

FELSEFE DÜNYASI

The Philosophy World

Sayı/Issue: 83 (Yaz/Summer 2026)

ISSN 1301-0875 - eISSN 2822-2970

Yayıncı ve Sahibi/Publisher and Owner

Türk Felsefe Derneği Adına/On behalf of Association for Turkish Philosophy

 Prof. Dr. Murtaza Korlaelçi (Türk Felsefe Derneği, Ankara, Türkiye)

Yazı İşleri Müdürü/Managing Editor

 Dr. Merve Nur Sezer (Ankara Üniversitesi, Ankara, Türkiye)

Editör/Editor

 Prof. Dr. Hasan Yücel Başdemir (Ankara Üniversitesi, Ankara, Türkiye)

Yazı Kurulu/Editorial Board

 Prof. Dr. Celal Tüzer
(Ankara Üniversitesi, Ankara, Türkiye)

 Prof. Dr. Christof Rapp
(Ludwig Maximilians Universität, München, Germany)

 Prof. Dr. Dean Zimmerman
(Rutgers University, New Jersey USA)

 Prof. Dr. Duncan Pritchard
(University of California Irvine, Irvine USA)

 Prof. Dr. Jason Brennan
(Georgetown University, Washington DC, USA)

 Prof. Dr. Muhammet Enes Kala
(Ankara Yıldırım Beyazıt Üniversitesi, Ankara, Türkiye)

 Prof. Dr. Fulya Bayraktar
(Ankara Hacı Bayram Veli Üniversitesi, Ankara, Türkiye)

 Assoc. Prof. Dr. Eros Carvalho
(Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil)

 Assoc. Prof. Dr. Zalan Gyenis
(Jagiellonian University, Kraków, Poland)

 Doç. Dr. Fatih Özkan
(Ankara Hacı Bayram Veli Üniversitesi, Ankara, Türkiye)

 Dr. Gary N. Kemp
(University of Glasgow, Glasgow, United Kingdom)

Alan Editörleri/Section Editors

 Prof. Dr. Ahmet Emre Dağtaşoğlu
(Trakya Üniversitesi, Edirne, Türkiye)

 Prof. Dr. Sebile Başok Diş
(Necmettin Erbakan Üniversitesi, Konya, Türkiye)

 Doç. Dr. Mehtap Doğan
(Ankara Yıldırım Beyazıt Üniversitesi, Ankara, Türkiye)

 Doç. Dr. Nihat Durmaz
(Selçuk Üniversitesi, Konya, Türkiye)

 Doç. Dr. Yurdağül Kılınç
(Selçuk Üniversitesi, Konya, Türkiye)

 Doç. Dr. Üyesi Nazan Yeşilkaya
(Şırnak Üniversitesi, Şırnak, Türkiye)

 Dr. Öğr. Üyesi Ferhat Taşkın
(Boğaziçi Üniversitesi, İstanbul, Türkiye)

 Dr. Öğr. Üyesi Kenan Tekin
(Boğaziçi Üniversitesi, İstanbul, Türkiye)

 Dr. Öğr. Üyesi Muhammet Çelik
(Ankara Sosyal Bilimler Üniversitesi, Ankara, Türkiye)

 Dr. Öğr. Üyesi Övge Öztürk
(Yozgat Bozok Üniversitesi, Yozgat, Türkiye)

Yazım Editörleri/Spelling Editors

 Ahmet Hamdi İşcan
(Ankara Üniversitesi, Ankara, Türkiye)

 Zehra Eroğlu
(Ankara Üniversitesi, Ankara, Türkiye)

Dil Editörleri/ Language Editors

 Dr. Hatice İpek Keskin
(Ankara Sosyal Bilimler Üniversitesi, Ankara, Türkiye)

 Abdussamet Şimşek
(Ankara Sosyal Bilimler Üniversitesi, Ankara, Türkiye)

Dizinlenme Bilgileri/Indexed by

ULAKBİM TR Dizin (2003'dan itibaren/since 2003)

Philosopher's Index (2004'dan itibaren/since 2004)

ERIH PLUS (2023'den itibaren/since 2023)

PhilPapers (2023'dan itibaren/since 2023)

DOAJ: Directory of Open Access Journals (2024'ten beri/since 2024)

EBSCO Database (2025)

Ulrich's Periodicals Directory (2025)

Felsefe Dünyası uluslararası, süreli ve hakemli bir dergidir; yılda iki sayı yayımlanır; elektronik versiyonu, açık erişimlidir.

Derginin online versiyonu ücretsizdir.

The Philosophy World is an international, periodical biannually, and peer-reviewed journal; its digital version is open access.

The online version of the journal is free of charge.

Fiyatı/Price: 500 TL | Basım Tarihi/Publication Date: Temmuz/July 2026, 200 Adet/Copies

Adres/Address

Necatibey Caddesi No: 8/122 Çankaya/ANKARA

Tel: 0 (312) 231 54 40

<https://dergipark.org.tr/tr/pub/felsefedunyasi>

Hesap No / Account No: Vakıf Bank Kızılay Şubesi

IBAN: TR82 0001 5001 5800 7288 3364 51

Tasarım / Design: Turku Ajans

Baskı / Printed: Uzun Dijital

İstanbul Çarşısı, Zübeyde Hanım, İstanbul Cd. No:48

D:48, 06070 Altındağ/Ankara

Tel: (0312) 341 36 67 | Sertifika No: 47865

A Critique of Practical Identity: Korsgaard's Conditional Motive*

Aslıcan TÜTER

<https://orcid.org/0009-0004-7690-138X> | aslican.tuter@std.bogazici.edu.tr

Boğaziçi University, Institute of Social Sciences, İstanbul, Türkiye

<https://ror.org/03z9tma90>

Abstract

This paper provides a critical analysis of Christine Korsgaard's constructivist attempt to ground moral obligation in the reflective structure of "practical identity." Korsgaard posits that the normative force of moral laws stems from the agent's internal necessity to maintain a coherent self-conception. The central argument of this paper is that in Korsgaard's framework, moral motivation devolves into a form of "conditional self-interest," as the authority of the moral law becomes contingent upon the agent's psychological desire for self-preservation. This paper examines this structural weakness through the "shmagency" challenge (Enoch) and the case of the "idealized Mafioso" (Cohen), including a reconstruction of Korsgaard's own responses to these objections, in order to show that her replies do not dissolve the fundamental problem. Crucially, the paper frames this critique not as a failure to satisfy Kantian standards, for Korsgaard consciously departs from Kant, but as an internal failure: her constructivist framework cannot secure the unconditional normative force it sets out to establish, even by its own lights. Furthermore, when compared against rival accounts of normativity, including moral realism and the Fact of Reason, the constructivist grounding proves less adequate. The paper concludes by defending Dieter Schönecker's intuitionist reading of the "Fact of Reason" (*Factum der Vernunft*), arguing that moral authority must be recognized as an objective, non-deductive *a priori* structural necessity rather than a subjective product of reflective endorsement.

Keywords: Kant, Korsgaard, Practical Identity, Autonomy, Categorical Imperative, Fact of Reason

Citation

Tüter, Aslıcan (2026). A Critique of Practical Identity: Korsgaard's Conditional Motive. *Felsefe Dünyası*, 83, 180-199. <https://doi.org/10.58634/felsefedunyasi.1874332>

Date of Submission	29.01.2026
Date of Acceptance	16.06.2026
Date of Publication	15.07.2026
Peer-Review	Double anonymized - Two External
Ethical Statement	It is declared that scientific and ethical principles have been followed while carrying out and writing this study and that all the sources used have been properly cited.
Plagiarism Checks	Yes – intihal.net
Conflicts of Interest	The author(s) has no conflict of interest to declare.
Complaints	felsefedunyasi@tufed.com.tr
Grant Support	The author(s) acknowledge that they received no external funding in support of this research.
Copyright & License	Authors publishing with the journal retain the copyright to their work licensed under the CC BY-NC 4.0 .

* I would like to express my gratitude to Dr. Saniye Vatansever and Dr. Ferhat Taşkın for their contributions to this article.

Pratik Kimliğin Eleştirisi: Korsgaard'ın Koşullu Motifi*

Aslıcan TÜTER

<https://orcid.org/0009-0004-7690-138X> | aslican.tuter@std.bogazici.edu.tr

Boğaziçi Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul, Türkiye

<https://ror.org/03z9tma90>

Öz

Bu makale, Christine Korsgaard'ın Kantçı otonomiye yönelik inşacı (constructivist) yorumunu, ahlaki zorunluluğu pratik kimlik üzerinden temellendirme girişimi bağlamında eleştirel bir analize tabi tutmaktadır. Makalenin temel tezi, Korsgaard'ın sisteminde ahlaki motivasyonun failin kimlik bütünlüğünü koruma arzusuna bağlı hale gelerek "koşullu bir kişisel çıkara" (conditional self-interest) dönüştüğüdür. Bu eleştiri, Korsgaard'ın Kantçı standartları karşılayıp karşılamadığı sorusundan değil, zira Korsgaard Kant'tan biliçli olarak uzaklaşmaktadır, kendi kurduğu konstrüktivist çerçevenin kendi standartlarını karşılayıp karşılamadığı sorusundan hareket etmektedir. "Shmagency" (failimsi) itirazı ve "idealize edilmiş Mafioso" örneği üzerinden Korsgaard'ın bu itirazlara verdiği yanıtlar da dahil edilerek çerçevedeki olumsuzluk ortaya konmaktadır. Sonuç olarak bu çalışma, Dieter Schönecker'in sezgici okumasını esas almaktadır. Bu okuma çerçevesinde ahlaki otorite, failin yansıtıcı onayına bağlı bir kazanım olarak değil, failden bağımsız olarak kendisini dayatan bir "Akıl Olgusu" (Fact of Reason) olarak kavranmalıdır. Buna göre ahlak yasası, rasyonel doğamızın sarsılmaz bir *a priori* yapısal zorunluluğunu oluşturmaktadır.

Anahtar Kelimeler: Kant, Korsgaard, Pratik Kimlik, Otonomi, Kategorik Buyruk, Akıl Olgusu

Atıf Bilgisi Tüter, Aslıcan (2026). A Critique of Practical Identity: Korsgaard's Conditional Motive. <i>Felsefe Dünyası</i> , 83, 180-199. https://doi.org/10.58634/felsefedunyasi.1874332	
Geliş Tarihi	29.01.2026
Kabul Tarihi	16.06.2026
Yayın Tarihi	15.07.2026
Değerlendirme	İki Dış Hakem / Çift Taraflı Körleme
Etik Beyan	Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.
Benzerlik Taraması	Yapıldı – intihal.net
Çıkar Çatışması	Çıkar çatışması beyan edilmemiştir.
Etik Bildirim	felsefedunyasi@tufed.com.tr
Finansman	Bu araştırmayı desteklemek için dış fon kullanılmamıştır.
Telif Hakkı & Lisans	Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmalarını CC-BY-NC 4.0 lisansı altında yayımlanmaktadır.

* Bu makaleye katkılarından dolayı Dr. Saniye Vatansever ve Dr. Ferhat Taşkın'a teşekkürlerimi sunarım.

1. Introduction

In the *Groundwork for the Metaphysics of Morals* (henceforth, the *Groundwork*), Immanuel Kant establishes the uncompromising demand that the moral law must possess unconditional supremacy: its principles must be subsistent *a priori* for themselves and must be valid for all rational beings in general, not merely under contingent conditions or based on the particular quality of human nature. If the moral law fails to meet this demand for unconditioned necessity, the entire system crumbles and falls into heteronomy, where the will is determined by empirical factors external to the rational will itself - such as desires, emotions, inclinations, authority, or divine command - rather than by itself. Within the *Groundwork*, Kant thus tries to establish how the fundamental moral command, the Categorical Imperative (henceforth, CI) is possible. As this imperative represents an action as objectively necessary for itself, without any reference to another end, Kant explicitly states that gaining insight into this would be an arduous task:

Secondly, the reason why it is difficult to gain insight into the possibility of this categorical imperative or law of morality is a very serious one. It is an *a priori* synthetic practical proposition; and as it is so difficult to make sense of the possibility of such propositions when it comes to theoretical knowledge, it is plain that it will be no less difficult when it comes to practical knowledge (Kant, *GMS*, 4:420–421).¹

The difficulty lies in demonstrating how the CI, a principle not analytically contained in the concept of a free will, can generate a pure, universal interest in the agent without relying on contingent desires or inclinations, thus avoiding heteronomy. This conceptual gap creates the tension in modern Kantian ethics, epitomized by what Pauline Kleingeld and Marcus Willaschek call the Kantian paradox: “the more one emphasizes the impersonal, *a priori*, timeless character of autonomy, the harder it is to make literal sense of self-legislation as an *act* or *activity* of the will” (Kleingeld & Willaschek, 2019, p. 3).

This unresolved problem of validation is precisely the crisis into which Christine Korsgaard intervenes. It is crucial to distinguish between Kant's primary concern, which is how the CI is objectively possible as a synthetic *a*

1 All references to Kant's works in this paper refer to the volume and page numbers of the standard German edition (Akademie-Ausgabe, Berlin: Walter de Gruyter, 1900–). The English translations used are as follows: *Groundwork for the Metaphysics of Morals* (*GMS*), translated and edited by C. Bennett, J. Saunders, and R. Stern (Oxford University Press, 2019); and *Critique of Practical Reason* (*KpV*), translated and edited by M. Gregor (Cambridge University Press, 1996).

priori law and Korsgaard's shifting of the focus toward the problem of moral motivation: the urgent practical "why" we are bound to morality. In her lecture "The Authority of Reflection" within *The Sources of Normativity*, Korsgaard addresses this motivational normative question. She seeks to vindicate the supreme principle of morality by grounding its necessity not in abstract metaphysical principles, but in the inescapable requirements of human practical identity. Korsgaard argues that because human consciousness has a reflective structure, we cannot simply act on unendorsed impulses. Reflection "forces to have a conception of ourselves" (Korsgaard, 1996, p. 100), compelling the agent to identify with "some law or principle which will govern your choices" (Korsgaard, 1996, p. 104). By rooting the law in the nature of "the authority of your own mind and will," Korsgaard aims to provide an account of why the CI is necessary for beings like us.

A critical clarification is necessary at the outset, and it shapes the entire argumentative structure of this paper. Korsgaard's *The Sources of Normativity* is not an exercise in Kant exegesis; it is an original metaethical project that consciously departs from Kant in important respects. While Korsgaard shares with Kant the conviction that normativity must be located in the autonomous agent's relation to her own will, she replaces Kant's appeal to pure reason as such with an appeal to the reflective structure of practical identity, a move Kant himself would have regarded as a regression to contingency. It would therefore be a mischaracterization to criticize Korsgaard for failing to satisfy standards she never set out to meet.

The critique in this paper operates on a different level. It asks whether Korsgaard's constructivist framework succeeds on its own terms. The central argument is that it does not. By establishing the moral motive in the agent's own self-maintenance, the theory introduces a fatal contingency. It transforms pure duty into conditional self-interest, thereby reducing the moral law's authority and creating an inherent structural risk of moral failure whenever the preservation of identity conflicts with what morality demands. Moreover, when placed alongside rival accounts of normativity, for example moral realism and the Fact of Reason reading, which will be introduced progressively throughout the paper and defended in full in Section 5, Korsgaard's constructivism proves less adequate as a grounding of unconditional moral authority. The original contribution of this paper lies in its analysis of Korsgaard's own responses to her critics. While scholars such as Cohen, Enoch, and FitzPatrick have identified the contingency problem in Korsgaard's account, the specific question of whether Korsgaard's own replies to these objections succeed has received

comparatively little systematic attention. This paper argues that both of Korsgaard's principal responses, that is the appeal to humanity as the ground of all practical identities, and the appeal to a hierarchy of identities culminating in a universal human identity, import substantive Kantian premises that her constructivist framework cannot itself derive without circularity. The problem is thus not merely that contingency infects the system from outside, but that Korsgaard's own attempts to close that gap reproduce the very problem they are meant to solve.

This paper is structured as follows. Section 2 outlines Korsgaard's core constructivist thesis. Section 3 presents the central internal critique, arguing that her framework reduces the CI's motive to a conditional interest and examining her own responses to major objections, particularly the Mafioso case. Section 4 examines the realist counter-proposals of Kleingeld and Willaschek, showing how the dispute over self-legislation reveals an impasse that itself points toward the need for an alternative grounding. Section 5 develops this alternative by reconstructing the Fact of Reason introduced in the preceding sections as the horizon of the problem and defending Schönecker's intuitionist reading as the most adequate resolution. Section 6 concludes.

2. Korsgaard's Account: Autonomy as Reflective Self-Constitution

"Why should I act according to the moral law?" is the central question that Korsgaard and many other scholars, including Andrews Reath (2006) and Sharon Street (2012), attempt to illuminate. The problem has the structure of an antinomy with two conflicting logical requirements. First, the problem of antecedent reasons: if we impose a principle on ourselves, we must have a reason to do so; but if that reason exists prior to our act of willing, then it would not be self-imposed, rather a pre-existing normative fact we simply recognize. Second, the problem of sovereign release: if the law's authority comes entirely from the agent's act of will, the agent possesses the power to unmake the law at any time, making moral obligation arbitrary rather than unconditional. As Robert Stern summarizes: "Korsgaard's Kantian solution to this antinomy is to argue that these principles or laws must be self-imposed" (Stern, 2015, p. 58).

Reath's account of autonomy emphasizes the legislative aspect of the moral agent's rational authority, in which autonomy is not merely a psychological capacity but the agency's capacity to legislate through its rational will, functioning as a sovereign legislator (Reath, 2006, p. 4). Sharon Street provides a metaethical foundation for the constructivist school,

arguing that the authority of moral reasons is internal to the structure of practical reasoning: “The truth of a normative claim consists in that claim’s following, as a logical or instrumental matter, from within the practical point of view, where the practical point of view is given a formal characterization” (Street, 2012, p. 41).

In *The Sources of Normativity*, Korsgaard argues that any account of moral authority relying on external facts faces an explanatory dead-end. She contends that the authority of a reason cannot be found in the object of desire itself, but must be located in the act of reflection:

“Reason” means reflective success. So if I decide that my desire is a reason to act, I must decide that on reflection I endorse that desire. And here we run into the problem. For how do I decide that? Is the claim that I look at the desire, and see that it is intrinsically normative, or that its object is? Then all of the arguments against realism await us... What brings such a course of reflection to a successful end? (Korsgaard, 1996, p. 97).

Korsgaard argues that realism fails because it attempts to halt the regress of justification by declaring that some things are intrinsically normative, yet ultimately relies on the realist’s confidence in his own moral beliefs. When an agent asks “Why must I do so and so?”, the realist asserts that obligation is “simply there, part of the nature of things” (Korsgaard, 1996, p. 30), an answer that merely reasserts rather than grounds the claim. Similarly, she rejects voluntarism (the view that obligation derives purely from the command of a legislator, divine or otherwise), because the authority of the commander itself requires prior moral justification, pushing the source of normativity endlessly backward.

Since authority cannot be found outside the agent, it must be found inside the act of the agent’s practical deliberation. Korsgaard explains that human consciousness has a reflective structure that grants us the capacity to turn our attention onto our own mental activities. We can step back from our desires and impulses. For Kant, while we can act on permissible inclinations, we must subordinate them to the moral law by endorsing them as maxims. Korsgaard adopts this by stating we cannot simply act on a desire without reflectively endorsing it as a reason for action. The search for the unconditioned is successfully terminated when the agent realizes that the source of authority resides in their own self-legislating will.

The conceptions of ourselves that are most important to us give rise to unconditional obligations, as violating them means losing integrity and

identity: "An obligation always takes the form of a reaction against a threat of a loss of identity" (Korsgaard, 1996, p. 102). The concept of "practical identity" refers to the conception of oneself under which one values one's life and actions, here we can refer to the sphere of roles, commitments, and self-conceptions that provide reasons and obligations. As she puts it:

An agent might think of herself as a Citizen of the Kingdom of Ends. Or she might think of herself as someone's friend or lover, or as a member of a family or an ethnic group or a nation. She might think of herself as the steward of her own interests, and then she will be an egoist. Or she might think of herself as the slave of her passions, and then she will be a wanton. And how she thinks of herself will determine whether it is the law of the Kingdom of Ends, or the law of some smaller group, or the law of egoism, or the law of the wanton that will be the law that she is to herself (Korsgaard, 1996, p. 101).

Korsgaard's view determines practical identity as complex and multifaceted. It is this conception of how one values and interprets oneself that underpins the normativity of moral action and obligation. Crucially, moral identity does not erase other forms of identity: "no one is simply a moral agent and nothing more" (Korsgaard, 1996, p. 125). Rather, the reflective and universal structure of moral identity gives it a unique authority that provides the framework within which more local identities operate. Tamar Schapiro (1999) supports this account by framing agency itself as a constitutive task: the individual must adjudicate between conflicting motivational claims to determine which conform to "the law of her will" (Schapiro, 1999, p. 729), substituting reason for instinct. In this way, moral agency involves actively valuing one's identity as a rational and moral being, an endorsement that shapes one's motivations and justifies moral actions.

This distinction between descriptive self-conception and normative commitment clarifies how practical identity generates motivational force. For Korsgaard, an identity is never a passive psychological fact; it is a normative framework that an agent must actively endorse to remain a functional chooser. Consider the example of a physician deciding to uphold patient dignity against institutional pressures. The physician does not merely occupy a socio-professional role descriptively; she reflectively affirms her identity as a moral agent committed to medical ethics. This active endorsement structures her subsequent choices, giving her a reason to refuse unethical commands because she values her existence under that specific description.

A similar mechanism operates in parental obligation. A parent's care for their child is not an instinctual behavior or a social script passively followed; it is grounded in a reflectively endorsed identity as a protector. If the parent considers neglecting this duty, the resulting internal crisis — manifested as guilt or psychological conflict — is a direct reaction against the literal disintegration of their self-conception. In both scenarios, the internal demand to maintain a unified identity is what imparts authority to the action. These cases illustrate Korsgaard's radical claim: moral agency is a continuous process of self-constitution, where acting responsibly is identical to protecting the very description under which you find your life worth living.

We can now see the core problem Korsgaard's account must address: what gives moral identity its special normative authority over other practical identities? It is here that Korsgaard's account diverges from Kant's original formulation — and it is here that the central difficulties arise.

3. Challenges to Korsgaard's View: An Internal Critique

The critical core of this paper does not target Korsgaard from an external, orthodox Kantian position; she is explicitly constructing an independent metaethical alternative. Rather, the pressing question is whether her constructivist architecture can sustain the very weight it demands, namely a non-contingent, stable grounding for unconditional obligation. When tested against its own structural commitments, Korsgaard's framework reveals a vulnerability as it risks collapsing the impersonal necessity of the moral law into the subjective, contingent domain of individual maxims.

The structural problem lies in the direct link Korsgaard establishes between obligation and identity. If an obligation is fundamentally a reflective reaction against the threat of losing one's practical identity, then the normative authority of the moral law becomes inevitably bound to the agent's desire for internal coherence. This conceptual move introduces a heteronomy. By reducing the moral motive to a mechanism of self-constitution and identity-preservation, the ultimate justification for moral action is deflated into a conditional imperative: "if you wish to remain a unified, coherent agent, you must follow the moral law." In short, pure duty is transformed into a sophisticated form of conditional self-interest. The categorical status of the command is replaced by a functional demand for psychological survival.

This vulnerability becomes clear when we look at how the framework handles moral conflicts. Consider a variation of the traditional "killer at the

door” scenario. An agent who identifies fundamentally as a protector or a loyal friend faces a profound crisis when a killer asks for the whereabouts of a friend hiding upstairs. Within Korsgaard’s apparatus, the choice to lie to the killer is not just an impulsive desire; it is a reflectively endorsed action rooted in a deeply valued, albeit contingent, practical identity. If the agent chooses to tell the truth, they face the disintegration of the very self-conception that makes their life worth living. Because the normativity of the action is drawn from the local identity of “friendship,” the framework struggles to explain why an objective, universal demand against lying should unconditionally override this contingent commitment. As Christopher W. Gowans points out, practical identities are intrinsically dependent on the fluctuating social worlds individuals find themselves in, meaning they vary radically from person to person (Gowans, 2002, p. 548).² By shifting the source of normativity to these individualized structures of self-valuation, Korsgaard opens the door to subjective exceptions, compromising the unconditioned supremacy required by a robust theory of obligation. Because practical identities are socially constructed, they are inherently susceptible to the epistemic defects of their environments—such as ideological conditioning, systemic biases, or closed echo chambers. Consider an agent whose primary practical identity is forged within a highly polarized or radicalized epistemic bubble (for instance, a member of a supremacist group). When this agent steps back to reflect on their desires, the cognitive tools, conceptual frameworks, and social values they rely upon to evaluate their maxims are already structurally corrupted by their environment. In such cases, the mechanism of reflective endorsement does not liberate the agent into objective moral truth, it effectively traps them, rubber-stamping the prejudices of their epistemic community. If the authority of the moral law depends on an identity that can be so easily manipulated by external epistemic distortions, Korsgaard’s framework does not provide a stable, independent ground capable of correcting deeply internalized social errors. This epistemic vulnerability showcases why a robust moral theory requires a grounding that is immune to the fluctuations of social and ideological environments a demand that, as will be shown, can be satisfied by recognizing the unconditional authority of the moral law as a Fact of Reason.

This exact structural weakness is what G.A. Cohen exposes through his challenge of the idealized Mafioso, just as the example of ideologically

2 Thomas E. Hill Jr. (1992) offers an essential discussion on the potential conflict between the formal duty to be truthful and the requirement to respect the dignity of persons, providing a more nuanced Kantian alternative to rigid interpretations of the CI. See Thomas E. Hill Jr., *Dignity and Practical Reason in Kant’s Moral Theory* (Ithaca: Cornell University Press, 1992).

conditioned agent above. Cohen introduces an agent who operates under a strict, deeply internalized code of strength, omertà, and honor. This Mafioso does not lack reflective capacity; he consciously looks at his desires, tests them against his criminal identity, and willingly endorses immoral actions, for instance murder or extortion, as laws for himself to maintain his self-conception. He is reflectively consistent, unified, and perfectly willing to protect his identity at all costs. Cohen's example demonstrates that the formal procedure of reflective endorsement can validate deeply immoral lives just as easily as moral ones. By focusing on the reflective process, Korsgaard captures only the internal phenomenology or experience of feeling obligated, while failing to ground an authentic, objective source of moral authority.

3.1. Korsgaard's Response to the Mafioso Objection and Its Limits

Korsgaard is acutely aware of this challenge, and in her reply lectures, she seeks to neutralize the Mafioso counterexample through a two-layered defense³. First, she argues that local, contingent identities cannot be evaluated in isolation because they are structurally parasitic on a deeper human identity. To pursue any local code, even a criminal one, the Mafioso must exercise rational agency. In doing so, he implicitly values his capacity for choice, which means he must value his rational humanity. For Korsgaard, this foundational value "stands behind our other roles and imparts normativity to them" (Korsgaard, 1996, p. 121). Since valuing one's own humanity logically requires recognizing and respecting the humanity of others, the Mafioso's local identity is caught in a structural contradiction: it systematically violates the very source of value that makes his own choices possible.

Second, Korsgaard relies on a reflective hierarchy to separate local identities from universal moral content. When the Mafioso is pushed by the reflective question "why should I care about my code of honor?", he cannot find a stable answer within the code itself. He is forced to retreat to a deeper level of self-conception, eventually reaching his identity as a human being as such. Because this fundamental human identity carries universal moral content, Korsgaard concludes that "his obligation to be a good person is *deeper* than his obligation to stick to his code" (Korsgaard, 1996, p. 257). If an agent reflects completely and honestly, moral identity must take precedence over any local, conflicting commitments.

3 For Korsgaard's own response to Cohen's Mafioso objection, see her reply in the same volume (Korsgaard, 1996, pp. 204-258). She argues that the Mafioso's error consists in a failure to recognize his own humanity but as argued in the body of this paper, this response presupposes a universalist standard that her constructivist framework cannot itself ground without circularity.

While these arguments are highly instructive, they do not dissolve the underlying problem. The first defense that rational agency inherently carries an implicit commitment to the value of humanity, presupposes the exact universalist moral content that Korsgaard set out to derive constructively. By embedding the moral law into the very definition of rational agency from the start, her response falls into circularity.

The second defense suffers from a structurally identical flaw. The claim that an agent will inevitably prioritize their universal human identity over a local one assumes a specific and predefined hierarchy of reasons. Nothing within the formal, procedural mechanism of reflective endorsement guarantees this ordering. If the Mafioso simply insists that his identity as a member of the family matters more to him than his abstract identity as a human being, Korsgaard's formal framework lacks the internal resources to prove him rationally mistaken. To force the moral identity to the top of the hierarchy, Korsgaard must import a substantive Kantian value choice into what was supposed to be a pure, procedural constructivist derivation.

David Enoch (2006) presses this vulnerability to its logical conclusion with the "shmagency" challenge. Even if we grant Korsgaard that one cannot be classified as an agent without aiming at self-constitution, a normative gap remains: why should I care about being an agent? As Enoch puts it:

Perhaps I cannot be classified as an agent without aiming to constitute myself. But why should I be an agent?... I am perfectly happy being a shmagent - a nonagent who is very similar to agents but who lacks the aim... of self-constitution (Enoch, 2006, p. 179).

Korsgaard confuses a metaphysical necessity (the fact that we find ourselves trapped in agency) with a normative necessity (the requirement to follow the moral laws of agency). Being an agent is a fact we find ourselves in, not a self-vindicating reason to follow any particular law. Building on this, William FitzPatrick (2005) turns Korsgaard's own regress argument against her: if she points to reflective identity as the source of obligation, the agent can simply ask "Why must I care about my reflective identity?" which restates the very regress Korsgaard accused realism of generating (FitzPatrick, 2005, p. 673). Without an objective moral fact to anchor it, reflective endorsement remains a psychological exercise rather than a source of authentic unconditional obligation.

Ultimately, Korsgaard's framework does not bridge the gap between the structural features of agency and the binding authority of the moral law

using its own internal resources. This is not a failure to satisfy Kantian standards from the outside; it is a failure to satisfy the precise standard Korsgaard herself set: providing a stable, non-contingent grounding for moral obligation. Each attempt to resolve the contingency problem, like the appeal to humanity, the appeal to the hierarchy of identities, either borrows from Kantian premises the framework cannot itself derive, or pushes the normative question back one level without answering it. This structural impasse points toward the need for a different kind of grounding: one where the law's authority does not depend on any further fact about the agent.

4. Kleingeld and Willaschek's Challenge: The Limits of the Realist Counter

One natural response to the failure of constructivism is to return to a more objectivist reading of Kantian autonomy. Kleingeld and Willaschek (2019) challenge the subjective, constitutive reading of autonomy, arguing that self-legislation in Kant functions as a counterfactual criterion for determining whether maxims are morally permissible, not as a statement about the origin of the Moral Law's authority. Their interpretation shifts the focus from self-legislation as a metaphysical act of the will to a methodological device that helps us assess the universality and moral permissibility of maxims through counterfactual reasoning (Kleingeld & Willaschek, 2019, p. 7).

On their reading, Kant's reference to self-legislation is instrumental as it articulates a necessary condition for moral legitimacy, namely, that moral laws are such that rational agents could regard themselves as giving these laws rationally, without implying that they actively create or enact them in a metaphysical sense. However, this interpretation struggles with Kant's own language:

According to this principle, all maxims are rejected which are not consistent with the will's own universal legislation. Thus the will is not just subject to the law, but subject to it in such a way that it must also be viewed as self-legislating — and only for this reason subject to the law (of which it can consider itself the author) in the first place (Kant, *GMS*, 4:431).

This passage states that the Moral Law's authority derives from the rational agent's act of self-legislation. The *Urheber* distinction provides a conceptual exit from voluntarism by decoupling the epistemic content of the moral law from its normative enforceability. Under this methodological reading, the moral law remains an objective, synthetic *a priori* template, yet it only binds the finite will because the agent counterfactually recognizes it as

the constitutive condition of rational freedom. Patrick Kain (2004) similarly models this effort, noting that Rawls' famous reading of Kant, that "moral philosophy 'Kant's main aim' was 'to deepen and to justify Rousseau's idea that liberty is acting in accordance with a law that we give to ourselves'" (Kain, 2004, p. 258) supports a legislative rather than constitutive reading of self-legislation.

Yet even on this revised reading, a difficulty remains. If the will is denied an active and constitutive role in creating the law's authority, the CI risks reappearing as an external imposition, the very form of heteronomy that Kant's theory of autonomy was designed to escape. Korsgaard is right to insist that for a law to have authentic authority over a reflective agent, the agent must be able to see the law as a law of her own will, as she writes: "So the reflective structure of human consciousness establishes a relation here, a relation which we have to ourselves. And it is a relation not of mere power but rather of authority. And that is the authority that is the source of obligation" (Korsgaard, 1996, p. 104). Kleingeld and Willaschek's methodological reading mitigates voluntarism but does not sufficiently explain why the law's binding normative force is felt as genuinely one's own rather than as an alien demand.

This dispute between Korsgaard's constitutive architecture and Kleingeld and Willaschek's methodological realism creates a profound metaethical stalemate. Neither account fully resolves the Kantian paradox. Korsgaard secures the agent's motivational ownership of the law, but she does so at the cost of rendering moral authority contingent upon individual self-preservation. Conversely, the methodological reading preserves the absolute objectivism of the law, but it struggles to explain how a counterfactual test can generate a binding, internal necessity that does not feel like an alien imposition.

This impasse reveals that the traditional derivation project, the attempt to infer the authority of the moral law from either psychological reflection or formal procedures of the will, has exhausted its resources. The structural gap between the facts of agency and the unconditional supremacy of duty remains unbridged. It is precisely this crisis that necessitates a radical shift in how we conceive of moral validation. To resolve this dilemma, we must abandon the assumption that the moral law requires an underlying justification, and instead recognize it as a foundational given. This is the precise philosophical turn Kant executes in his mature practical philosophy through the doctrine of the Fact of Reason.

5. The Fact of Reason as Resolution: Schönecker's Intuitionist Reading

The resolution to the stalemate identified above requires stepping back from the framing that both constructivism and the methodological reading share: the assumption that the Moral Law's authority must be *derived* from something else, whether reflective identity or counterfactual universalizability. Kant's own decisive move in the *Critique of Practical Reason* is precisely the abandonment of this derivation project:

Moreover the moral law is given, as it were, as a fact of pure reason of which we are *a priori* conscious and which is apodictically certain, though it be granted that no example of exact observance of it can be found in experience (Kant, *KpV*, 5:47).

Kant asserts that the Moral Law does not need a complex logical deduction from more fundamental principles. It is a foundational truth of reason that presents itself to us with immediate, binding authority, "apodictically certain", and this certainty is not dependent on its conformity with any reflective endorsement procedure. This is why the primary motive behind identifying the law as a "fact" is to secure its absolute, apodictic certainty as a principle that forces itself upon us independently of contingent desires or empirical circumstances.

Dieter Schönecker (2013) develops this into a full intuitionist reading. The validity of the Moral Law is recognized not through deductive reasoning but through a form of self-evidence mediated by the feeling of respect (*Achtung*). Schönecker reinterprets the Fact of Reason as a concept of non-deductive justification, that is we do not derive the law's authority from something more basic; rather, the law's authority is directly given to moral consciousness through respect, and this givenness itself constitutes the justification. This reading is significantly bolstered by Onora O'Neill (2002), who argues that the Fact of Reason must be understood not as a theoretical premise, but as the fundamental, inescapable reality of our autonomy in practical deliberation. By recognizing the Fact of Reason, we acknowledge that autonomy is not something we construct, but the necessary standpoint of practical reason. Schönecker thereby resists the compromise inherent in interpretations that anchor the law in the human will's subjective activity or contingent constitution (Schönecker, 2013, p. 3).

This intuitionist reading, while powerful, requires careful positioning within broader Kantian scholarship, as it converges with yet remains

distinct from other influential accounts. Owen Ware (2014), for instance, emphasizes that the Fact of Reason is an experiential realization of our autonomy (Ware, 2014, pp. 11–13), while Karl Ameriks (2003) frames it as Kant's definitive structural maneuver to secure moral authority against the threat of an infinite regress (Ameriks, 2003, p. 20). Unlike Ware and Ameriks, whose accounts focus primarily on the experiential or logical halting of the regress, Schönecker's "intuitionism" makes a more controversial epistemic claim that the feeling of respect (*Achtung*) itself serves as a non-deductive justification. However, it is precisely this robust and contentious reading that is uniquely equipped to counter Korsgaard's framework. While Ware and Ameriks establish that the regress must stop, Schönecker explains how the law's authority fundamentally bypasses the need for the agent's active endorsement. Against Korsgaard's constructivism, which demands a subjective, procedural act of self-constitution to generate normativity, Schönecker's focus on *Achtung* demonstrates that moral authority is grounded in a fundamental, affective receptivity to the law, a self-evident givenness that renders the very mechanism of reflective identity normatively redundant. As Jenine Grenberg (2013) aptly summarizes, the Fact of Reason is "neither an epistemic mystery nor an act, but instead a forced phenomenological fact" (Grenberg, 2013, p. 190): so, not something constructed through a procedure of the will, but a pre-existing authority that forces itself upon our common moral consciousness.

This resolution come closer to a resolution in comparison to the constructivist alternative for reasons that emerge directly from the critique developed in Sections 3 and 4. The core failure of constructivism is that it cannot close the gap between the structural features of agency and the binding authority of the law without importing substantive Kantian premises that its own framework cannot justify. The Fact of Reason reading dissolves this problem structurally because there is no gap to close because the law does not derive its authority from agency, it presents itself directly to practical consciousness as the condition of possibility of all genuine practical deliberation. The regress that defeated the Mafioso response and the shmagency challenge does not arise, because the law's authority is not grounded in any further fact about the agent that could be refused or found contingent.

One might object that appealing to the Fact of Reason is merely a retreat into dogmatism, a refusal to justify the moral law by simply declaring it a "given" premise. If the central argument of this paper is that Korsgaard's framework fails because it cannot close the gap between agency and moral authority without importing unjustified premises, why should the Fact of

Reason escape the same charge? The response is the following: the difference between constructivism and the Fact of Reason reading is not that one has a stopping point and the other does not. Both must stop somewhere. The decisive difference is that constructivism's stopping point is concealed, while the Fact of Reason's stopping point is explicit. Korsgaard's framework does not ask why reflective endorsement itself should be normatively binding. It assumes that the agent's need for self-coherence generates genuine moral obligation. The shmagency challenge and the circularity objections of Sections 3 and 3.1 reveal precisely this. The Fact of Reason, by contrast, does not pretend to derive the moral law's authority from something more fundamental. It acknowledges that practical justification has a terminus, a point at which one reaches not an arbitrary axiom, but the very condition of possibility for all normative questioning. To ask "why should I accept the Fact of Reason?" is already to invoke the framework of practical reason that the Fact of Reason constitutes. The regress does not continue. In this sense, the Fact of Reason reading is not more dogmatic than constructivism, and it is more transparent about where the grounding must ultimately rest.

Hence, the structural difference between the two positions becomes fully visible. As mentioned, Korsgaard's framework does not explain what makes reflective endorsement normatively binding in the first place. The question "why does the agent's need for self-coherence generate genuine moral obligation rather than mere psychological pressure?" is never posed within the constructivist framework. To answer that question from within the constructivist framework would require appealing to something beyond reflective endorsement itself, which is precisely what the framework misses. The Fact of Reason, by contrast, does not merely declare a stopping point: it explains why the stopping point is where it is. Schönecker's reading through *Achtung* is decisive here. The feeling of respect is not a psychological byproduct of moral cognition but rather is the mode through which the moral law announces its authority to a finite rational being.

Consider a concrete case for it to be clearer: a judge who, under significant political pressure, is about to sign an unjust verdict. At the moment she recognizes what she is about to do, something stops her, not a calculation of consequences, not a reflection on her professional identity, but an immediate sense that this action is simply impermissible. This is the experience of *Achtung*. She does not first consult her practical identity as a judge and then endorse the law's demand; the law's demand presents itself before any such reflective procedure begins. When an agent experiences respect in this way, she is not constructing the law's authority through an act of reflective

endorsement, but she is recognizing an authority that was already there, that presents itself as the condition under which practical deliberation is possible at all. Consider the same judge on Korsgaard's account. What arrests her is the threat to her self-conception as an honest person, a deeply held practical identity. But this means the law's authority over her is conditional because it holds as long as she identifies with honesty, as long as that identity is more central to her self-conception than, say, her identity as a loyal party member. The authority is produced by her endorsement, which means it could, in principle, be unproduced by a different endorsement. Schönecker's reading shows that the judge's experience of arrest is better described differently. The feeling of respect is already given and not produced but recognized. The law feels one's own not because the agent constructed it, but because the agent's rational nature is the very faculty through which the law announces itself. The difference is not merely terminological, it is the difference between a normativity that depends on what the agent happens to value and a normativity that is constitutive of what it means to be a valuing agent at all. The Fact of Reason is not an arbitrary theoretical postulate or an unjustified external imposition but it is the inescapable phenomenological bedrock of practical deliberation itself. The demand for a deeper, deductive derivation of the moral law misapplies theoretical criteria to a purely practical reality. The law is self-authenticating through the immediate, first-personal experience of respect (*Achtung*).

Crucially, this resolution does not simply dismiss Korsgaard's contribution. Her account correctly identifies that moral obligation must be felt as genuinely one's own, not as an alien external imposition, and that the phenomenology of moral consciousness involves an identification of the self with the demands of practical reason. What she cannot explain constructively is why this identification carries unconditional authority. The Fact of Reason reading provides the answer to this, the identification is unconditional precisely because the law is not something we construct but something we recognize, and the recognition, as the feeling of respect attests, is not contingent on any particular practical identity we happen to hold.

When placed alongside other accounts examined in this paper - Korsgaard's constructivism, Kleingeld and Willaschek's methodological realism, and the shmagency challenge - the Fact of Reason reading proves most adequate because it satisfies all three desiderata that any grounding of moral obligation must meet: first, it explains how the law is genuinely the agent's own (against external heteronomy), second, it explains how the law's authority is unconditional (against constructivist contingency), and third, it explains why

the regress of justification terminates (against the infinite regress problem). No other account surveyed here satisfies all three simultaneously.

6. Conclusion

This paper has argued that Korsgaard's constructivist account of moral obligation, for all its ingenuity, fails to achieve its own stated aim. By grounding the authority of the moral law in the agent's need to maintain a coherent practical identity, the framework reduces pure duty to a conditional self-interest: the law is binding only for those who happen to value being agents, and only insofar as following the law is required to preserve the identity they hold. The Mafioso, the shmagent, and the friend-protecting liar all reveal different facets of this fundamental contingency, and Korsgaard's own responses to these challenges, examined carefully, rely on Kantian premises that her constructivist framework cannot derive without circularity.

This critique is not made from outside Korsgaard's project by appeal to Kantian standards she never accepted. It is an internal critique as her framework cannot satisfy the standard it sets for itself. When compared to other accounts of normativity, including the Fact of Reason reading defended here, the constructivist grounding proves to be structurally inadequate.

Schönecker's intuitionist reading of the Fact of Reason, supported by Ware and Ameriks, provides a coherent resolution to the Kantian paradox as it has been reconstructed throughout this paper. Rather than deriving the law's authority from anything more fundamental, whether reflective endorsement or counterfactual universalizability, it recognizes the law as a foundational given of pure practical reason, presented to moral consciousness through the feeling of respect. This reading establishes the CI as a non-negotiable structural truth of our rational nature, not a subjective creation of the reflective mind.

Korsgaard's account nonetheless represents an irreplaceable contribution to the literature. She identifies that the phenomenology of moral obligation involves an identification of the self with practical reason's demands, and she provides the most philosophically serious attempt to show how this identification generates binding obligation. The limitation of her account is the fact that unconditional authority of the moral law cannot be derived from the contingent structure of any mind, however reflective. The original contribution of this paper has been to show that this impossibility is not merely visible from outside Korsgaard's framework as Cohen, Enoch, and FitzPatrick have demonstrated in their respective ways, but is reproduced

within it. Korsgaard's own responses to the Mafioso objection, analyzed closely in Section 3.1, reveal that every attempt to close the contingency gap from within the constructivist framework requires importing Kantian premises that the framework cannot itself ground. This finding strengthens the case for the Fact of Reason reading. If the law's authority cannot be derived even by the most sophisticated constructivist strategy, the most plausible conclusion is that it does not need to be derived it is already given. As Onora O'Neill (1985) reminds us, we must look beyond the psychological activity of the will, for the Moral Law is "a demand which we must take seriously if we doubt that we are, even in the maturity of our faculties, merely consenting adults" (O'Neill, 1985, p. 277).

References

- Ameriks, Karl (2003). *Interpreting Kant's Critiques*. Oxford University Press.
- Cohen, G. A. (1996). Reason, humanity, and the moral law. In C. M. Korsgaard (Ed.), *The Sources of Normativity* (pp. 167–187). Cambridge University Press.
- Enoch, David (2006). Agency, shmagency: Why normativity won't come from what is constitutive of action. *The Philosophical Review*, 115(2), 169–198.
- FitzPatrick, William J. (2005). The practical turn in ethical theory: Korsgaard's constructivism, realism, and the nature of normativity. *Ethics*, 115(4), 651–691.
- Gowans, Christopher W. (2002). Practical identities and autonomy: Korsgaard's reformation of Kant's moral philosophy. *Philosophy and Phenomenological Research*, 64(3), 546–570.
- Grenberg, Jenine (2013). *Kant's Defense of Common Moral Experience: A Phenomenological Account*. Cambridge University Press.
- Hill, Thomas E., Jr. (1992). *Dignity and Practical Reason in Kant's Moral Theory*. Cornell University Press.
- Kain, Patrick (2004). Self-legislation in Kant's moral philosophy. *Archiv für Geschichte der Philosophie*, 86(3), 257–306.
- Kant, Immanuel (1996). *Critique of Practical Reason* (M. Gregor, Trans. & Ed.). Cambridge University Press. (Original work published 1788; cited as *KpV* with Academy edition pagination).
- Kant, Immanuel (2019). *Groundwork for the Metaphysics of Morals* (C. Bennett, J. Saunders, & R. Stern, Trans. & Ed.). Oxford University Press. (Original work published 1785; cited as *GMS* with Academy edition pagination).
- Kleingeld, Pauline, & Willaschek, Marcus (2019). Autonomy without paradox: Kant, self-legislation and the moral law. *Philosophers' Imprint*, 19(6), 1–18. Retrieved from <https://quod.lib.umich.edu/p/phimp/3521354.0019.006>

- Korsgaard, Christine M. (1996). *The Sources of Normativity*. Cambridge University Press.
- O'Neill, Onora (1985). Between consenting adults. *Philosophy & Public Affairs*, 14(3), 252–277.
- O'Neill, Onora (2002). Autonomy and the Fact of Reason in the *Kritik der praktischen Vernunft*. In O. Höffe (Ed.), *Immanuel Kant: Kritik der praktischen Vernunft* (pp. 81–97). Akademie Verlag.
- Reath, Andrews (2006). *Agency and Autonomy in Kant's Moral Theory*. Clarendon Press.
- Schapiro, Tamar (1999). What is a child? *Ethics*, 109(4), 715–738.