/Vol-18/11-uschold.pdf

Organon F 31 (4) 2024: 411–431
https://doi.org/10.31577/orgf.2024.31405
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Acting as Process

Yuchen Guo*

Received: 14 March 2024 / Accepted: 17 November 2024

Abstract: What is acting? What type of action does an actor perform when portraying their character? Is it possible to conceptually distinguish acting from other types of activities? This paper aims to answer these types of questions about defining acting. My work involves two aspects. On one hand, I argue that the three current popular theories defining acting (the pretense theory, the display account, and the game model) are implausible; namely, acting cannot be reduced to a simple state or event. On the other hand, I argue that acting should be understood as a process and try to clarify the distinctive characteristics of this process.

Keywords: Acting; theater; process; pretense; display; game.

1. Introduction

Consider an actor portraying Oedipus, Hamlet, Dom Juan, Willy Loman, or any other character. To successfully play the roles, the actor should say and do what the character says and does—to act as if they were the character; moreover, the actor also often enters the inner world of the char-

* Lanzhou University
https://orcid.org/0000-0002-7441-0948
School of Philosophy and Sociology, Lanzhou University, TianShui Southroad 222, Chengguan District, Lanzhou, 730000, China,
yuchenguo17@gmail.com

acter or becomes emotionally involved in the character's situations. Generally, this type of behavior—portraying another individual—is called "acting." What is the nature of acting? What is it to portray the character? What type of behavior is acting?[1]

Currently, three competing theories aim to answer these questions about defining acting. One is the pretense theory proposed by John Searle and his adherents. According to this theory, acting is a kind of pretense: actors pretend to do and say what their characters do and say or pretend to be their characters (Searle 1975, Lewis 1978, Austin 1979, Walton 1990, and Alward 2009). Another alternative is David Saltz's game model: onstage acting is like a game: actors' onstage actions are derived from the goals associated with the roles they occupy (Saltz 1991). The most recent theory is James Hamilton's display theory of acting, according to which acting is a kind of demonstration: actors display and hide their own features to enable the audience to figure out the content of the narrative (Hamilton 2009, 2013).[2] [3]

[1] Does acting always occur in the context of theater? I do not think so. I admit that acting is typically embedded in the context of theater. However, this does not apply to all types of acting. Film actors also portray their characters. In role-playing therapy, the therapist plays a role to treat patients. In mental health education, students simulate patients to explore their feelings and attitudes. Indeed, most books about actor training or acting techniques mention issues unrelated to theater and script, such as body postures, private emotions, and respiration.

[2] Alward (2009) argues that Gregory Currie's account of fictional utterance can be applied to the context of theatrical illocution and implies a *sui generis* account of dramatic acting. According to this, the actors on stage perform *sui generis* illocutionary actions; that is, actors intend their audiences to imagine or make-believe the proposition expressed by their utterance. The point I defend also implies that acting requires a communicative intention; however, the actor does not intend audiences to imagine the relevant proposition but to imagine that the actor is identical to the character.

[3] Recent philosophy of theater in the analytic tradition concerns the nature of theatrical performance (Bennett 2020, Osipovich 2006, Stern 2014, Stern ed. 2017, Hamilton 2007, Krasner & Saltz eds. 2006, and Woodruff 2008). This involves acting-related issues. However, these issues are more closely related to the context of theater. Acting is discussed to better understand theater as an art form rather than

The first goal of this paper is to show that the three above-mentioned theories (the pretense theory, game model, and display theory) fail to explain some cases of acting, such as those occurring in a group rehearsal or preparation. Hence the three theories are not a plausible account of dramatic acting. The paper's other goal is to provide an alternative account of the nature of acting. I argue that acting should be understood as a process and characterize this process to distinguish acting from other human activities. In Section 2, I use real examples to clarify my position and explain how my point differs from other theories. In Sections 3 and 4, I rebut the three popular theories of acting and clarify my point that acting is a process.

2. Four ways of understanding

2.1 Sui generis understanding versus reductive understanding

We are curious animals and always attempt to understand, define, theoretically grasp, and interpret the external world. Generally, two popular methods—*sui generis* understanding and reductive understanding—help define and grasp a thing or concept.

First, consider reductive understanding. Understanding something reductively involves placing it in a well-known category and using characteristics of this category to interpret this thing. For instance, attempting to interpret archaeological objects often involves a reductive method: they are placed in a familiar category of artifacts and thereby accounted for. The understanding of a "*jue*," an ancient Chinese vessel, is an obvious example. A *jue* is a major archaeological find in Chinese history. It has an ovoid body supported by three splayed triangular legs, with a long curved spout on one side and a counterbalancing flange on the other side, one or two handles on the side, and two columnar protuberances on the top.[4] Generally, a *jue* is considered a traditional Chinese vessel used to warm wine and drink. That

acting as an intentional act. Yet, I suggest acting is a special type of activity occurring not only in the context of theater but also in other cases, such as filmmaking, psychotherapies, and teaching techniques.

[4] For a photo of a jue, visit https://en.wikipedia.org/wiki/Jue_(vessel)

is, this archaeological object is categorized as a drinking vessel. It follows that every component of a *jue* can be understood to be used to serve drinks: the flange catches the poured wine, the spout is used to take a sip, and the handles are for grasping. The characteristics of a drinking vessel are used to understand the *jue* as an archaeological find.

Many similar examples exist in biology. Faced with unidentified or newly found animals, a biologist must classify them—find their places in the system of biological classification. A relevant example is the debate about the classification of the Tully monster, an extinct genus of soft-bodied animal that lived during the Pennsylvanian geological period about 300 million years ago. Some studies show that the Tully monster is a vertebrate close to a lamprey. Yet, other biologists reject the identification of the Tully monster as a vertebrate, calling it a mollusk.[5] This suggests that the biologists appeal to a reductive understanding of this type of animal; the essence of this debate is to categorize the Tully monster as a vertebrate or a mollusk, two types of known phyla.

A *sui generis* understanding also exists about certain things. Understanding something in a *sui generis* way involves constructing a *sui generis* category and attributing it to a *sui generis* type rather than placing it in a known category. For example, in linguistic studies, languages unrelated to any other language are called "language isolates." Such language isolates, such as Basque, Korean, Sandawe, and Haida, are often not classified into well-known language families but made the only language in their own language family. Similar examples exist in biology. The platypus is a mammal native to Australia. Unlike other typical mammals, the platypus is egg-laying, duck-billed, beaver-tailed, otter-footed, venomous, and senses prey through electrolocation. These features make it difficult to classify into a well-known mammal family. Hence biologists no longer attempt to classify the platypus into well-known mammal families and genera but propose an Ornithorhynchidae family and Ornithorhynchus genus including only the platypus. The two examples appear to suggest that, apart from the reductive understanding, people also often appeal to a *sui generis* understanding

[5] For relevant debate, see https://en.wikipedia.org/wiki/Tullimonstrum#Classification.

to interpret some things: they do not place them in a well-known category but propose *sui generis* categories to classify them.[6]

2.2 State-approach understanding versus process-approach understanding

People sometimes use a distinctive state to explain certain things; however, people also sometimes use a distinctive *way* of combining different states to explain something. I call the former "state-approach understanding" and the latter "process-approach understanding."

Consider process-approach understanding. An obvious example is proposed by Peter Goldie, who claims that when trying to understand what cheque-writing is, you cannot use any single state to define it (Goldie 2012, 61). For example, one cannot say that cheque-writing is only a set of ink marks on paper—a publicly observable state. The latter is only an output of cheque-writing and can be reproduced by photocopying; yet photocopying is not called "cheque-writing." Thus, cheque-writing should be viewed as a dynamic process involving different states and unfolding over time.

Unlike process-approach understanding—using a distinctive process to understand something—state-approach understanding means that a kind of distinctive state defines and distinguishes something. For example, when seeing a red apple, the red conscious experience should be understood as a distinctive state—a visual state with qualitative feeling. It is certain that red conscious experience involves many parts: a memory trigger, visual attention, and an action tendency. However, the visual experience itself, a distinctive mental state, defines the nature of red conscious experience (Goldie 2012, 60–61).

Regarding human conscious experience, a debate exists concerning how to understand the nature of grief—a particular emotion—showing the differences between state-approach understanding and process-approach understanding. Traditional accounts imply grief is a distinctive mental state. For instance, the feeling theory identifies emotion with a feeling (James

[6] Evidently, one can place Basque in a category of language or the platypus in a category of animal, but that is too trivial and cannot help understanding and define something.

1884); the judgment theory identifies emotion with evaluative judgments (Nussbaum 2001); the perception-based theory identifies emotion with a particular perception (Roberts 2003). Yet, regardless of whether emotion is a perception, judgment, or feeling, all these theories imply that grief is a distinctive mental state (perception, judgment, and feeling are different types of mental state). Conversely, Peter Goldie (2012) argues that grief should not be understood as a distinctive state but a process with distinct stages that unfolds over time because grief has characteristics classifying it as a mental process rather than a mental state.

In this paper, I first suggest that acting should be understood in a *sui generis* way, not reductively. In Section 3, I defend this point and explain why other human activities such as pretense, display, and games cannot help understanding the nature of acting. Second, I suggest that acting should be viewed as a distinctive process involving distinct states rather than a single state. In Section 4, I defend that point and aim to explain what characterizes the process of acting.

3. Pretense, display, game, and acting

3.1 Preparation and rehearsal

Consider the following three cases:

Preparation: *A* is an actor ready to portray Hamlet, a tragic character, whom he has never tried to act out. After reading and analyzing the script, *A* attempts to imitate Hamlet's facial expressions and postures using a mirror, recite his lines, and, alone, act all the scenes in which Hamlet appears as if he were on stage with other actors.

Group Rehearsal: After *A* and other actors have prepared everything necessary for their performance and memorized their lines, they meet to act out the scenes as if they were on stage with their audiences. On encountering a difficulty, they stop their actions and repeat this scene several times to improve and perfect their acting.

Performance: When actually stepping onto the stage and facing the audience, via the routine developed and built in preparation and rehearsal,

A and other actors successfully portray these characters in *Hamlet* and show audiences the content of the play.

These cases can be distinguished. The case of preparation refers to what an actor does when preparing the roles; it involves imitating the character's expressions, repeating the lines, or acting out the scenes alone. Group rehearsal is a collective activity undertaken by a group of people; it involves all actors meeting to act out the scenes as if they were on stage. Performance is what actors do when they actually face audiences on stage. The aim of preparation and rehearsal is to build a model or routine to help actors successfully portray their characters and attract audiences in performance.

Preparation, rehearsal, and performance should all be viewed as acting because what actors do in all the three cases are to portray the character, to play the roles, and to act the scenes. They are doing and saying what other individuals say and do; they act as if they were their characters. For instance, when *A* repeatedly says and does what Hamlet says and does, such as fighting with Claudius or drinking poisoned wine, *A* neither believes he is Hamlet nor actually performs these actions. He aims to act out Hamlet rather than become Hamlet. Another reason for my point is that no qualitative distinction exists between preparation, rehearsal and real performance. Some actors say, "I didn't do well on stage. When I prepared the roles alone, I did it even better." Indeed, actors' actions in rehearsal and preparation may be better than in real performance (perhaps due to the pressure and stress). This also suggests rehearsal, preparation, and performance are types of acting.

In rehearsal and preparation, the actor's intentions differ from those in onstage performances. In performance, the actor could intend to display, pretend, and attract audiences, as stated by the pretense account, display theory, and game model. Yet, when preparing roles and rehearsing the scenes, the actor intends to find and remove imperfections, become familiar with the scenes, arouse emotions, overcome shyness, or do something else that aims to improve his portrayal of the character. The difference concerning the content of the intention reveals that these popular theories of acting incorrectly identify acting with performance but ignore rehearsal and preparation. I explain and defend this point in the next section.

3.2 Pretense, display, and game

Currently, three theories of acting (the pretense account, display theory, and game model) address the problem of the nature of acting. First, consider the pretense account and display theory. Both face the same problems. The pretense account is a family of theories that regard acting as a pretense (Searle 1975, Lewis 1978, Austin 1979, Walton 1990, and Alward 2009). The earliest version of the pretense theory is supposed to originate from John Searle's account of illocutionary actions in fictional discourse, which implies a point about theatrical discourse: onstage actors do not actually perform illocutionary acts but pretend to perform them. For example, when the actor says, "The king is dead," the actor pretends to assert this. Searle's point implies that acting is viewed as pretending-to-do; actors pretend to do and say what their characters do and say (Searle 1975, see also Walton 1990 and Alward 2009). Another version of the pretense theory says that acting is "pretending-to-be." That is, when actors portray their characters, their behaviors can be viewed as "pretending to be someone" (Osipovich 2006, 468). For example, when Hamlet sits down in the play, the actor can actually sit down. Here the action the actor performs is not pretending-to-do but pretending to be someone: the actor does not pretend to sit down but pretends to be Hamlet sitting down. Briefly, the pretense account implies that acting is understood as pretend play.

Unlike the family of pretense theories, the display theory of acting claims that dramatic acting is more like the display behaviors of individuals or collectives rather than pretend play because actors portray their characters *for their audiences* to observe their performances in particular ways. According to Hamilton, when portraying characters, actors display and hide their own features to enable the audience to figure out the content of the narrative. That is, actors choose some of their own features that are relevant to the characters and present them to the audiences so that they can grasp the content of a play. Acting is viewed as a display for audiences.

The pretense and display theories face the same problems. Neither theory explains why preparation and rehearsal are acting because these are neither pretense nor display. In preparation and rehearsal, actors intend to repeat the lines and become familiar with the play, find the problems and perfect their performance, and arouse emotions and overcome shyness.

Simply, actors intend to improve and perfect their performances. It is difficult to say that actors are conscious of pretending in such cases. They intend to neither pretend to do what the characters do nor pretend to be the characters. Additionally, actors do not want to display to their audiences at that moment. When someone approaches them, actors may even stop their acting since they do not want to show their imperfections or create spoilers for audiences. In preparation and rehearsal, actors focus on improving their performance—the technical aspect of acting—rather than consciously pretending or displaying for audiences.[7]

Finally, consider Saltz's game model, according to which, when portraying the character, what an actor does is to play a game; they adopt the goals associated with the roles of the character. Saltz considers that agents sometimes act on behalf of someone else: a secretary acts on behalf of a boss, and a ticket inspector acts on behalf of the public transport company. A secretary apparently has no personal reason to place an order to buy certain equipment, and an inspector has no personal reason to check tickets. The reasons for their action derive from their desire to fulfill the function of their roles within the organizations. When a person occupies the role of an inspector, she or he adopts the goal of the transport company—to check tickets and ensure no loss; when a person occupies the role of secretary, her or his goal is to convey the boss's ideas. Saltz calls the intentional states adopted by the agents acting on behalf of someone else "borrowed intention"—this derives from the goals associated with the roles they occupy (Saltz 1991, 38). Hence, Slatz claims that when portraying the characters, actors act on behalf of their characters; their intentions for acting derive from the goals associated with the roles of the characters they portray.

[7] Proponents of the display theory might suggest a distinction between display and display preparation. While acting in performance is display, acting in preparation and rehearsal is display preparation—actors are not displaying but preparing for their display. However, this revised theory might be less parsimonious. It separates acting into two types of actions (actions in performance and actions in preparation) and uses two concepts (display and preparing for display) to interpret acting. Conversely, according to the theory I endorse, acting is a complete process rather than two distinct cases. Whether acting occurs in performance or preparation, it is at least a representation (Guo 2022). In this sense, my theory is more parsimonious.

Yet, the characters are fictional entities; unlike the boss of a company, the characters do not exist in the real world. Therefore Saltz further introduces "game intentions" to explain dramatic acting (Saltz 1991, 38-39). He claims that actors are more like game players rather than secretaries and inspectors. An example is chess. A chess player has no desire to actually fight with their opponent's king; they are aware that it is only a game and not the real king. Since they have a desire to play chess, they adopt the goals associated with the roles of the pieces and act on behalf of the pieces. Moreover, the pieces, such as the king, queen, or knight, are fictional and not real individuals. Therefore, Saltz suggests that in the context of playing a game, "to act on behalf of the game" should be interpreted as "to simply play according to the rules." The player's intentions derive from the rules of a game. Similarly, when actors portray characters, what they do are like what a game player does in a game. "Actors act on behalf of the characters" is interpreted as "to act according to the rules concerning the drama." The intentional states adopted by the actors derive from the rules governing the practice of dramatic acting. Briefly, according to Saltz, acting is viewed as a game; actors adopt the goals associated with the roles of the characters they portray, which is equivalent to the fact that actors act according to the relevant rules.

Nevertheless, the game model faces three serious difficulties. First, we can ask whether an actor should adopt the goal of the character (on behalf of the character) or merely act according to the theatrical rules. Which is more important to act the scene? The two interpretations conflict. Agents sometimes act according to the rules but do not adopt someone else's goal. Saltz imagines the following example (Saltz 1991, 44, Note 25). An employee believes making a phone call will result in making their boss unhappy but still does this because making requested phone calls is part of their job. Here, the employee acts according to the rules concerning their job but does not adopt the goal of the boss (their boss does not want to answer the phone). Similar examples exist in theatrical contexts. Suppose the actor portraying Romeo says to audiences, "Please look at me, Romeo, so pitiable and doomed!" Such an expression might produce excellent aesthetic effects and induce compassion in audiences. This conforms to the current theatrical conventions and rules (if this is what the script requires). However, at this

moment, the actor does not adopt Romeo's goal; Romeo does not say this because he believes no audience exists to observe him. The actor's acting is determined by the rules governing theatrical conventions and not by the character's goal.

Agents sometimes act on behalf of someone else but not according to the relevant rules. For example, Diego Maradona hit the ball into England's goal with his hand in the 1986 FIFA World Cup. This was named "The hand of God"—a famous event in the history of football. Maradona's action was illegal according to football rules. Yet, he acted on behalf of the Argentine football team, and what he did conformed to the interests of the Argentine team. In theatrical contexts, the actor's improvisation is a persuasive example. For example, in a performance of *Romeo and Juliet*, an actor portraying Romeo kisses the letters Juliet sent to him before his death. This could move audiences through Romeo's love. However, this behavior does not appear to conform to theatrical rules. It is inconsistent with the content of the play. Shakespeare's texts contain no such plot. Additionally, we could suppose that the script does not indicate that Romeo kisses the letters and that actors do not prepare for that event in rehearsal. Specifically, this behavior is improvised: the actor kisses Juliet's letters only because he has a sudden brilliant idea during the performance. Therefore, actors do not do this according to the rules governing the performance of *Romeo and Juliet*. However, the actor's behavior can satisfy Romeo's goal. It is reasonable to imagine that Romeo kisses Juliet's letters in the play to express his love.

Two other difficulties concern rehearsal and preparation. If the actor uses the "trial and error" method when preparing the roles, he intends to do something inappropriate for the character. For example, in *Hamlet*, when Hamlet drinks the poisoned wine and begs Horatio to live on and retell his story, he feels a sense of relief. The actor portraying Hamlet might struggle to show Hamlet's relief to audiences. Thus, when preparing the roles, he might attempt to act the scene with an entirely negative emotion excluding relief to understand why the latter is essential in this scene. Here the actor does not intend to act on Hamlet's behalf; Hamlet does not act with negative emotions excluding relief. Moreover, the actor does not act according to the rules governing the play. The play requires the actor to feel a sense of relief, not merely negative emotions. Unlike a game player, the actor in

rehearsal and preparation can intend to act the scene wrongly via the "trial and error" method, according to neither the theatrical rules nor the character's goal.

The final difficulty stems from the repetitiveness of dramatic acting. Actors often live and experience the same scene many times in rehearsal and performance. Although group rehearsal sometimes involves something imperfect, actors appear to portray the same characters, play the same roles, and act out the same scenes during a rehearsal and a performance. The actor portraying Romeo needs to see Juliet's death many times, whether in rehearsal or performance; the actor playing Hamlet also "kills" Claudius and "dies" in Horatio's arms countless times, whether in rehearsal or performance. The actors often experience and re-experience the same events in rehearsal and performance. However, players do not experience the same scene in most games. Consider basketball or chess. In rehearsal and preparation, the players do basic training, and they do not actually fight with another team or player. Two teams or two players may become familiar with each other via a warm-up match. However, warm-up matches still differ from real matches. They involve different results, players, formations, and procedures. No two matches are identical. Therefore, the events occurring in rehearsal (training) cannot be viewed as a replication of those occurring in the final performance. Based on this difference, a game is not an appropriate concept for understanding acting.

Certain states may appear appropriate for explaining the nature of acting; for example, a pretense, display, and game appear obvious candidates for understanding acting. However, the whole process of acting, including preparation, rehearsal, and performance, is not best characterized by reference to these distinctive states. As demonstrated, in rehearsal and preparation, actors' behaviors should not be viewed as pretense and display, and the game concept cannot explain certain characteristics of theatrical rehearsal, such as repetitiveness. Hence, these distinctive states are unsuitable candidates for understanding the nature of acting. Acting involves many states, some of which can be viewed as a pretense or display, with others resembling a game; however, the acting, since it involves different types of states, cannot be reduced to a certain distinctive state. Therefore, I conclude that the reductive understanding of acting is implausible. A *sui generis*

account of acting is required: acting is a *sui generis* human activity. Then, what characteristics does acting have? The next section answers this question.

4. Acting as a process

In Section 3, I argue that acting should not be viewed as a pretense, display, or game. Then, what is the nature of acting? I suggest using a process approach to consider the nature of acting: acting should be viewed as a distinctive process, not a distinctive state. That is, acting, as a *sui generis* human activity, is a long-term process comprising the input and output states, the operation and transitions between states, and the laws or rules. Therefore, the input, output and relevant rules or laws of acting should be characterized to understand what this distinctive process is. In this section, I explain the distinctive characteristics of the acting process.

First, I consider why acting is a process. Answering this question requires understanding what being a process is. Here, I do not want to provide a necessary and sufficient definition of a process but aim to understand what a process is by considering the distinction between a process and a state or event.

Peter Goldie (2012) provides a detailed account of the distinction between a state or event and a process. According to him, a process differs from a state or event based on three aspects. First, a process is something that persists by perduring, as "its identity is not determined at every moment of its existing," and a state or event is something that persists by enduring, as "its identity is determined at every moment at which it exists" (Goldie 2012, 61). For instance, red conscious experience is a mental state and not a process because it endures over time and is fully determined by what the subject feels during every moment it exists. Writing a cheque is a process because its temporal parts, such as depositing the drop of ink, are insufficient to determine which type of process it is. In this sense, cheque-writing, as a process, involves distinct states, and none of these states continue throughout the process. Second, a process, unlike an event, can be interrupted and restored. Yet, an event or state can only be stopped (Goldie 2012, 63). For instance, the process of writing a cheque can be interrupted

by a concern that the recipient is a cheat and is restored by dispelling this concern. Third, we appeal to a causal explanation to understand a state or event. For example, for red conscious experience, we look for the cause as whatever it was triggered, such as the fact of seeing a red apple. Yet, with a process, we seek "what sustains it, what keeps it on course, what prevents it from ceasing or disintegrating" (Goldie 2012, 63).

Acting appears to satisfy the three features. First, acting is not determined at any moment of its persistence. Every moment of acting is different; it is sometimes a preparation or rehearsal and sometimes an onstage performance. Actors have distinct intentions at different moments. In rehearsal, they intend to improve their onstage actions; in preparation, they intend to memorize the lines; in performance, they intend to attract audiences. Each single moment cannot determine it is acting. For instance, when preparing the roles, actors can repeat the lines aloud. This behavior cannot be determined as acting because other behaviors, such as foreigners reading English aloud, have similar characteristics. Second, acting can be interrupted. When tired, actors can walk off the stage and rest. At this moment, actors do not engage in dramatic acting. However, they can restore their acting after resting well. Third, the question of why actors give a fascinating performance does not imply a causal explanation. We do not explain that actors do the final performance only because they step out on the stage; many elements determine the actors' performances. We prefer to claim that not only the actor's long-term preparation and rehearsal but also their energy and emotional experience determine whether they can give an attractive performance. We can even consider the role of audiences: it is said that the audience's attention gives actors a vital force to act.

The thesis of acting as a process implies that mere pretense and display cannot determine what acting is. Pretense and display only occur during certain stages, such as onstage performance, not the whole process. Onstage performance can be viewed as a type of display; however, neither preparation nor rehearsal is display. Onstage performance involves pretense but also activities that are not pretense, such as basic actions—only raising one's arm, sitting down, or walking. That is, pretense and display are only episodes constituting the process of acting—they are not equivalent to the process of acting.

The next task is to characterize the acting process. Initially, consider the concept of process. Generally, a process involves four key parts:

1) The inputs entering the process from outside
2) The operations and transitions between different states
3) The laws or rules governing transitions and operations
4) The output derived but distinct from the inputs

If it is possible to successfully characterize the main features of the inputs, outputs, operations and transitions, and the norms or rules of a process, this process can be successfully distinguished from other types. Consider the example of a cheque-writing process. Its inputs are interpreted as one's intention to pay. The operations of cheque-writing are characterized as depositing each successive drop of ink. The rules governing the operations are understood as the social conventions concerning the circulation of money and psychological facts such as believing writing a cheque is a means of paying and desiring to pay money. Finally, the output is viewed as the amount appearing on the cheque. Cheque-writing is distinguished from other processes by the contents of the input and output states and the characteristics of the operations and rules.

The distinctive features of the acting process can be considered similarly. The inputs of the acting process are parts of what actors do in preparation, such as analyzing the texts and scripts, understanding the mental states of the characters, and learning about the characters' world. These include objective knowledge actors can acquire by learning. Additionally, when preparing the roles, actors also need to imagine what they would do if they were in the character's situations, recall their sensory and emotional memories consistent with the characters' situations, and imitate the characters' expressions, postures, and behaviors. Here, actors adjust their mental and behavioral states to live the characters. Therefore, they acquire subjective experiences. I claim that both the objective knowledge and subjective experience are important for preparing for the roles. Yet, actors do not need to do all the things I listed; due to different social and historical contexts, characters, and acting techniques, actors make different preparations. I adopt a *disjunctive* explanation to specify the determinable features of the inputs of the acting process: actors acquire objective knowledge about the

character and play, *or* acquire subjective and personal experience, *or* acquire both.

The rules governing the transition from inputs to outputs are interpreted based on two aspects. The first is the aspect of the play. In the acting process, actors' actions should conform to the relevant conventions of the play. Suppose an actor is asked to portray Hamlet in Shakespeare's classic tragedy *Hamlet*. His behaviors on stage must be consistent with the relevant conventions of the classic tragedy. If actors broke the conventions and portrayed Hamlet as a licentious and brutal guy, they would fail to act out Hamlet.[8] The second is the aspect of the audiences. Actors should be bound by certain communicative intentions. I suggest that this intention is not to communicate the content of the play to audiences because mere storytelling also involves such intentions. The content of this intention is instead an imaginative identification; actors intend their audiences to imagine or make-believe they are identical to the characters (Guo 2022). For instance, when portraying Romeo, the actor intends his audiences to imagine him as Romeo. If the audiences fail to imagine this when watching his performance, his acting is unsuccessful.

Additionally, I do not think the imaginative identification intention is explicit in rehearsal and preparation. In the latter cases, actors might not consciously have such intentions; instead, they intend to improve their performances. Yet, here, the imaginative identification intention still affects the actor's acting. Instead, this intention is converted to a normative standard, which generates other intentions, such as memorizing the lines, improving the performance, and overcoming difficulties. Their aim is to make audiences better imagine the actor as the character.

[8] An anonymous reviewer suggests this explanation is circular: theater acting is a process involving theatrical conventions. I do not agree. These conventions can be understood in a way that does not concern the concept of theater. They can be viewed as texts involving instructions for actions. Alternatively, they can be viewed as suggestions teaching how people act, such as expressing emotions or adopting body postures. Moreover, given that acting can occur in non-theater contexts, the conventions can also be understood as rules of psychotherapy, education, or filmmaking. I cannot list all the conventions related to acting; therefore, I simply call them "the conventions of the play."

I still adopt a disjunctive explanation to specify the rules governing the acting process: actors act *either* according to the relevant conventions, *or* consistently with their imaginative identification intentions, or with both.[9]

The operation of the process of acting can be characterized as a representation; simply, actors represent someone else's features. When preparing the roles alone, actors do not intend to pretend or display to audiences; however, they are still representing the character's features. This type of representational relationship involves variants. First, actors use their own features to represent features they do not have but another individual—often a fictional character—would have in the fictional world. Second, actors attribute their own features to a different individual; namely, the actors and their characters have the same features. For example, when Hamlet raises his arm, the actor portraying Hamlet also raises his arm. Here the actor and his character Hamlet have the same feature; he attributes his own feature to Hamlet—a fictional character. Third, the actor and the character they portray are sometimes the same individual: actors play themselves on stage. I suggest that actors use their own features in this case to represent the features that they do not actually have but would have at a different moment. For example, Donald Trump plays himself in the TV series *Days of Our Lives*, and he refuses to trade sex for a job in the fiction. Here Trump does not actually refuse sex trafficking; instead, he represents himself refusing sex trafficking in a counterfactual situation. Briefly, when portraying

[9] Notably, an intention to act according to the convention of the play can be derived from an intention to imagine an identification. When an actor intends to make their audiences imagine they are the character, the actor can consciously consider whether their acting conforms to the relevant conventions. Otherwise, the audiences would fail to imagine the actor as the character. For example, if the actor broke the conventions and portrayed Hamlet as a licentious and brutal guy, the audiences would resist imagining him as Hamlet. However, actors do not have an imaginative identification intention in all cases. For instance, Brechtian theaters require a particular type of acting in which actors fail to portray their characters in performance. I suggest that Brechtian actors act only according to the relevant conventions of the Brechtian theater, not with their imaginative identification intentions. Thus, I suggest a disjunctive explanation. I do not require actors must have an imaginative identification intention. This point differs from Guo (2022), who claims that imaginative identification is indispensable.

the characters, actors represent their own features as those other individuals would have or the features the actors would have at a different time.

Moreover, this representation process is governed by theatrical conventions or the actor's imaginative identification intentions. When actors aim to represent characters, their representations should conform to theatrical conventions or be consistent with their intentions to make audiences imagine an identification. If Hamlet is described as indecisive in the play, the actor should aim to represent an indecisive person, or he should act to make audiences more easily imagine he is indecisive.

The final onstage performance, which the audiences watch, can be viewed as the outputs of the process of acting. By rehearsing and preparing the roles, actors build patterns, models, and routines. Via the latter, they successfully portray the characters in the final performance.

Briefly, acting is a process from preparation to performance. Initially, actors aim to understand the content of the play or trigger their own experiences; then, via their understanding and experiences, actors prepare and rehearse the scenes and lines. Finally, via the routines built in rehearsal, actors portray their characters to their audiences in onstage performances. This process is triggered when the actors start to make preparations to enter the roles and ends when the onstage performance is over and the actors walk off the stage. This process is characterized as a representation—actors represent the characters' features and are governed by certain norms, such as the relevant theatrical conventions and the actors' imaginative identification intentions. Acting is distinguished from other processes by the features of its inputs, outputs, and operations governed by the relevant rules. One of those elements cannot alone distinguish acting.[10]

[10] Certain states may be especially appropriate for describing certain types of acting process. For example, the character's emotion felt by actors evidently defines method acting. However, acting is not best characterized by reference to only a certain single state because this cannot distinguish acting from other types of processes. For instance, for Brechtian theaters, actors are not encouraged to feel the character's emotions; in this sense, the character's emotion—a distinctive state—is unnecessary for the process of acting. Alternatively, in people's engagement with fiction, they can be immersed in the narrative and feel the character's emotions; here, the character's emotion is insufficient for distinguishing acting from reading. Hence, I suggest acting

Consider a child's pretend play and a liar. In pretend play, unlike actors, children do not need to understand the play and experience the characters when preparing the play; in telling a lie, unlike actors, the liar does not act with the intention to make audiences imagine an identification. Reconsider the ordinary actions. They are not representations. Unlike actors, when someone does something in daily life, they do not represent the other's features. In this sense, the process-approach differentiates acting.

The process-approach can also differentiate different types of acting. Consider two actors, one portraying Hamlet and another portraying Romeo. Both are still distinguished by the features of their inputs, outputs, and operations governed by the relevant rules. When portraying Hamlet, the actor analyzes the text of *Hamlet*, represents Hamlet's features, and intends his audiences to imagine that he is Hamlet. Moreover, when portraying Romeo, the other actor analyzes the play *Romeo and Juliet*, represents Romeo, and intends the audiences to imagine he is Romeo and not other characters.

Reconsider the role-playing therapy. If a patient has phobias, a therapist can play the role of someone causing these phobias to help the patient overcome them. Compare this case with dramatic acting. Different rules are involved. One involves the rules of psychotherapy, such as ensuring patients' safety; another involves the rules of theater, such as attracting audiences.

5. Conclusion

There are four different kinds of ways to grasp a thing. Reductive understanding states something is placed in a well-known category and uses the features of the things in this category to understand it. *Sui generis* understanding states something is not placed in the well-known category but classified into a *sui generis* category. I argue that the three current popular theories of the nature of acting (the pretense theory, the display account, and the game model) are implausible and cannot explain some

is differentiated by the whole process comprising inputs, outputs, the operations between states, and the relevant norms.

cases of acting. Therefore, I suggest that other concepts should not be used to understand the nature of acting; a *sui generis* account of acting rather than a reductive account is required. Acting is to portray roles and act out scenes, not other kinds of activities.

I also introduce the distinction between a state approach and a process approach. I further argue that acting should be viewed not as a state but a process because it has some characteristics of a process but not a state. Therefore, acting can be distinguished from other processes by specifying the characteristics of its inputs, outputs, operations, and rules.

Notably, I do not provide a sufficient and necessary definition of acting (I do not even assert that it is a "definition") because I do not explain the conventions, the performance, or what preparation involves in detail. This might be future work. This article only aims to provide a new way to understand acting and argue that this new way helps distinguish acting from other concepts.

References

Alward, Peter. 2009. "OnStage Illocution." *Journal of Aesthetics and Art Criticism*, 67(3): 321–31. https://doi.org/10.1111/j.1540-6245.2009.01362.x

Austin, John Langshaw. 1979. *Philosophical Papers*. New York: Oxford University Press. https://doi.org/10.1093/019283021X.001.0001

Bennett, Michael Y. 2017. *Analytic Philosophy and the World of the Play*. London: Routledge. https://doi.org/10.4324/9781315294735

Davies, David. 2011. *Philosophy of the Performing Arts*. Hoboken, New Jersey: Wiley-Blackwell. https://doi.org/10.1002/9781444343458

Goldie, Peter. 2012. *The Mess Inside: Narrative, Emotion, and the Mind*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230730.001.0001

Guo, Yuchen. 2022. "What is Acting." *The Journal of Aesthetics and Art Criticism*, 80(1): 58–69. https://doi.org/10.1093/jaac/kpab066

Hamilton, James R. 2007. *The Art of Theater*. Hoboken, New Jersey: Wiley-Blackwell. https://doi.org/10.1002/9780470690871

Hamilton, James R. 2009. "Pretense and Display Theories of Theatrical Performance." *Organon F*, 16(4): 632–54.

Hamilton, James R. 2013. "Acting." *Journal of Dramatic Theory and Criticism*, 28(1): 35–39. https://dx.doi.org/10.1353/dtc.2013.0020

James, William. 1884. "What is an Emotion?" *Mind,* 9(2): 188–205. https://doi.org/10.1037/11304-033

Lewis, David. 1978. "Truth in Fiction." *American Philosophical Quarterly*, 15(1): 37–46.

Nussbaum, Martha C. 2001. *Upheavals of Thought: The Intelligence of Emotions.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511840715

Osipovich, David. 2006. "What is a Theatrical Performance?" *The Journal of Aesthetics and Art Criticism*, 64(4): 461–70. https://doi.org/10.1111/j.1540-594X.2006.00224.x

Roberts, Robert C. 2003. *Emotions: An Essay in Aid of Moral Psychology.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610202

Saltz, David Z. 1991. "How to Do Things On Stage." *Journal of Aesthetics and Art Criticism*, 49(1): 31–45. https://doi.org/10.2307/431647

Saltz, David Z., and David Krasner, eds. 2006. *Staging Philosophy: Intersections of Theater, Performance, and Philosophy.* Ann Arbor, Michigan: University of Michigan Press. https://doi.org/10.3998/mpub.147168

Searle, John. 1975. "The Logical Status of Fictional Discourse." *New Literary History*, 6(2): 319–32. https://doi.org/10.2307/468422

Stern, Tom. 2014. *Philosophy and Theatre: An Introduction.* London and New York: Routledge. https://doi.org/10.4324/9781315887517

Stern, Tom, ed. 2017. *The Philosophy of Theatre, Drama and Acting.* Lanham, Maryland: Rowman & Littlefield International.

Walton, Kendall. 1990. *Mimesis as Make-Believe.* Cambridge, Massachusetts: Harvard University Press. https://doi.org/10.2307/1772866

Woodruff, Paul. 2008. *The Necessity of Theater: The Art of Watching and Being Watched.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195332001.001.0001

Organon F 31 (4) 2024: 432–433
https://doi.org/10.31577/orgf.2024.31406
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

REPORT

Pavel Materna
(1930 – 2024)

On October 25, prof. Pavel Materna, aged 94, passed away peacefully in Prague. This notice, though awaited, afflicted us a lot. Pavel was a Czech logician, philosopher and musician. He was a follower of Pavel Tichý, the founder of Transparent Intensional Logic (TIL). Pavel Materna contributed a lot to the development of TIL, the system within which he developed an original theory of concepts. In this theory, concepts are defined as abstract procedures that can be viewed as 'instructions' encoded by natural-language expressions on obtaining an object (if any) denoted by a given expression. These results yielded numerous papers and three books. They are "Svět pojmů a logika" (in Czech), Praha, Filosofia 1995, "Concepts and Objects", Helsinki, Acta Philosophica Fennica vol. 63, 1998, and "Conceptual Systems", Berlin, Logos Verlag 2004.

Applying TIL to natural-language processing was Pavel's primary professional interest. He was known internationally for his influential research in this area, where he achieved significant results and published numerous papers and books. This work culminated in the book published with co-authors Marie Duží and Bjørn Jespersen, "*Procedural Semantics for Hyperintensional Logic*" with the subtitle "*Foundations and Applications of Transparent Intensional Logic*", Springer 2010. The book provides a much-needed comprehensible, elaborate and systematic exposition of the foundations of Transparent Intensional Logic, demonstrates how TIL lends itself to a broad range of applications and presents new results concerning knowledge representation, attitude logic, incomplete meanings (anaphoric references), and philosophy of mathematics. In July 2011, the Council of the Academy of Sciences of the Czech Republic presented Procedural Semantics for Hyperintensional Logic with its prestigious award for outstanding results of major scientific importance.

Personally, Pavel was a good friend, joyful in discussions and a great musician and pianist. We had a lot of fun listening to his piano playing on conference evenings. He had an unbelievable musical memory and could play any motive

from Wagner's operas, as well as works of other great composers like Leoš Janáček ("On an Overgrown Path") and Bohuslav Martinů ("Puppets"). In addition, anybody could ask for any anthem, even Chinese or Japanese, and Pavel played it for the great joy of the listeners.

Besides logic, the author of this short obituary shared another common hobby with Pavel: classical music. We played four-hand piano compositions and performed them at the conference SOFSEM, which used to be an extraordinary phenomenon in then Czechoslovakia. The conference lasted a fortnight and was a winter school of informatics and other close areas, such as philosophy and logic. In those dark times of "normalization" after the Soviet invasion of Czechoslovakia in 1968, SOFSEM was an island of freedom with an extremely nice and friendly atmosphere.

Pavel loved life and women. He was married four times, and he had six children with three of his wives. He spent the longest time, more than fifty years, with his last wife Helena and their children Klára and Štěpán. Helena is a medical doctor who devoted her care to Pavel till his last days. Pavel Materna died fifty years after the death of his close friend Pavel Tichý.

He will be greatly missed by all who love him.

Marie Duží

Call for Papers

2026 marks the 50th anniversary of the publication of David Lewis famous paper "The Paradoxes of Time Travel" in the APQ. This paper has dominated the philosophy of time travel since its publication. So much so that a philosopher's stance on time travel largely consists in the extent of his agreement or disagreement with Lewis.

Organon F would like to mark the occasion by devoting a special issue to Lewis's paper. This issue will appear in February 2026. We thus call for papers of between 2000 – 8000 words which directly engage with one or more of the claims that Lewis makes in his paper.

All papers should be submitted by **May 2025**.

Guest Editors

Brian Garrett (The Australian National University)

Jeremiah Joven Joaquin (De La Salle University)

Call for Papers

ISSUES ON THE (IM)POSSIBLE XI

May 27-29, 2025 (Bratislava, SLOVAKIA)

Keynote Speakers

Andrea Iacona (Università di Torino)

Edward N. Zalta (Stanford University)

We invite submissions for a 30-minute presentation, discussion excluded. The conference focuses on analytic philosophy. The areas of interest include, but are not limited to modal metaphysics, the metaphysics of fiction, hyperintensionality, modal epistemology, and metaontology.

An abstract of approximately 500 words should be prepared for blind review and include a cover page with the full name, institution, and contact information. Abstracts can be submitted in pdf or doc(x) and should be sent to impossible@savba.sk

Informal queries: martin.vacek@savba.sk

Deadline for submission: **January 31, 2025**
Notification of acceptance: **February 28, 2025**

Venue: 19 Klemensova, Bratislava, Slovakia
Webpage: www.impossible.sk

Contents

RESEARCH ARTICLES