

Philosophic

# Fragments

philosophiques

Volume 26, Spring 2010

Undergraduate Journal of Philosophy

Revue de philosophie des étudiants de premier cycle

McGill University

# Acknowledgements

## Editors

Wilson Blakley  
Lara Bourdin  
Edward Brook  
Rachel Rudolph

## Undergraduate Readers

Melissa Gail Cabigon	Éléna Choquette
Anna Cook	Jonathan Fine
Joseph Krongold	Robyn Penney
Emily Perry	Sierra Yates Robart
Rebecca Rodin	Aadam Tejpar
Jonathan Wald	

## Postdoctoral Reviewer

Tracey Nicholls

## Graduate Reviewer

Yvan Tétrault

## Copy Editor

Daniele Degano

Thanks to the Department of Philosophy, the Dean of Arts, the Arts Undergraduate Society and the Philosophy Students' Association for funding.

Special thanks to Mylissa Falkner and Claudine Lefort for their organisational help.



# Contents

<b>Ricœur's Criticism of Autonomy as Self-Positing</b> <b>The Epistemic Mode of Attestation</b> <i>Ilya Shodjaee-Zrudlo</i>	5
<b>Artworks or Artifacts?</b> <b>A Critical Examination of Arthur Danto's "Artifact and Art" through</b> <b>Denis Dutton and Larry Shiner</b> <i>Alysa Batzios</i>	19
<b>Behind the Nietzschean Mask</b> <b>The Pragmatic Epistemology of Value-for-Life</b> <i>David Feldman</i>	33
<b>Science as Regulative Judgments</b> <b>A Discussion of Kant's Response to Humean Skepticism</b> <i>David Brooke Struck</i>	46



# Ricœur's Criticism of Autonomy as Self-Positing The Epistemic Mode of Attestation

*Ilya Shodjaee-Zrudlo*

RICŒUR's dialogue with the history of philosophy aims to draw out the full implications of the Aristotelian analysis of praxis, whilst also bearing in mind the tradition of self-reflection that is marked by the *cogito* of Descartes and the transcendental "I" of German Idealism. Ricœur bases his theory of self-recognition in action, specifically via "a reflection on those capacities which together sketch the portrait of the capable human being" (*Recognition* 91). The term *attestation* is used by Ricœur to "characterize the epistemic mode of assertions having to do with capacities" (*Recognition* 91). A theory of self-recognition grounded in attestation incorporates a certain suspicion towards the concept of autonomy itself, placing it at odds with the pride of the *cogito* and the self-positing of Kant's autonomy.

Ricœur's critique of the self-positing of autonomy is the subject of this essay; a critique that, in Ricœur's own words, fundamentally "doubt[s] the autonomy of autonomy" (*Oneself as Another* 215). Firstly, this essay will summarise Ricœur's understanding of Kant's notion of autonomy; secondly, it will discuss Ricœur's critique of Kant's self-positing autonomy and his reclamation of autonomy as based on attestation; thirdly, it will contextualise attestation within the larger scheme of a hermeneutics of the self; and, finally, it will dwell on a few practical implications of a philosophy of self-recognition grounded in capacity-building.<sup>1</sup>

---

<sup>1</sup>The expression "capacity-building" is not a recurring Ricœurian term, but I will employ it in the final section of the paper and discuss its possible relation and contribution to Ricœur's image of the capable man.

The Kantian notion of autonomy springs from an exploration of the human will and its constraints, within the context of universality. This will is fundamentally characterised by its relationship to law in the tradition of “nomological” morality. As defined by Kant, the will is, due to its very nature, constrained by inclinations. On the other hand, the good will, which Kant defines as that which is “good without qualification” (Kant 61), is one that acts solely out of duty: out of a constraint that is self-imposed. Thus, the good will *transcends*, in a sense, the limitations set by regular inclinations and unlocks the power of practical reason. Autonomy is this self-legislation of the will, a will which imposes upon itself moral constraints (or, positively phrased, *adopts* the appropriate imperatives) and is thus the good will without qualification. Under the banner of this deontological method, imperatives, which answer the question “What ought I to do?” are examined one by one, until the “will that is good without qualification will equal the self-legislating will, in accordance with the supreme principle of autonomy” (Ricœur, *Oneself as Another* 207). How are these imperatives to be examined, and how do we distinguish the categorical imperative from the inclinations, the latter of which are to be rejected? Universality itself returns to judge between imperatives in the form of the test of “universalisability,” by which maxims are accepted or rejected. This process culminates in the discovery of the universal law, of the categorical imperative: “[a]ct only on that maxim through which you can at the same time will that it should become a universal law” (Kant 88). Autonomy is submission to this purely formal law.

Another important question is implied in this discussion of autonomy: if one does not self-legislate through this test of universalisation and abide by the narrow principle of autonomy, what is one left with? Because of the singular (and universal) character of autonomy (the good will), all that is left is the diametrically opposite concept of *heteronomy*. Kant effectively gives us a binary choice between autonomy and heteronomy, the former of which represents true freedom, self-legislation and obedience, whereas the latter embodies complete dependence on inclinations and passivity. As Ricœur comments, the “opposition between autonomy and heteronomy has thus appeared as constitutive of moral selfhood” (*Oneself as Another* 211). Kant has posited a legislating self at this “universal” stage of analy-

sis.<sup>2</sup> Ricœur warns us not to critique the positing of this autonomous, legislating self as an essentially "egological" move, since this would be contrary to Kant's own description of the abstract nature of this level of the analysis. In other words, "the principle of autonomy is held to escape the alternatives of monologue and dialogue," but does *not* imply the positing of a "solitary ego" (Ricœur, *Oneself as Another* 211). However, Ricœur will still critique autonomy for being, in a sense, proud.<sup>3</sup>

Now we will briefly explore the foundation of Kant's principle of autonomy, which is simultaneously the target of Ricœur's first strike against a *strong* autonomy. The principle of autonomy is accepted as a *fact of reason*, due to the "consciousness that we have of the self-legislating capacity of the moral subject" (Ricœur, *Oneself as Another* 212). This consciousness posits autonomy itself, bringing it to a level of certainty that is primordially unquestionable. This fact of reason is established at the level of pure reason, and then it is properly fleshed out and made robust at the level of practical reason, making it a fact of pure, practical reason. Before moving on to Ricœur's critique of a strong sense of autonomy, it is interesting to note that in this critique, "Ricœur remains within a Kantian architectonic" (Anderson 16). Anderson in fact maintains that Ricœur not only never completely discards Kant's three stages of analysis but also actually works within a Kantian framework when identifying and analysing the discontinuities related to the account of autonomy. Ricœur's purpose is not to behead autonomy; on the contrary, he is attempting to "reclaim the autonomy of self as *ipse*, not *idem*; that is, selfhood, not sameness" (Anderson 16), and thus to preserve autonomy in its full sense, but not in its strong sense.

Moving now to Ricœur's identification of the impasses associated with the Kantian notion of autonomy, it is important to keep in mind that Ricœur seeks to reclaim autonomy, not as a fundamental

---

<sup>2</sup>These different stages of analysis mirror the unfolding of the categorical imperative in its different formulations. The stages of analysis are, briefly, unity, plurality and totality. We are dealing here with the level of unity.

<sup>3</sup>Although this is the term reserved for the critique of the Cartesian cogito, I prefer it to the vague meaning of a "strong" sense of autonomy. I will use both terms, "strong" and "proud," to simply denote the distance between the original meanings of the terms "autonomy" and "cogito" and a self-recognition based on attestation, ultimately characterised by acting and suffering.

principle of reason, but as a capacity of the capable and moral human being. In other words, autonomy is recast as the capacity to self-legislate, to command one's own self to obey the universal moral law. This essay will gradually unfold the notion of attestation which is central to Ricœur's critique of a proud autonomy. Moreover, in the context of a continuous dialogue with the history of the philosophy, Ricœur is attempting to soften the so-called radical break with the Aristotelian tradition that Kant had inaugurated. He does this by stressing the implications of universality in the former and revealing the teleological character of aspects of the latter. It is this latter method that Ricœur employs to deconstruct Kant's notion of autonomy, given that he wishes to expose its teleological foundations and the weakness of its claim to adequate selfhood. Ricœur shows that autonomy is not as strong as it claims to be by highlighting three specific aporias that are implied in the Kantian text.

The first aporia of autonomy is the way in which freedom is associated with law, which relates to a deduction made by Kant from the principle of autonomy. Ricœur maintains that the consciousness described above—the consciousness of our capacity to self-legislate, or to freely submit ourselves to the law—is the “specific form taken by the attestation of ‘who?’ in the moral dimension” (*Oneself as Another* 212, my emphasis). This consciousness is taken by Kant to be a “declaration” of autonomy, but Ricœur suggests that there is a concealed element of attestation in this “proud” declaration. Attestation, in the sense of “believing-in” and incorporating a definite “feeling” of suspicion (as will be explored below more thoroughly), is thus at the core of Ricœur's interpretation of autonomy. The positing of autonomy by a fact of reason that Kant describes is revealed to be more akin to the level of certitude associated with attestation. The crux of this aporia lies in the fact that there is an “avowal of a certain receptivity, to the extent that the law, in determining freedom, *affects it*” (Ricœur, *Oneself as Another* 213). Since the law determines freedom, autonomy is, in a sense, *passive* in this relationship. Lying at the very root of autonomy, this passivity contributes to a “weaker” sense of autonomy, grounded in attestation, which Ricœur is seeking to establish, as opposed to an “arrogant” and self-positing autonomy.

Springing from the first aporia, the second explores the role of

respect as *feeling*, which Ricœur believes is affecting the principle of autonomy. The full meaning of respect, of course, only unfolds at the second level of analysis: that of plurality, corresponding to the second formulation of the categorical imperative, which involves the respect owed to persons as ends in themselves and not as means. However, this is the peculiarity that Ricœur wishes to highlight: that respect is often used by Kant at the *first* stage of analysis, at the level of universal law, in the context of the strong respect for the law that autonomy carries. Respect for the law is, in a sense, a motive or an inclination itself, and the will is fundamentally affected by it. This respect implies a certain passivity that Ricœur highlights as the second aporia. As Michel points out, this problem concerns "l'inadéquation entre la pure autoposition consubstantielle de l'autonomie et la dimension de réceptivité impliquée dans le respect" (Michel 312).<sup>4</sup> There is a disconnect between the supposed spontaneity of reason and the affect of respect as a feeling. Abiding by his logic of exclusion, Kant attempts to solve this problem by dividing affectivity itself into two parts: respect and self-love. The former is accepted by Kant as a motive of the good will, whilst the latter is cast aside as perverted inclination. In a very compelling manner, Ricœur interprets these two parts of affectivity as two different versions of self-esteem, which constitutes the *ipseity* (selfhood) of the ethical aim. The type of respect that Kant approves of is understood by Ricœur to be "self-esteem that has passed through the sieve of the universal and constraining norm—in short, self-esteem under the reign of the law" (*Oneself as Another* 215). Self-love, on the other hand, is self-esteem that has been perverted by a penchant for evil, which is the subject of the third aporia of autonomy. To summarise, the second aporia relates to respect, which contributes to a certain sense of passivity. Again, this problematises the possibility of a strong, self-positing autonomy.

For the third aporia, Ricœur seeks assistance from Kant's essay on radical evil, which opens the latter's *Religion within the Limits of Reason Alone*. According to Ricœur's summary of this essay, Kant is arguing that the origin of evil does not lie in desire or inclination. Rather, evil is a reversal, or a perversion, of the respect for the law:

---

<sup>4</sup>"the disconnect between the pure self-positing of autonomy and the dimension of receptivity implied in respect" (my own provisional translation).

it seeks to fulfill desires instead of obey the law. Evil can symbolically be reduced to a single “evil maxim,” which *inclines* one towards evil at the level of choice, and is, in Kant’s words, a propensity for evil. In “ethico-universal” opposition to this propensity for evil is a predisposition for good. These two “forces” seem to be active at the level of choosing maxims, and this propensity for evil most definitely affects our ability to choose freely and thus to be autonomous. The “tug-of-war” between propensity for evil and respect (and perhaps even our predisposition for good) creates a “subjective ground [for] the use of freedom” (Ricœur, *Oneself as Another* 217). Evil is not, however, completely original, nor is it entirely foreign to human nature (Ricœur, *Oneself as Another* 217–18); rather, it is *radical* in that it twists the root of choice at the very foundational level of maxims. Ricœur sides with Kant in ascribing to religion the role of ensuring that this propensity for evil is dealt with, and that its effects are limited. Despite the therapeutic role of religion, the fact that radical evil is lying in wait still places a significant limitation on freedom, or the free will. This “wound,” as Ricœur characterises it, contributes to a weak sense of autonomy. With this final aporia, autonomy is violently brought down from its throne of pride, self-sufficiency and absolute self-positing, only to be reclaimed by Ricœur as a “dependent” autonomy in the third part of the ninth study of *Oneself as Another*, entitled “Autonomy and Conflict,” to which we will now turn.

Reaching the final stages of Ricœur’s critique of Kant, we will trace the “transition” of the three aporias of autonomy from interiority to externality, involving the notions of tutelage and heteronomy. We will then demonstrate the role of attestation by looking at an illustration taken from the second part of the ninth study, in which Ricœur expands on a traditional Kantian thought experiment. Ricœur’s project to reclaim autonomy begins with his reaffirmation of the possibility of a nonegological autonomy. To save autonomy from being warped by ego one has to proceed backwards from the rule of justice, to respect and finally to autonomy, reversing the order of Kant’s own analysis in the *Groundwork*. The autonomy that one finds at the end of this regressive path, however, is fundamentally different. The three aporias encountered above—relating to receptiveness, passivity and powerlessness—that left gaping wounds in

autonomy, are now cauterised, but at the cost of a new, *dependent* notion of autonomy. "Dependency as 'externality,' related to the dialogic condition of autonomy, in a sense takes over from dependency as 'interiority' revealed by these three aporias" (Ricœur, *Oneself as Another* 275). Externality refers generally to autonomy's relationship to a real and concrete other, a relationship which will be explored in the following paragraph. Moving to the opposition between autonomy and heteronomy, Ricœur suggests that freedom is more complex and inherently dependent than a self-positing autonomy. Thus, a fundamental revision of the concept of autonomy is required, a revision that reconsiders the dichotomy created by Kant between autonomy and heteronomy. The latter seems to be indistinguishable from the state of "tutelage,"<sup>5</sup> which consists in having one's judgment in some way determined by the judgment of another. Accordingly, Kant assigns to autonomy the definition "responsibility for one's own judgment" (Ricœur, *Oneself as Another* 275), the binary opposite of tutelage-heteronomy. However, within the context of the rule of reciprocity, one's judgment can in fact be influenced by the judgment of others without being pejoratively labeled as tutelage-heteronomy. Ricœur is arguing that the rule of reciprocity *can* and *should* affect autonomy.

We can begin to see here why Ricœur chose to take the regressive path towards autonomy, beginning from the rule of justice, and how this reveals the dependency of autonomy on a heteronomy of a certain type. Ricœur's reference to the essay on tutelage also reveals the originally political nature of the concept of autonomy that, in a sense, justifies this regressive path. We will now see how the revision of autonomy incorporates the other, both internally and externally. The three aporias of autonomy are each characterised by a certain type of "otherness" that autonomy depends on: "the other of freedom in the figure of the law, which freedom nevertheless gives itself; the other of feeling in the figure of respect; the other of evil

---

<sup>5</sup>This notion derives from an English translation of Kant's essay entitled *What is the Enlightenment?* Essentially, tutelage refers to the situation of being a minor, in which you need a guardian to make decisions for you. Kant's essay is aimed at critiquing this collective age-old immaturity in favour of a *political* notion of autonomy, in which all moral agents dare to voice their opinions in reason socially. This notion of autonomy is, however, obviously different from the autonomy posited as a fact of reason at the first stage of analysis in the *Groundwork*.

in the figure of the penchant toward evil" (Ricœur, *Oneself as Another* 275). These three aporias are thus the aspects of otherness *interior* to autonomy. Autonomy's dependence on the rule of reciprocity and the other in the figure of Lévinas' master of justice are aspects of otherness *exterior* to autonomy. This last element of externality represents the kind of "heteronomy" that Ricœur wishes to incorporate into the new, weakened notion of autonomy. Autonomy has now been rescued from its potential egological predicament, and has been properly harmonised with a robust sense of otherness. The rescuing of autonomy contributes to the ever-present dialectic between identity and otherness at work in Ricœur's hermeneutics of selfhood. Through this de- and re-construction of autonomy, Ricœur has effectively created an autonomy which unfolds dialogically and authentically recognises the other in the strong sense implied by irreplaceability.<sup>6</sup>

To fully tease out autonomy's external dependence on the rule of reciprocity, as well as the role of attestation at the very root of this dependency, it is helpful to turn to Ricœur's careful analysis of the Kantian example of false promises. Ricœur's rigorous examination of the test of universalisation applied to the maxim of false promises is part of a section in the ninth study that examines the conflicts that arise between respect for the law (autonomy) and respect for others (in the second formulation of the categorical imperative).<sup>7</sup> In this particular example, testing the maxim results in a *logical* contradiction at the level of the *will*. However, this logical contradiction is recast by Ricœur as an exception made to the rule "on behalf of the agent by reason of self-love" (*Oneself as Another* 264). Essentially, because the universalisation of the maxim of false promises would be *undesirable* by the agent performing the thought experiment, it is to be rejected as a logical impasse. This brings up the question, which Ricœur asks, of whether one can make an exception in order to respect the other.

Ricœur rejects the idea that the thought experiment above, out-

---

<sup>6</sup>Irreplaceability, as Ricœur explains it, involves authentically understanding or acknowledging that the other, outside of one's self, is irreplaceable; in other words, that every single person is an objective end simply by virtue of being a person.

<sup>7</sup>This is where Ricœur suggests that practical wisdom should be allowed to give priority to others, "in the name of the solicitude that is addressed to persons in their irreplaceable singularity" (*Oneself as Another* 262).

lined by Kant, involves an authentic respect for the other. Ricœur analyses the dialogic structure of promising itself. He shows that attestation plays a key role in promising, and thereby bolsters his critique of the monologic,<sup>8</sup> strongly autonomous structure of the self that is reflected in the application of the second categorical imperative. First, the commitment to keep a promise reflects a strong sense of self-constancy through time that is taken to be constitutive of self-hood. Second, at the risk of falling into a monologic self-constancy, Ricœur explores the idea of *availability*, advanced by Gabriel Marcel, and the tie between the one who promises and the one to whom the promise is made. In making a promise, you make yourself available to the other. In the same moment, the other responds by committing to your response in the form of *counting on you*. Finally, one can reply by asserting that the other *can* count on you. This exchange is steeped in "the moral tenor" of solicitude (Ricœur, *Oneself as Another* 268). Moreover, and this was the ultimate point of bringing up this example, one can see in this exchange multiple examples of assertions related to capacities, all referring directly to attestation of capacities. In making one's self genuinely available, one says: "I *can* be faithful." Conversely, in a true example of the authentic nature of the other in the exchange, one asserts in response: "You *can* count on me." This example illustrates the role that attestation plays in a critique of the limitations of a self-positing autonomy. Self-attestation is at the root of promise making, just as it was at the root of the positing of autonomy itself in the first internal aporia.

We will now take a step back and consider more holistically the epistemological and practical considerations at work in Ricœur's notion of attestation. We will do this in particular by adding to Ricœur's critique of autonomy his critique of the Cartesian *cogito*. In a sense, these considerations, although consisting of a step backwards, bring us directly to the heart of the first aporia of autonomy, wherein Ricœur recognised an essential element of attestation, relating to "the 'who?' in the moral dimension" (*Oneself as Another* 212). Playing on the meaningful connection between attestation and self-recognition, Ricœur proposes that agents of action, in "recognizing that they have done something ... implicitly attest that they were capable of doing it" (*Recognition* 92). The following excerpt from *The*

---

<sup>8</sup>Self-contained, as opposed to dialogical.

*Course of Recognition* illustrates beautifully the way Ricœur suggests that attestation can be meaningfully used in the context of a philosophy of self-recognition:

In intersecting in the certitude and the assurance of the “I can,” the two semantic fields of attestation and self-recognition bring to bear their respective harmonics, in this way lending richness and density to what I propose to call recognition-attestation. From this mixture comes the certitude of assertions introduced by the modal phrase *I can*. (92)

Turning now to epistemological considerations, what kind of “certitude” is being evoked by attestation; in other words, what is the “epistemic mode” of attestation? We will first dwell on the epistemological strengths of attestation, and then move to its weaknesses.

Ricœur wishes to distance himself from foundational philosophies such as Descartes’, which ultimately rely on a notion of *scientific* or absolutely certain knowledge.<sup>9</sup> Attestation is not, however, as distant from foundational knowledge as doxic belief is.<sup>10</sup> Ricœur maintains that “doxic belief is implied in the grammar of ‘I believe-that,’” whilst “attestation belongs to the grammar of ‘I believe-in’” (*Oneself as Another* 21). This encourages Ricœur to exploit the etymological connection between attestation and testimony, lending greater strength to the type of belief implied in attestation. Within the larger structure of the hermeneutics of the self, attestation implies a sort

---

<sup>9</sup>Ricœur uses the term *episteme* to refer to science and this type of knowledge (*Oneself as Another* 21). As an aside, it is interesting to note that the existence of even this kind of knowledge has been the source of controversy ever since the notion of “brute data” or “fact” has been shown dubious by Karl Popper: facts are always stated within the context of a theory, hypothesis or a problem. If one were to accept this account of scientific observation, one would have to place scientific knowledge on the same “epistemological” plane, so to speak, as the rest of knowledge, since observation does not give any kind of *fundamentally* different grounding for scientific knowledge. This would probably bring “belief-in” closer, epistemologically, to the status of *episteme*, or, rather, the latter would be brought closer (in a sense, “down”) to the former. What ramifications a plunging into the domain of the philosophy of science could carry for the epistemic strength of attestation, I do not know, but they could be interesting to pursue. It would definitely alter attestation’s epistemic strength *relative* to that of so-called fact-based foundational philosophies.

<sup>10</sup>Doxic belief, sometimes called “opinion,” is often understood as being diametrically opposed to knowledge.

of verification, which lends it more strength than belief, in that it employs a "detour by way of analysis"<sup>11</sup> (Ricœur, *Oneself as Another* 21) in the stage of reflection. Attestation finally "translates" into "the conviction that selfhood is a dialectic of the self, the same, and the other, and the proof that this is indeed the intimate reflective structure of selfhood" (Venema 161).

The weaknesses of the epistemic mode of attestation are reflected throughout the hermeneutical approach. As implied above, attestation's first weakness is its lack of a fundamental and sure foundation, the hallmark of the Cartesian *cogito*. There is a constant suspicion<sup>12</sup> that surrounds attestation. Referring back to testimony to illustrate this suspicion, Ricœur points out that "there is no 'true' testimony without 'false' testimony" (*Oneself as Another* 22), thus creating a type of "hermeneutical circle"<sup>13</sup> that cannot fundamentally be broken out of, adding to the weakness of attestation. Attestation also carries the same fragility as trust. Finally, within the context of action, attestation "can be defined as the *assurance of being oneself acting and suffering*" (Ricœur, *Oneself as Another* 22).

On one level, the *justification* for an "acting and suffering" self and the weak epistemic strength of attestation in relation to the certainty of the *cogito* is sought in a critique of this same *cogito* as monologic and absolutely egological. This is reflected in the consistent attribution of "pride" to the *cogito*, as it is established by Descartes. Ricœur's first critique of the self-positing act of "I think therefore I am" is that it "remains as abstract and empty as it is invincible" (Jervolino 7).<sup>14</sup> The *real* task of the reflexive self

---

<sup>11</sup>This simultaneously makes it more fragile, in that it imposes an "indirect and fragmentary mode" in the return to the self.

<sup>12</sup>Ricœur specifically uses suspicion as opposed to doubt, since doubt is associated with the Cartesian meditations. In a strictly Ricœurian sense, the "other" of doubt is certainty, and the "other" of suspicion is the epistemic strength of attestation, both of the latter lying somewhere between the former pair.

<sup>13</sup>I use the term hermeneutical circle here in Taylor's sense. The statement: "...there is no recourse against suspicion but a more reliable attestation" (Ricœur, *Oneself as Another* 22), obviously leads to the question of what the standards for reliability could be. In this context belief in general is more starkly contrasted against *episteme*. The hermeneutical circle is further implied in Ricœur's exhortation for the reader to "judge whether the investigations that follow live up to this claim" (23) in the closing section of his introduction to *Oneself as Another*.

<sup>14</sup>Although the critique of autonomy is different in its particulars, it is essentially

is to make “my concrete experience equal to the positing of ‘I am’” (Jervolino 7), which, as Jervolino argues, is fundamentally an act of interpretation, belonging to the realm of hermeneutics. Briefly, the self-positing of the *cogito* does not take into account any genuine sense of otherness, is developed in complete isolation, and is thus an empty concept. However, Ricœur does not, as we have seen, employ the Nietzschean route of “hyperbolic doubt.” Instead, he develops a hermeneutics of selfhood that grounds the self in action, and asserts its *self* in attestation. Again, as was illustrated above, the *kind* of self that is asserted is no longer proud, but shattered; an acting and suffering self.<sup>15</sup>

The role of attestation in the critique of the self-positing of autonomy begins and, in a sense, culminates, or is most clearly expressed, in autonomy’s first (internal) aporia, where Ricœur recognises an element of attestation at the very core of the positing of autonomy. A further illustration of the role of attestation in the critique of autonomy was sought in the example of promise-making. This essay has also attempted to present Ricœur’s general critique of the vacuity of self-positing, and his suggestion of a new ground for the philosophy of self-recognition: attestation, which characterises an acting and suffering self.

As a personal reaction to Ricœur’s notion of attestation, the critique of the self-positing “I” and human capacities, I would like to conclude with two reflections on Ricœur’s thought, and introduce the idea of capacity-building. First, the critique of the self-positing I and the importance of the other seem to completely reject any form of personal development which takes place in isolation.<sup>16</sup> How is one

---

of the same character.

<sup>15</sup>I would like to mention, on a personal note, that leaving aside all the philosophically rigorous argumentation that Ricœur elaborates upon, this description of the self is fundamentally in accord with our common sense experience and the impression we have of the human being as generally reflected in history. We continually move from crisis to victory in what seems to be a condition endemic or, positively stated, inherent to human nature. How else could we progress in the teleological sense? If we attained a perfect understanding all of a sudden, there would be no room for progress, and if success always eluded us, we would have no means of finding direction. The dynamics of crisis and victory allow for progress while closing the door to absolute certainty, both in practice and in understanding.

<sup>16</sup>There are a number of ways to think of such an isolation; the most evident being a solitary form of ascetic monasticism, completely “detached” from the world.

able to truly *recognise* one's self without the other, and thus without the chance to build one's capacities? When can I even articulate the phrase "I can" in complete isolation? Ricœur is very explicit, I believe, about the necessity of the other in one's self-definition, especially in terms of one's narrative identity, that incorporates others' narratives. Second, what Ricœur seems to be less explicit about is the aspect of *process* relating to self-recognition, which I would like to call capacity-building. One's sense of self (ultimately, "l'estime de soi") becomes stronger—not in a proud sense, but in the sense of confidence and reliance—as one's capacities grow. As one becomes involved in the lives of others, as one's narrative weaves into others' in more meaningful ways, and as one suffers with others, one's self-esteem becomes more "concrete."<sup>17</sup> This notion of capacity-building does not translate into some ever-increasing "power" within one's self, in a deeply egological sense, but is rather more akin to a profound understanding of one's limitations, an increase in one's ability to sacrifice with and for others, and fundamentally a strong sense of *humility*. An appreciation for *development*, expressing itself as capacity-building, can, so to speak, "save" Ricœur's acting and suffering self from a true tragedy, by reinforcing the notion of service in the ethical aim.<sup>18</sup>

---

<sup>17</sup>I am alluding, but not completely ascribing, to the Hegelian description of determination, moving from the abstract to the concrete.

<sup>18</sup>As an aside, if reinforcing the notion of service and sacrifice for others in the ethical aim threatens to destabilise the self in favour of the other, in a Lévinassian sense, so be it. I do not believe, however, that the destabilisation of the self is a necessary consequence of this preference for the other. Since self-recognition is to be achieved only through the presence of another, the self is already destabilised from the outset. In a sense, preferring the other above one's self "sublimates" the Golden Rule.

## Works Cited

- Anderson, Pamela Sue. "Ricœur's Reclamation of Autonomy: Unity, Plurality, and Totality." *Paul Ricœur and Contemporary Moral Thought*. Eds. John Wall, William Schweiker and W. David Hall. New York: Routledge, 2002. 15–31. Print.
- Jervolino, Domenico. *The Cogito and Hermeneutics: The Question of the Subject in Ricœur*. Trans. Gordon Poole. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1990. Print.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Trans. H. J. Patton. New York: Harper & Row, 1964. Print.
- Michel, Johann. *Paul Ricœur: Une Philosophie de l'Agir Humain*. Paris: Editions du Cerf, 2006. Print.
- Ricœur, Paul. *The Course of Recognition*. Trans. David Pellauer. Cambridge, Mass: Harvard UP, 2005. Print.
- . *Oneself as Another*. Trans. Kathleen Blamey. Chicago: Chicago UP, 1992. Print.
- Venema, Henry Isaac. *Identifying Selfhood: Imagination, Narrative, and Hermeneutics in the Thought of Paul Ricœur*. Albany, N.Y.: State U of New York P, 2000. Print.

## Artworks or Artifacts?

### A Critical Examination of Arthur Danto's "Artifact and Art" through Denis Dutton and Larry Shiner

*Alysa Batzios*

IN "Artifact and Art," an essay from the catalogue for the *ART/artifact* exhibition held at the Center for African Art in New York City in 1988, Arthur Danto explores the difference between artworks and artifacts. On his view, objects' perceptible properties play no role in determining whether or not they are artworks. Rather, he argues, the difference lies in objects' imperceptible properties, namely those that relate to their roles and statuses in their particular cultural contexts. In this essay, I will critically assess the evidence that Danto uses to support his argument, with a particular focus on the Pot People/Basket Folk thought experiment. In line with Larry Shiner's response, " 'Primitive Fakes,' 'Tourist Art' and the Ideology of Authenticity," and Denis Dutton's scathing critique, "Tribal Art and Artifact," I will argue that Danto's account is weak on three counts: firstly, because it is out of touch with real tribal artistic practices; secondly, because it is based on a humanly impossible example; and thirdly, because it fails to accommodate the importance of the perceptible properties of tribal objects in determining their artistic status, as well as the relevance of perception in understanding "primitive" objects in general.<sup>1</sup> I will in effect suggest that both the

---

<sup>1</sup>In this paper, the term "primitive" is used simply to designate objects made outside of the Western tradition. It is worth noting, further, that the term has a loaded history. It has frequently been used in a pejorative sense to describe objects as less sophisticated or developed, and more "authentic" or "pure" than those produced in the Western tradition. I do not intend to draw or comment upon the connotative history of the term.

perceptible *and* imperceptible properties of objects bear upon their artistic status.

It is worth mentioning that Danto seems to navigate around one of the key problems of his Institutional theory of art—that is, the need to bring an object under a concept of art (as defined by a certain historical set of institutions, practices, etc.) in order to make it into an artwork. In doing so, he offers a more inclusive account of what art is. Nevertheless, I insist that his account would be immeasurably strengthened by considerations of real African tribal artistic practices and traditions.

Danto begins his discussion of the artistic status of African tribal artworks with a critical assessment of the infamous exhibition "*Primitivism*" in *20th Century Art*, which was presented at the Museum of Modern Art in New York in 1984. The exhibition featured over 200 tribal objects, which were displayed alongside 150 modern artworks. One of the central claims of the exhibit was that the creators of the latter works—artists like Paul Gauguin and Pablo Picasso—had been influenced by the traditions of tribal cultures. Notably, Picasso was said to have been inspired by a trip he made to the ethnographic museum in Paris' Palais du Trocadéro in 1907. According to Danto, the tribal objects that Picasso found on his visit were indeed works of art. However, he urges, Picasso's discovery of these objects did not *transform* them into artworks since he, like anyone else, could have only discovered "what is already there but has remained up until then unknown or misrecognized" (Danto, "Artifact and Art" 19). Picasso, furthermore, did not transform these objects into artworks by creating artworks that shared some of their perceptible properties. Finally, Danto insists that the curators of "*Primitivism*" did not turn the objects into art by placing them alongside works of modern art.

The major methodological flaw of the exhibit, for Danto, was its reliance on the idea that tribal objects could—and had indeed—become artistically enfranchised based on their perceptual resemblance to modern artworks and on the fact that they had inspired the creation of those works. Danto's charge is that on this logic, any object could be made into an artwork if an artist created a piece that resembled it or that was in some way inspired by it. What makes the objects of the "*Primitivism*" exhibition artworks is not, furthermore,

the fact that we can take an aesthetic interest in their physical properties. Many objects that are not artworks can provoke an aesthetic response from us; for instance, we can respond aesthetically to a hunter's net. The hunter's net is not, however, a work of art on that basis. On this note, Danto adds that one cannot identify an object as an artwork simply by scrutinizing it from a detached "aesthetic" perspective. Anything, he argues, could be treated in such a way—indeed, even a dead dog's teeth, or a dead woman's eyelids—but that does not mean that anything can be an artwork or that anything can become an artwork by virtue of one's looking at it in that way.

If we cannot identify the difference between artworks and artifacts by perception, then, how can we? Danto suggests that we consider a thought experiment involving two African tribes in order to answer that question. Tribe A and Tribe B live independently of one another, separated by a physical barrier. They are completely out of contact with one another and are unaware of each other's existence. They are known, respectively, as the Pot People and the Basket Folk. Each tribe produces pots and baskets that are, to our eyes, indistinguishable. Both produce baskets that are simple in design and tightly woven, while their pots are elegantly curved in shape. There is no great difference between the external properties of the pots of Tribe A and the pots of Tribe B, nor between the baskets of Tribe A and the baskets of Tribe B. Thus, it is not easy for us to tell simply by looking at these objects which of them have been produced by Tribe A and which by Tribe B.

For Danto, despite their superficial similarities, the two pot-making traditions are significantly different from one another. The pots of the Pot People are highly valued within the tribe as spiritual objects. They are believed to contain magical powers and to embody the essential principles of the Pot People's cosmology. Furthermore, those who create these pots are seen to stand in a special relation to the God of the Pot People, who is known as the first pot-maker. By contrast, the pots of the Basket Folk do not carry any special significance for the group. They are valued for their ordinary utilitarian function and nothing more. Their creators are regarded simply as artisans. On the other hand, the baskets of the Basket Folk, like the pots of the Pot People, are objects of spiritual worth to the tribe. The Basket Folk cosmology is constructed around the notion of the

basket and baskets are believed to embody the principles of the universe. Moreover, basket weavers are revered in the Basket Folk tribe in a similar way that pot-makers are in the Pot People tribe. Danto concludes that, "if they had the words," the Basket Folk and the Pot People would describe their baskets and pots as "artworks," respectively, and their pots and baskets as "artifacts," respectively ("Artifact and Art" 23). The grounds for this conclusion will be made explicit later on.

Danto imagines that the Fine Arts Museum and the Natural History Museum of a western country acquire these objects. The baskets of the Basket People and the pots of the Pot People are, for him, appropriately placed in the Primitive Wing of the Fine Arts Museum, while the pots of the former and the baskets of the latter are placed in the Natural History Museum. The Natural History Museum chooses to create dioramas of scenes of daily life of the Basket Folk and the Pot People. A class of children comes to visit the museum and views both dioramas. A guide informs them that the baskets of the Basket Folk and the pots of the Pot People are artworks. The problem arises when one of the children proclaims that she can see no difference between the baskets or between the pots in either diorama. The guide insists that there *is* indeed a difference, but that he cannot explain it in terms of the physical properties of the objects. The two objects are so alike externally that it would be "impossible, or nearly impossible," to distinguish the difference between two based on their appearances. The guide assures the child that the distinction between the two could be made, but only by the "experts"—presumably the museum's curators and anthropologists. Danto, furthermore, dismisses the possibility that we could arrive at this distinction by scientifically probing each object: "proton-bombardment, chemical traces, vegetable differentia" will not help us identify the difference ("Artifact and Art" 26). Moreover, even if such extreme measures *did* indeed allow us to identify physical differences between the two objects, Danto holds that these differences could not ground the difference in the objects' artistic statuses.

Danto uses this example to explore what he calls the "absolute" difference between a tribal artwork and an artifact. If we cannot distinguish the two in terms of their perceptible properties, then we must be able to do so by reference to their imperceptible properties.

He argues, then, that the absolute difference between artworks and artifacts relates to the differences in their roles and statuses in their given societies. To be an artwork, Danto argues, is to possess higher meanings, to have a content, to embody a concept or a thought. It is to be part the "Realm of the Absolute Spirit."<sup>2</sup> For Danto, artworks are most often treated with great respect in their cultural and historical contexts. To appreciate an artwork, therefore, one must grasp the meanings, content, concepts or thoughts that it is meant to express or embody. To do so, one needs to have knowledge of the work's history of making and cultural context of production. Artifacts, in contrast, do not possess meanings, have content, embody concepts or thoughts. At most, they can serve utilitarian purposes; nevertheless, they remain disposable. They are unsophisticated components of the "Prose of the World" (Danto, "Artifact and Art" 26). Although one can certainly take an aesthetic interest in their external qualities, one still cannot appreciate them as one could an artwork. Knowledge of these objects' history of making and context of production is sufficient to explain why this is the case, and it is necessary if one is to correctly identify them *as artifacts*. If the distinction between artworks and artifacts is absolute, then nothing can be done—especially by the curators of "primitive" exhibits—to change their artistic or artifactual status.

In the last footnote of the essay, Danto notes that he is essentially carrying forth a line of philosophical analysis that parallels one found in his seminal 1968 volume, *The Transfiguration of the Commonplace*. Specifically, Danto's art versus artifact problem is similar to the art versus forgery problem that he lays out in the latter text; namely, if a forgery and an artwork are indiscernible in terms of their visible, external properties, what distinguishes one from the other? Danto argues that the answer, as in the art versus artifact problem, is in the imperceptible properties of the objects.

In the case of artworks vs forgeries, this involves several inter-related considerations. For something to be an artwork, it must be made into a particular work of art. A particular work of art is one that has been brought under the concept of "art." For Danto, the

---

<sup>2</sup>For a critique of Danto's use of Hegelian concepts in this article, see Shiner's " 'Primitive Fakes,' 'Tourist Art' and the Ideology of Authenticity,' " especially pages 231–2.

concept of “art” is shaped by the history of art making and art theory; accordingly, it changes over time, following changes in artistic production and theory. For something to be brought under the concept of art, the concept of art must be able to accommodate it as an instance of art. For example, a pile of hemp *à la* Robert Morris would almost certainly not have been a work of art if it had existed in Antwerp in the seventeenth century (Danto, *Transfiguration* 57). For Danto “certain artworks simply could not be inserted as artworks into certain periods of art history;” an object’s status as an artwork depends upon the historical possibility of its having been brought under the concept of art (*Transfiguration* 57). Furthermore, bringing an object—a pile of hemp or anything else—under the concept of art also involves an act of artistic intention. To perform this act, one must be closely aware of the concept of art (*Transfiguration* 60). Morris could not have made his pile of hemp into a work of art, had he not been somehow aware of the concept of art at the time; that is, had he not been familiar with the history of art and its theories as well as with the possibility of making something like a pile of hemp into a work of art. In that sense, Danto’s theory of art is institutional.<sup>3</sup> An artwork cannot be produced outside of a context of artistic practices, theories and traditions. Furthermore, an object’s status as an artwork is determined precisely by reference to this context—thus, by reference to its imperceptible properties. Consequently, forgeries cannot have the same artistic status as the works of art they are meant to resemble. Forgeries are missing the essential imperceptible properties that make things works of art.

There is an important difference to note between Danto’s two texts. In “Artifact and Art,” Danto seems to get around one of the major problems that arise out of the *Transfiguration*: namely, the need for an individual to be working with a concept of art in order for his creation to be an artwork. As we have seen, in “Artifact and Art,” Danto argues that something is a work of art if it expresses, embodies or represents a higher spiritual meaning or content for a given culture. Given that he makes no reference to the necessity of

---

<sup>3</sup>Danto’s Institutional theory of art also depends upon the notion of an “artworld.” The status of an object as an artwork depends on its relation to the artworld, which is defined by the existence of a tradition of artistic practices and theories. The concept of art is similarly defined in relation to the “artworld” in which it is operative.

a “concept of art”—defined, as we have seen, in relation to the kind of tradition of art history, theory and institutions that we find in the West—he appears to have effectively manoeuvred around a crucial problem in understanding and discussing the works of art of other cultures that do not share this Western concept of art. An object need not be brought under a concept of art in order for it to be an artwork.

While I do not wish to dismiss this move towards a more inclusive approach as insignificant, I intend focus the rest of this essay upon the problems with Danto’s account. I will base my argument largely on two critical responses to Danto’s “Artifact and Art:” Larry Shiner’s “ ‘Primitive Fakes,’ ‘Tourist Art’ and the Ideology of Authenticity ” and Denis Dutton’s “Tribal Art and Artifact.” I will use some of the evidence raised in both of these responses to mount a case against Danto’s claims. My critique has two parts. In the first part, I refer to the examples of real African tribal artistic practices presented in Shiner and Dutton’s papers in order to ask whether it is possible for Danto to posit an “absolute” difference between artworks and artifacts.<sup>4</sup> The second part of my critique involves an assessment of the Pot People/Basket Folk thought experiment, which is similarly inspired by Dutton’s and Shiner’s views.

In the debate over the status of primitive artifacts and artworks made outside of the Western tradition, scholars have tended to bring in specific anthropological evidence to defend their claims. The absence of specific anthropological evidence, I argue, is precisely one of the major weaknesses of Danto’s essay. In his essay, Danto claims that he intends to grapple with issues relating to African tribal artworks and artifacts, yet he does not offer a single example of real African artistic practices to support his argument. Instead, he uses a thought experiment to argue that the difference between artworks and artifacts is absolute. However, this experiment does not provide any useful support for Danto’s argument. It illustrates, instead, some of the problems involved in making an absolute claim about the difference between artworks and artifacts, and, more generally, about the status of objects produced outside of the Western context.

---

<sup>4</sup>The first part of my critique is not exclusively dependent on the arguments put forth by Dutton and Shiner in their papers. That is, it is more of an expression of my views than those of Dutton or Shiner.

Examples described in anthropological studies of specific African tribal art practices throw into question the absoluteness of Danto's distinction between artworks and artifacts. For one, some tribal objects can embody thoughts, express higher meanings, have contents, while also serving a specific utilitarian purpose that draws in no way upon those meanings or contents. In addition, in some cases, an object's status or role within a society can change depending on its context. Shiner raises the example of objects that are only spiritually meaningful to the members of a small-scale African tribe during the ritual ceremony in which they are actively used as part of a dance. Outside of the ritual, they lose all meaning and are often abandoned or thrown away.

Let us suppose that a curator adopts Danto's method of distinguishing artworks and artifacts and maintains that the difference between the two is absolute. For the sake of adhering to this approach, they might argue that this object at the very least has the potential, through its involvement in a ritual performance, to be imbued with spiritual significance. Accordingly, they decide to display it in a fine arts museum as an artwork. As Thomas McEvilley argues in "Doctor, Lawyer, Indian, Chief," this move would be inappropriate, ignorant, childish, and "absurd" (48). For McEvilley, it would be equivalent to a non-Western tribal group taking items from our garbage and displaying them in a museum as art objects of spiritual value.

There are also examples that involve the even more complex issue of cultural authenticity.<sup>5</sup> Objects that were once exclusively created for spiritual purposes are now made for sale to tourists and other "outsider" buyers. They result from a negotiation of traditional art practices in an "inter-tribal and inter-cultural world" (Shiner 232), where the makers and their tribes often struggle for cultural and economic survival. Shiner describes a specific example. A carver who produces ceremonial masks for sacred rituals in her tribe is asked to produce two masks for museum curators who intended to display them in a fine arts museum. It is not clear whether Danto would classify these objects as artworks or artifacts. While they exemplify some of the tribe's ritual and artistic traditions, can they

---

<sup>5</sup>I will not go into the details of this richly complicated issue. For interesting observations, see the Shiner article.

be viewed as artworks if they are uprooted from the context that gives them their meaning? Are they artifacts because they do not possess the meanings they would have possessed if they had been produced specifically for ritual or spiritual purposes? These examples not only illustrate the problem with making the types of strict distinctions between artworks and non-art objects that Danto would have us make; they also show how far Western and African cultures are from understanding one another.

Furthermore, not only does Danto fail to include relevant examples of specific African artistic practices in his account, the example that he does offer raises more problems than it resolves. Let us put aside for a moment the concern about Danto's categories of artworks and artifacts. Let us suppose that there is an absolute difference between artworks and artifacts as Danto describes and that this is made explicit in his Pot People/Basket Folk thought experiment. In his response to Danto's article, Dutton argues that Danto's example is weak not because it makes a claim to absolute difference between the two categories, but because it is profoundly unrealistic; for that reason, it is unhelpful for understanding African tribal artistic practices. He however admits that it would be wrong to dismiss a potentially instructive philosophical thought experiment because it seems outlandish or logically impossible. According to him, the value of these kinds of experiments comes either from the way they throw into light something in real life or from their capacity to exaggerate the impossibility of whatever they exemplify (Dutton 17). While Danto's example is not logically impossible on Dutton's view, it is "*humanly impossible*" (my italics) and, again, out of line with the reality of some African artistic practices (17).

For Dutton, the issue depends on our understanding of the patterns of human behaviour involved in artistic production and reception. On his view, it is simply humanly impossible that an art tradition could be visibly indiscernible from an artifact tradition. If pots have specific spiritual meaning for the Pot People and are the most prized of the tribe's possessions, then they would probably be made with regard to a canon of pot-making excellence. Since people tend to take great care in producing items that have spiritual significance to them, Pot People probably pay careful attention to the decorative details they include and to the kinds of clay and firing techniques

they employ when they produce their pots. By contrast, Basket People would probably not pay as much attention to the decorative elements or to the kinds of materials they use to create their pots. Danto makes the mistake of ignoring the reality of tribal artistic production, in which objects that have the greatest spiritual significance are often identifiable by their rich perceivable content.

Could we then assume that Pot People pots and Basket People pots are visually indiscernible from one another? It is possible that two groups could produce objects that, to our eyes, would be indiscernible. But could we maintain that no one could see the difference between them, not even the Pot People or Basket Folk? Dutton suggests that Danto crucially omits this group of experts from his thought experiment. Consequently, Danto also fails to explore the relevant “and determining role of perception [of someone properly-acquainted with the traditions of non-Western cultural production] in understanding primitive art” (Dutton 20). For Dutton, differences between tribal artworks and non-artistic artifacts can be seen with the eye. For example, there are visible differences between African tribal carvings that are produced for ritual purposes and are deeply meaningful to the tribe, and the nearly identical carvings sold as tourist trinkets. The carver intentionally expresses the ritual object’s meanings through its visible properties. For instance, in a given tribe, the carvings meant to appease the spirits may have both initiation and dorsal scars, while the ones made for sale will be made intentionally without dorsal scars. To the carver, the spirit will see that this detail has been omitted. Moreover, the spirit will also notice that the tourist carving is made with a different kind of wood. The person with an untrained eye will miss these differences. With training, however, they can learn to perceive them. For Dutton, understanding primitive artistic practices and distinguishing artworks from artifacts therefore requires gaining knowledge of the relevant artistic and cultural traditions. Once we learn these traditions as best as any “insider” can, we will be able to detect the crucial aesthetic qualities that are intentionally placed in these objects and understand the ways in which these properties indicate the object’s artistic status.

A counter-objection can be raised against Dutton’s line of argumentation. There is no question that Dutton is indebted to many of

the observations made by Nelson Goodman in his response to the forgery problem, and he repeatedly refers to Goodman in his article. In *Languages of Art*, Goodman argues that the crucial differences between forgeries and the works they are meant to imitate can be perceived by the trained eye. In light of its similarities to the latter argument, Dutton's argument merits the kind of rebuttal that Danto offers against Goodman in *The Transfiguration of the Commonplace*. Danto suggests first that Goodman (like Dutton) sets up a logical line of thought that follows: "If a is not identical with b, then there must be a property F such that a is F and b is not F" (*Transfiguration* 43). Danto suggests that Goodman jumps to conclude that the property F must be a perceptual property. For Danto, Goodman's greatest error is in assuming that all aesthetic differences are perceptual differences. Goodman makes the mistake of dismissing the relevance of non-perceptible properties, like history of making and context, in understanding the distinctions between artworks and forgeries.

Though Dutton owes a lot to Goodman, he seems to have successfully navigated around this problem. For Dutton, we must take an object's perceptible properties into account if we are to make distinctions between tribal artworks and artifacts. However, Dutton does not maintain that differences in perceptible properties are the only properties we must consider in understanding primitive artworks and artifacts and in making distinctions between artworks and non-artistic artifacts. Other factors—namely, imperceptible properties, like the object's history of making, significance and role within its specific culture—also matter and must also be learned if we are to understand these objects. For Dutton, Danto's account is correct on this account; however, it ultimately fails by discounting perceptual differences as unimportant and irrelevant in considerations of tribal artistic practices. Indeed, it seems that Danto wants to argue that all relevant aesthetic differences are non-perceptible differences. For Dutton, however, both perceptible and non-perceptible differences are relevant.

In line with Dutton and Shiner's critiques, I have argued that Danto's account of the differences between tribal artworks and artifacts is problematic. It is out of touch with real tribal artistic practices and it is weakened by its heavy reliance on the humanly impossi-

ble example about the Basket Folk and Pot People. Danto fails to accommodate the importance of the perceptible properties of tribal objects and the relevance of perception in understanding primitive art practices. Though Danto avoids one of the major philosophical issues of his Institutional theory and thereby presents what is ostensibly a more inclusive view of art, his account nevertheless suffers from his misunderstandings of real African artistic traditions and practices. A better approach to tribal artifacts and, more generally, non-Western artifacts, would arguably involve a consideration of these objects' imperceptible as well as perceptible properties.

## Works Cited

- Danto, Arthur C. "Artifact and Art." *Art/artifact*. New York: The Center for African Art, 1988. 18–32. Print.
- Dutton, Dennis. "Tribal Art and Artifact." *The Journal of Aesthetics and Art Criticism* 51.1 (1993): 13–21. Print.
- McEvilley, Thomas. "Doctor Lawyer Indian Chief." *Art and Otherness: Crisis in Cultural Identity*. Kingston, NY: Documentext/McPherson, 1992. 26–56. Print.
- Shiner, Larry. "'Primitive Fakes,' 'Tourist Art' and the Ideology of Authenticity.'" *The Journal of Aesthetics and Art Criticism* 52:2 (1994): 225–234. Print.



# Behind the Nietzschean Mask

## The Pragmatic Epistemology of Value-for-Life

*David Feldman*

“Sir,” the philosopher will perhaps give him to understand, “it is improbable that you are not mistaken, but why insist on the truth?”

---

*Beyond Good and Evil* §16

IN *Beyond Good and Evil* §25, Nietzsche counsels fellow free-spirited philosophers to “flee into concealment . . . always have your masks and subtlety, that you may be mistaken for what you are not, or feared a little” (226). And indeed, perhaps no thinker has been more widely mistaken for what he was not than Nietzsche himself: his work has been used to justify militarism, fascism, misogyny, and worse. The hints that Nietzsche might at times be having us on, intentionally leading us astray with stylistic ambiguities, circuitousness, self-contradiction, obscurantism, and riddling—epitomised in the recurring image of the mask—may justifiably frustrate us. But what if Nietzsche’s masks were something more than either a failure or a refusal to communicate? In the following pages I show why the most common interpretations of the problematic mask image are deeply inadequate and propose a new one, firmly grounded in the spirit of Nietzsche’s value theory, which has profound implications for our understanding of his epistemology and metaphysics, and of the Nietzschean project as a whole.

W. D. Williams suggests that the mask image is Nietzsche’s *warning* to the reader that “his writing is never to be taken at face value”

(83); for Kaufmann, this means that even concepts like “beyond good and evil,” “the will to power,” and others “may be masks that elicit reactions quite inappropriate to what lies behind them” (note in Nietzsche, *Beyond Good and Evil* 241). In a similar vein, Williams imagines that Nietzsche’s use of metaphor, dramatisation, and hyperbole is a way of forcing the reader to “follow the elusive scent that is being laid and attempt to discover the ‘real’ Nietzsche underneath the mask” (84). But we should be careful. After all, why should we be so sure that there is some “hard nugget of naked honesty” (Williams 91) behind the mask at all?

The image of the mask, representing posture, persona, deception, illusion, and concealment, is of particular interest to the reader of Nietzsche precisely because it is so emblematic of the paradoxical nature of Nietzsche’s thinking. Ultimately, the mask is problematic because it cannot help but imply some true identity lying behind it, in the same way that Nietzsche’s use of metaphor and ambiguity leads Kaufmann and Williams to seek the “real” philosophy that lies beneath the cryptic gnomicism and mischief. However, Nietzsche telegraphs his intention throughout his works to transcend the notions of fixed “reality” and positive “truth” entirely; and this confounds the function of masks in general. To study the mask image, which may be understood to stand for Nietzsche’s guarded approach, we must first study Nietzsche’s epistemology. If we are to correctly understand Nietzsche’s concept of truth, we must find a way to reconcile the mask with the epistemic doctrine of perspectivism, which holds that nothing is fact and everything is nuance and interpretation. In this paper, I suggest a new understanding of the mask image, which is more comprehensive than the one endorsed by Kaufmann and Williams. To do this, I place the mask in context as part of the will to illusion, the Apollonian dream-image, and the necessity for life of keeping “the truth” in concealment. The secret function of Nietzsche’s masks, I contend, is *not* to conceal difficult truths, but to protect the *usefulness* of certain falsehoods, which, though false, are nevertheless valuable while properly masked. Moreover, I show why a strict perspectival epistemology is ultimately an untenable position, even according to Nietzsche’s own understandings of truth and falsity. I go on to explain the important role it plays in his philosophy in spite of this untenability. As we shall see, Nietzsche’s

ontological and epistemological claims are intended not to *replace* our existing cognitive paradigm or epistemic schema, as his dramatic style often implies, but rather to force the question of value-for-life in order to *improve* that paradigm and help to create a healthier culture, society, and individual.

Typical explanations of Nietzsche's use of the mask image are armchair psychoanalysis—cf. Williams' "defence-mechanism" (91) thesis and de Almeida's musings that "perhaps foolishness was the last rampart behind which Nietzsche could entrench to better laugh at the world, to better laugh at himself" (83). These do not extend much further than to caution the student of Nietzsche to bear in mind that the thinker might not always be perfectly sincere. Although it is true that in the passage quoted above and elsewhere Nietzsche writes of the philosopher's need for protection, in my view there is no justification for, or benefit in, assuming that he refers to his own insecurities. Indeed, he denies as much when he declares in *Ecce Homo* that "I am one thing, my writings are another matter" (715). Moreover, if the mask image is merely Nietzsche's coy sign that there is more to his work than meets the eye, with no significant role to play in the substance of the philosophy itself, then how shall we determine what is sincere and what is posturing? Kaufmann's and Williams' "cautious" interpretation gives the reader free reign to assign meanings to the text at his or her whim, disregarding any point of contradiction as red herring or smokescreen from the very beginning. "The worst readers," Nietzsche complains, "are those who proceed like plundering soldiers: they pick up a few things they can use, soil and confuse the rest, and blaspheme the whole" (*Aphorisms* 155). We should avoid this dead end of inquiry and instead be bold enough to try to take the mask seriously and holistically, on Nietzsche's own terms.

The essential difficulty with the image of the mask is epistemological, not psychological: it arises in Nietzsche's rejection and redefinition of the traditional notions of truth and falsity. In order to understand the recurring mask image, let alone the claim that "whatever is profound loves masks," we must first lay some groundwork to appreciate Nietzsche's notions of perspectival truth and falsity (*Beyond Good and Evil* 240). Our first question must be: how can the mask, whose nature is to deceive and conceal, exist in the absence of

an objective, statically true world, and what is its meaning then? To be precise: in what sense is the mask compatible with Nietzsche's famous formula in *The Will to Power* §481: "facts is precisely what there is not, only interpretations[?]" Is it compatible at all?

Perspectivism is the epistemological position that is the culmination of Nietzsche's criticisms of the "inherent limitations" of the current cognitive paradigm (i. e., the assumptive basis on which we perceive reality and thought), in particular, the "faith in opposite values" and the "fictions of logic;" and it is consequently of crucial relevance to the problem of masks and to the will to truth and illusion. Grimm shows how perspectivism derives from "the flux ontology of power-quanta" (45) which is the ontological significance of the will to power:

The correspondence theory must presuppose a world of (relatively) stable objects and facts about which we form statements or form ideas, and to which these statements and ideas must adequately correspond if they are to be "true" ... [But for Nietzsche], what there is is will to power, and since the will to power is in constant, turbulent flux, there exists nothing stable to which our statements or concepts could be said to apply. Since within this scheme there can be no substances, entities, things, fact, objects, etc. ... there is nothing which could ground that relationship of intentionality, i. e. the subject-object relationship, which must be valid if ideas and statements are to correspond to anything else. Facts and objects are idiosyncrasies of our "optics," not independently existing "things." (46)

This "flux ontology," characterised by the perpetual striving of distinct quanta to increase their power, leads directly to Nietzsche's "perspectival" epistemology. With this perspectivism, the correspondence structure of truth collapses in the face of the transient, turbulent character of reality implied by the ontology of the will to power. Every apparent state of reality is reduced to a mere subjective perspective, and thus there is a real sense in which, positively speaking, "everything is false" (Danto 35). Nietzsche says as much in *The Will to Power*: "There are many kinds of eyes ... and consequently there are many kinds of 'truths', and consequently there is no truth" (§540).

This ontological position, however, is negated by its own epistemological consequences, for it is only one of many false perspectives on a world in perpetual flux. If nothing is statically true, including the position that nothing is statically true, then why bother to write philosophy? More to the point, if perspectivism can only be explained by tearing down the entire structure of logical identity and non-contradiction, then how can the idea of truth persist? Somehow Nietzsche's texts are full of reference to truth and, in particular, falsity: in fact much of his creative energies were devoted to defrocking ideas that he saw as falsehoods that had come to be held as true because of tradition and weakness, including systematic metaphysics, Christian morality, and positivism. In the opening of the first essay in *On the Genealogy of Morals*, Nietzsche writes in praise of

fundamentally brave, proud, and magnanimous animals, who know how to keep their hearts as well their suffering in bounds and have trained themselves to sacrifice all desirability to truth, *every* truth, even plain harsh, ugly, repellent, unchristian, immoral truth.—For such truths do exist. (461)

References to truths like these are common in Nietzsche's work. As we have seen, the idea of positive, correspondence-based truth is incompatible with the kind of ontology of constant change Nietzsche seems to propose. Nietzsche writes as though his perspectivism is "true." However, invoking the formula "fact *is* precisely what there is not," he cannot mean that perspectivism exists positively, in the sense that it corresponds to fact, because in order to be consistent, his ontological and epistemological claims must only be one interpretation of a reality that may or may not be in constant flux (Nietzsche, *Will to Power* §481). If these claims are too forceful then they collapse. When he says logic is "false," he cannot mean that it does not exist, for objective falsity, too, relies on a world of fact, which Nietzsche denies. In spite of Nietzsche's commitment to a world without facts, a distinction is evidently made in his thought that preserves the significance of the language of truth and falsity.

We begin to see that Nietzsche's use of the terms "true" and "false" is different from those we are used to. Danto argues that Nietzsche employs a pragmatic criterion for truth, such that "the

criterion was always and only whether any of the structures ... enhanced and facilitated life" (32). This is instructive as a starting point, and it agrees with famous passages like "truth is the kind of error without which a certain species of life could not live. The value for life is ultimately decisive" (Nietzsche, *Will to Power* §493). According to this interpretation, Nietzsche's claims fall into place not so much as strictly ontological or epistemological, but as moral-behavioural. Whatever makes man "stronger, more evil, and more profound; also more beautiful" is called "true" (Nietzsche, *Beyond Good and Evil* 426); whatever denies life, makes man into "a herd animal, something eager to please, sickly, and mediocre," is called "false" (Nietzsche, *Beyond Good and Evil* 266). Those interpretations which enhance life should be held and those that do not, such as metaphysics, Christian morality, and positivism, should be thrown away. If a strict perspectival epistemology, founded on a "flux ontology of power-quanta," enhances life, then let us call it "true." This characteristic of "value-for-life" is not the same as comfort or ease, of course:

[M]aking unhappy and evil are no counter-arguments .... Perhaps hardness and cunning furnish more favourable conditions for the origin of strong, independent spirit and philosopher than that gentle, fine, conciliatory good-naturedness and art of taking things lightly that people prize .... (Nietzsche, *Beyond Good and Evil* 239-40)

This pragmatic criterion is the core of Nietzsche's theory of truth.

This new kind of truth, then, is a property of those particular positively erroneous propositions which are valuable to life. Nietzsche is perfectly comfortable to *mix* his new pragmatic kind of truth with the traditional positive one. Sometimes, as above in *The Will to Power* §493 ("truth is the kind of error without which a certain species of life could not live"), he even uses both in the same phrase. A thing can be a positive error but a pragmatic "truth." In fact, this is the standard to which he suggests all errors should be held. This is why he sneers at the *false* errors invented by metaphysicians, moralists, democrats, and Christians: their claims are positively false, like everything, but also life-denying. Nietzsche's criterion requires that these positive errors affirm life: the "dangerous maybe" that is the

keystone of *Beyond Good and Evil* is an abbreviation of the criterion: *Maybe this is positively true and maybe it is only one perspective; but, more importantly, does it enhance life?* (201). In this way we can understand Nietzsche's reminder that "the falseness of a judgement is for us not necessarily of an objection to a judgement," and his exhortations for a philosophy that would "recognize untruth as a condition of life" (Nietzsche, *Beyond Good and Evil* 202).

Truth, then, is for Nietzsche a pragmatic quality; it signifies that which enhances and cultivates human life. So the epistemological foundation for our question of masks is laid. Why is pragmatic, life-affirming truth shameful, delicate, or otherwise worth concealing? Why does *Beyond Good and Evil* begin with a caution against the will to truth? Above all, why should the will to truth, understood as the will to discover techniques for enhancing and cultivating life, ever be hidden from view? In his preface to the second edition of *The Gay Science* (1886), Nietzsche criticises those who "endanger temples by night, embrace statues, and want by all means to unveil, uncover, and put into a bright light whatever is kept concealed for good reasons" (§4). How can we understand this kind of sentiment? Why should we mask truth, or veil it, or decline to shine light on it? Why is it "characteristic of more refined humanity to respect 'the mask' and not to indulge in psychology and curiosity in the wrong place" (Nietzsche, *Beyond Good and Evil* 411)? Why is shame "profound?" In the context of a correct understanding of Nietzsche's conception of truth, this is baffling. How can this be reconciled with his obvious goal to promote that which promotes life? Is not his rallying call to free spirits—"let us clench our teeth! let [sic] us open our eyes and keep our hand firm on the helm!" (Nietzsche, *Beyond Good and Evil* 221)? And the mystery persists: there is a connection in many of Nietzsche's writings from the veil to *decency*, nakedness, and the truth as a woman whose modesty must be respected. This kind of imagery seems to support the claim that there is something about the will to truth "at any price" that is distasteful, cowardly, or even dangerous:

Today we consider it a matter of decency not to wish to see everything naked, or to be present at everything, or to understand and "know" everything. "Is it true that God is present everywhere?" a little girl asked her mother;

“I think that’s indecent”—a hint for philosophers! One should have more respect for the bashfulness with which nature has hidden behind riddles and iridescent uncertainties. Perhaps truth is a woman who has reasons for not letting us see her reasons? ... Oh, those Greeks! They knew how to live. What is required for that is to stop courageously at the surface, the fold, the skin, to adore appearance, to believe in forms, tones, words, in the whole Olympus of appearance! Those Greeks were superficial—out of profundity! (Nietzsche, *The Gay Science* §4)

There is a strong symbolic current that connects the mask, the veil, and the Olympian god-image: “truth,” in the Nietzschean sense, not only “loves” masks, nor only is tolerable when masked, but also *must* be masked, for the preservation of its honour, its profundity. The will to truth “at any price” is impetuous, unrefined, and “indecent.” The mature Greeks knew when to stop asking, and for Nietzsche their superficiality was a virtue. Indeed, the Olympian (in the language of *The Birth of Tragedy*, the Apollonian) dream-image is connected in *The Birth of Tragedy* to the Hindu “veil of māyā?” (35). In that early work, the Apollonian veil masks the Dionysian truth in order to preserve life:

The Greek knew and felt the terror and horror of existence. That he might endure this terror at all, he had to interpose between himself and life the radiant dream-birth of the Olympians. ... It was in order to be able to live that the Greeks had to create these gods from a most profound need. (42)

The Apollonian illusion, in Nietzsche’s later pragmatic sense, could be called a true error: though it was deception and not (positive) truth, it was necessary to protect the Greek from the “horror of existence” which *could not have been borne* without it being falsified. In *The Birth of Tragedy*, this unbearable “horror” is represented by the Dionysian essence, the “primal unity, eternally suffering and contradictory,” in which the *principium individuationis* disintegrates and life is revealed as chaotic, turbulent, essentially barbarous (45). What is this if not the nascent form of the “flux ontology of power-quanta”

that would be developed in Nietzsche's later works? That ontology is presented as the terrible, incontrovertibly "true essence of things," knowledge of which "nauseates" and "kills action" (Nietzsche, *Birth of Tragedy* 60). Indeed, nothing about the primal Dionysian truth Nietzsche offers in *The Birth of Tragedy* seems at all life-affirming; on the contrary, it teaches that "what is best of all is ... not to be born, not to *be*, to be nothing" (42). Here is a strong and singular claim: the terrible, unbearable Dionysian truth, which is the symbol for Nietzsche's perspectival ontology and epistemology, must be masked in order to enable life.

Now the problem of the mask, and its full epistemological-ontological significance, comes to full light. The Dionysian/power ontology that is called true in *The Birth of Tragedy* is life-denying and must be concealed by the Apollonian life-affirming error. It seems we are being offered at last a mask that is false, positively speaking, and a concealed ontology that is true. This usage, though, would undo all our work (and all Nietzsche's) toward a pragmatic criterion of truth and a functional perspectival epistemology: the Apollonian mask, understood this way, is pragmatically true but positively false, and it is given to cover an ontology which is positively true but pragmatically false. This is a final paradox that cannot be resolved if Nietzsche's claims are to hold because, as we have seen, the very concept of a positively true perspectival epistemology is self-negating; if it is true it must be false. If, as *The Birth of Tragedy* seems to give us, it is not life-affirming either, then in *no sense* can the Dionysian, in isolation, be called true: neither positively nor pragmatically. And yet truth is exactly the name it bears. Neither is this the clumsy error of a young, passionate Nietzsche. Although nearly fifteen years separate *The Birth of Tragedy* from *Beyond Good and Evil*, the theme of masking difficult "truths" in order to make life possible persists. Nietzsche suggests in the latter work that

it might be a basic characteristic of existence that those who would know it completely would perish, in which case the strength of a spirit should be measured according to how much of the "truth" one could still barely endure. (239)

Again, the Dionysian "truth," having failed the pragmatic criterion for truth, relies on the positive one which it precludes by construc-

tion. With that, the justification for suggesting an epistemology and ontology that is false in every way but name dissolves as well.

The solution, as Nietzsche warns, lies beyond the faith in opposite values. The insight in *The Birth of Tragedy* that a meaningful existence is possible only in the synthesis of the Apollonian and Dionysian, mask and masked, saves Nietzsche's whole work from collapse. The secret function of the mask image is to remind us of this critical synthesis. Truth, for Nietzsche, is too much for humans to bear, and must be "thinned down, shrouded, sweetened, blunted, falsified," and masked (*Beyond Good and Evil* 239). At the same time, however, it also *loves* masks, and requires superficiality and the will to illusion in order to maintain its dignity, its profundity, its very truthfulness. Both are true at once, because the Dionysian world-view that is Nietzsche's primary hypothesis is *only true when it is masked*.

Nietzsche's hypothesis of perspectival epistemology and power ontology, which corresponds to the Dionysian world-view, is not true in a positive sense. If it were, it would contradict itself. As such it could never be meaningfully *held* as true, except if it were a useful error. It might be true in this pragmatic sense only insofar as it enhances life, but as we have seen, a positive belief in the validity of the Dionysian world-view does *not* enhance life: on the contrary, it paralyses, and perhaps causes those who believe in it to perish. Thus it is *pragmatically*, as well as positively, false. A world organised according to a perspectival epistemology and the ontology of the will to power would be horrid and brutish. Nietzsche knew that such a world would resemble the barbarism of the Dionysian festival, in which "the most savage natural instincts were unleashed, including even that most horrible mixture of sensuality and cruelty" (*Birth of Tragedy* 39). It is obvious that Nietzsche did not intend his vision of the Dionysian world as a moral prescription. The consequences frightened him. Yet he never stopped writing about will-to-power ontology as though it were true. The only possible explanation is this: the Dionysian story becomes true only when the Apollonian veil of illusion, the mask, is introduced, not to conceal, but to be *thrown into relief* by the Dionysian lie. Only then can the lie *become practically useful*, and thus become pragmatically *true*, as part of a synthesis with the Apollonian "errors" that are perhaps not so

erroneous after all.

As we have seen, Nietzsche's Dionysian world-view of strict perspectival epistemology and an ontology of perpetual change, taken in isolation, is wrong in every way. It is logically untenable and pragmatically devastating. It is not meant, I think, to be taken seriously as a system. However, it provides an impetus and analytical ground on which to challenge the so-called "fictions of logic," the synthetic judgements *a priori*, religion, morality, positivism, art, science, and any number of other "errors" that form the foundation for our lives, and thus it is, in a pragmatic sense, "true." Perspectivism is a kind of thought-experiment which asks "What if nothing had any positive value? Would we still choose to believe in morality, or science? Do they cultivate life?" It is the relationship of the perspectivism to prior values, the Dionysian to the Apollonian, and the masked to the mask, that renders the whole pragmatically true. The Dionysian "truth," by presuming freedom from positivism, presents and demands the question of value-for-life. With this insight, we can understand when Nietzsche claims that "we no longer believe that truth remains truth when the veils are withdrawn; we have lived too much to believe this" (*The Gay Science* §4). Perspectivism, as an *attitude* rather than a coherent and earnest epistemology, first implicates all rational propositions on their own terms, with the rigour of positive truth, in order to give us the chance to ask the question: "does this enhance life?" A perspectival epistemology is worthless until it gives occasion to consider whether or not our held values benefit us as individuals and as a species. As an approach, therefore, the systematised Dionysian doctrine is an efficient agent of Nietzsche's general prescription: the revaluation of prior values according to their value-for-life.

This interpretation also helps us make sense of the "honour" of the mask. Perspectivism and the will to power are fantastic straw men. They are meant to cast doubt and not to be believed. But as such they are *fragile*. If some "knight of sorrowful countenance," ready to fight for the cause of truth on earth, was to tilt at this Dionysian windmill and knock it over outright, then he would destroy that noblest falsehood, which had theretofore driven men to self-scrutiny and philosophical maturation (Nietzsche, *Beyond Good and Evil* 226). Thus it is *noble* to respect this last meaning of the mask,

viz., the conceit that perspectivism is meant to be taken seriously. It is *profound* to be superficial and *honourable* to decline to tear the veils off of truth, because to remove the mask and expose the trick would be to remove the allure that gives the Dionysian its constructive power. The mask conceals the falsity of the absurd “flux ontology of power-quantia” so that, since “it is certainly not the least charm of a theory that it is refutable,” it will go on attracting “subtler minds” and causing them to consider the criterion of value-for-life (Nietzsche, *Beyond Good and Evil* 214).

Perspectivism is a buffoon’s philosophy: it refutes itself. In spite of this, however, it draws around itself a “veil of illusion,” demanding a value-for-life reckoning of the putative “errors” that have so far obscured the “horrible truth” of perspectival epistemology from view, and *thus becomes pragmatically valuable*. Nietzsche must have felt that the best way to lead us to criticise our own deeply held assumptions was to suggest that these were not truths at all, but rather masks, invented by those who could not bear to know the truth. The role of the mask is not to protect the individual from the horrible Dionysian truth of perspectivism, but to *legitimise* the Dionysian claims. The mask shrouds the *falseness* of Dionysian “truth” and thus protects its power to cast the truth of those very masks into doubt. Nietzsche’s entire work, ultimately, is bait. When we bite, we do so by asking whether or not, as Nietzsche claimed, our beliefs really do deny life. Nietzsche devoted his career to a *gambit*, whose goal was to use illusion and trickery to help his species to discover the path to a healthier, more vigorous life.

He was never satisfied with the results. As he wrote in his final book, *Ecce Homo*: “From this moment forward all my writings are fish hooks: perhaps I know how to fish as well as anyone?—If nothing was caught, I am not to blame. *There were no fish*” (766). Since that time, more than a few “subtler minds” have taken the bait and, in weighing the Dionysian vision against their own, found that their own prior values could be adjusted to better rise to this challenge about value-for-life. As for those who have ignored the gauntlet altogether, Nietzsche might say they are the ones *most* in need of new philosophies which can cultivate life.

## Works Cited

- Almeida, Rogério Miranda de. *Nietzsche and Paradox*. Trans. Mark S. Roberts. Albany: State U of New York P, 2006. Print.
- Danto, Arthur. "Nietzsche's Perspectivism." *Nietzsche: A Collection of Critical Essays*. Ed. Robert Solomon. Garden City, New York: Anchor Books, 1973. Print.
- Grimm, Rüdiger H. *Nietzsche's Theory of Knowledge*. Berlin: Walter de Gruyter, 1977. Print.
- Nietzsche, Friedrich. *Beyond Good and Evil. Basic Writings of Nietzsche*. Trans. and Ed. Walter Kaufmann. Modern Library ed. New York: Random House, 2000. 199–435. Print.
- . *The Birth of Tragedy. Basic Writings of Nietzsche*. Trans. and Ed. Walter Kaufmann. Modern Library ed. New York: Random House, 2000. 31–144. Print.
- . *Ecce Homo. Basic Writings of Nietzsche*. Trans. and Ed. Walter Kaufmann. Modern Library ed. New York: Random House, 2000. 673–791. Print.
- . *The Gay Science*. Trans. Walter Kaufmann. Vintage Books Edition. New York: Random House, 1974. Print.
- . *On the Genealogy of Morals. Basic Writings of Nietzsche*. Trans. and Ed. Walter Kaufmann. Modern Library ed. New York: Random House, 2000. 451–599. Print.
- . *Seventy-Five Aphorisms from Five Volumes. Basic Writings of Nietzsche*. Trans. and Ed. Walter Kaufmann. Modern Library ed. New York: Random House, 2000. 147–178. Print.
- . *The Will to Power*. Trans. Walter Kaufmann and R. J. Hollingdale. Ed. Walter Kaufmann. New York: Random House, 1967. Print.
- Williams, W.D. "Nietzsche's Masks." *Nietzsche: Image and Thought*. Ed. Malcolm Pasley. Berkeley and Los Angeles: U of California P, 1978. 82–103. Print.



# Science as Regulative Judgments

## A Discussion of Kant's Response to Humean Skepticism

*David Brooke Struck*

IN his *Critique of the Power of Judgment*, Kant discusses teleology in terms of regulative judgments. He claims that teleology is objective, although it consists in mere judgment by analogy to our own causality as agents. By contrast, mechanistic causality is claimed to be "blind," not being causality toward any determinate end (Kant, *CPJ* §61). With respect to the natural sciences, particularly with respect to physics, what does this discussion in Kant reveal? In the second part of his third *Critique*, Kant's system finally addresses the question raised by Hume about the application of causality to objects and the place of natural science. My interpretation of the text is that Kant's project to defeat Hume's skepticism about causality shows promise in some areas, mostly in the framework laid out in the first *Critique*. Nevertheless, I argue that his project as a whole miscarries, and finally, that it fails to establish whether the causality sought by science is metaphysically real or not. In order to elucidate this argument, it will be necessary first to discuss and clarify several terms used by Kant, most importantly: determinative, regulative, reflective and constitutive. The thrust of the argument is that natural science cannot yield a determinative explanation of objects, though it will first be necessary to discuss what that means, as well as to lay out the positive explanation of what natural science does accomplish.

Hume's famous example of billiards is a good illustration of how scientific theorisation operates: we witness the constant conjunction of one ball striking another, and through induction, conclude that the momentum of the first ball is what causes the second to begin

moving at the moment of impact. We are not given the formula for the conservation of momentum beforehand, which would make the judgment determinative, as it would merely be a subsumption of this instance under the law of momentum conservation. Rather, we are given a set of experiences piecemeal that we “subsume under a law that is not yet given” (Kant, *CPJ* §69). This is what Kant calls the reflecting power of judgment.

Here is where Hume’s objection gets serious consideration: we are given the experiences piecemeal, and we experience the motion of the billiard balls, and can even repeat the process through experimentation. However, where do we ever witness the motion of one billiard ball *causing* the motion of the other? We witness their constant conjunction, yet are not purely on that basis justified in claiming one to be the cause of the other (Hume 43).

The power of the first *Critique* is to respond to this argument by saying that the billiard balls are objects of experience, and hence must be subsumed under the concept of causality (Kant, *CPR* A80/B106). We are thus justified in saying that both billiard balls are subject to causality, and this causality is in the object of experience, not merely the subject; their behaviour is governed by a causal chain, as they are both objects of experience *in virtue of* their belonging to a causal chain. However, and this is where the *Critique of Judgment* will clear the muddied waters, how and to what extent are we justified in judging their behaviour to be governed by one causal chain and not another? This distinction between the moment of causality in general and specific causal explanations will serve to illustrate how Kant’s response to Hume is in my opinion clearly not as successful as Kant would have liked.

The concept of causality is constitutive of objects, in that no object could exist as an object of experience without being subject to it (Kant, *CPR* §76). As causality in general is constitutive of objects, we call it determinative; all objects given to us must already be determined by a causal relation, otherwise they could not be given to us. Bare objects of experience are not given prior to their causal relation, but rather in being experienced, it is already implicit beforehand that they are implicated in a causal relation. The same statement would hold about any moment of each of the categories, because it is in virtue of existence formed by the categories that objects of experience

are such. Therefore, judging that an object is causally implicated is a constitutive judgment, and it is thus objectively certain. Up to this point, Kant's refutation of Humean skepticism looks to be showing promise, as he has showed the necessity of causality in experience.

By contrast, teleological judgments about the world are merely regulative judgments about the internal purposiveness of an object by analogy to our own causality, and we cannot apply such a maxim with any objective certainty (Kant, *CPJ* §66). This assertion is quite complex, and requires deconstruction. We witness our own internal causality and understand ourselves as agents; our actions are carried out toward an intended end, and the unity of our actions toward such is grounds for the unity of agency. In order to consider objects in nature as wholes, we must also ascribe an internal end to them, towards which all of their actions can be seen as directed.<sup>1</sup> The ascription of this end is by analogy to our own internal causality. It is through this harmonious direction of action that objects in nature are seen as unified wholes; each part is interpreted as benefiting the whole, and in fact they "are possible [as parts] only through their relation to the whole" (Kant, *CPJ* §65), which in turn exists for the benefit of the parts. This two-way causality of the parts being for the benefit of the whole and vice versa grounds the unity of these objects as wholes (Kant, *CPJ* §64) because we take it that "nothing in such a creature is in vain" (Kant, *CPJ* §66), and so all the actions of the object are unified in their direction toward the end ascribed to that object.

These ends are not given in nature itself, however, and we can therefore only say that the causal chain interpreted as moving towards these ends are the product of regulative judgments (Kant, *CPJ* §63). The pivotal point here is that we need not take objects in nature as wholes; a tree can be explained biologically/teleologically *qua* tree,<sup>2</sup> or by the physical relations of its atoms, or by the chemical

---

<sup>1</sup>An end that is for the benefit of the object itself (e.g.: a tree absorbs sunlight through its leaves in order to allow it to grow, and to strengthen itself) is called internally purposive. Externally purposive ends are ends that are for the benefit of other beings (e.g.: the tomato plant grows tomatoes in order to feed the human). Externally purposive ends are contingent on the benefitted entity being an end of nature itself. Such an end is never given and thus the objective ascription of external purposiveness cannot be demonstrated.

<sup>2</sup>Teleology is, however, the necessary "ground for the cognition of the systematic

relations of its compounds, etc., simply *qua* object. Judgments about natural ends employ the reflective power of judgment rather than the determinative, as they serve as merely one possible way of cognising the object. Unlike the category of causality generally, there is no necessity in cognising objects as such internally purposive natural ends; therefore, the ascription of ends by analogy must take place in the subject rather than those ends being in the object itself. Because teleological ends are not given in nature, but rather only ascribed through analogy to our own internal causality, there is no objective ground for ascribing these ends to the objects themselves. We can still say that the explanation is objective, because it is formed using causality, which is itself an objective feature of our cognitive faculty; however, because the ascription of teleological ends is not necessary in order to cognise the object in nature,<sup>3</sup> the accompanying teleological explanation is only reflective, and never determinative (Kant, *CPJ* §§61, 69). Teleological explanations are but one possible way to describe the actions of the tree, whereas causality in general is not a mere possibility for the object to be presented to us, but is rather necessary, and therefore determinative.

We have thus begun to define the four keywords that Kant uses in his teleological explanation. Let us now make them explicit: constitutive judgments are judgments about the constitution of objects of experience (as formed by the categories); they represent the determinative power of judgment to subsume under given concepts and to give reality to these concepts of the understanding (Kant, *CPJ* §§61, 65, 69). As these concepts of the understanding are formative for experience, these judgments ground necessary objective explanations. Regulative judgments are judgments about the power of research and observation; they represent the regulative power to subsume under principles not yet given (Kant, *CPJ* §§61, 65). As these principles are

---

unity" (Kant, *CPJ* §65), but the necessity of that cognised unity is dependent on internally purposive ends, which are only problematically ascribed.

<sup>3</sup>Although Kant maintains that teleology is necessary in order to cognise these wholes as such: "The above fundamental principle of teleology thus stands, namely, that given the constitution of the human understanding, only intentionally acting causes for the possibility of organic beings in nature can be assumed, and the mere mechanism of nature cannot be adequate at all for the explanation of these products of it—even though nothing is to be decided with regard to the possibility of such things themselves by means of this fundamental principle" (*CPJ* §78).

not the objective grounds of experience (i. e.: are not the categories) but are, rather, “derived from experience” and observation (Kant, *CPJ* §66), these judgments are not necessary objective explanations; they are descriptions. These descriptions rely on causality, however, and are thus formed using objective principles. Still, because they are not constitutive of objects of experience, they have no necessity and remain merely possible. This is in line with Kant’s statement that teleology “cannot rest merely on grounds in experience, but must have as its ground some sort of *a priori* principle” (*CPJ* §66). Regulative judgments can be objectively understood, even though the ground of their application to experience is merely subjective. That is to say, they can be objective descriptions,<sup>4</sup> but given that their necessity is not objectively rooted, they cannot be objective explanations.

Teleological judgments are regulative judgments. Though they are objective descriptions, they do not come from the determinative power of judgment, and can thus be proven *only* as objective descriptions, because they are not necessary for cognition, unlike the categories. All of this is exposed by Kant in his discussions in paragraphs 61–69 of the *Critique of the Power of Judgment*. However, the task now is to determine what relation stands between teleology and the natural sciences, such as physics. Does this so-called “pure” science stand on a maxim, or in some other way lose its objective grounding as an explanation, as does teleology? Or is it rather a constitutive factor of experience, no experience being possible without the rules of science underlying the behaviour of objects? The answer to these questions will ultimately reflect Kant’s success or failure to conclusively refute Hume.

In paragraphs 69–71, Kant distinguishes the teleological from the mechanical conception of nature. Whereas teleology ascribes an end and thus wholeness to nature, mechanical conceptions assume no end to nature at all. In this way, the necessary ascription of an end in order to understand objects of nature as wholes is merely a result “of the peculiar constitution of my cognitive faculties” (Kant, *CPJ* §75), and not indicative of some determinate concept in the object itself. However, as teleological judgments are objectively regulative, they

---

<sup>4</sup>Kant uses the word “elucidation” instead of “description” as the reflective counterpart of “explanation,” which arises from determining concepts (*CPJ* §78).

are still “immanent and secure in their use and appropriate for the human point of view” (Kant, *CPJ* §76). In fact, all principles offered by reason are regulative, as their derivation is based on piecemeal experience. The understanding subsumes experiences under principles given beforehand, which are thus determinative. Furthermore, in order for the understanding to yield a determinative judgment about experience, it must be conscious of the given principle in intuition, and not merely think it, for that would be mere possibility (Kant, *CPJ* §76). By contrast to the teleology, which is regulative as an objective possibility, the causality of mechanical nature “certainly has objective reality” (Kant, *CPJ* §4).

It seems as though Kant is painting a metaphysical picture as follows: Objects in nature can be caused in two different ways, one of which is entirely mechanical, requiring no ascription of ends. The other (the teleological) relies on the maxim of ascribing ends to objects in order to understand them as wholes; however, as the ascription is only problematic, this causality is not determinative of the object, and its objectivity is a reflection of the necessities of our cognitive faculties, and not a reflection of the character of these objects. Therefore, the judgments about teleology are merely regulative, although they are necessitated by our cognitive faculties. Teleology is necessary because mechanism cannot explain the “generation of organized products of nature” (Kant, *CPJ* §71). Mechanism is constitutive; teleology is regulative. It seems here as though Kant’s project to refute Humean skepticism is successful, as he argues that a mechanistic causality must underlie the objects of experience, save for a special class of objects that reflect not a lack of causality, but a different causality that is the product/feature of our cognitive faculty rather than a separate metaphysical causality. However an epistemological issue in science is the pursuit of the rules of the natural mechanism and that is a slippery stepping-stone for Kant on which he must fall down, or else o’er leap.

In §78, Kant discusses the possibility of unifying the mechanistic and teleological conceptions and whether they could be reduced to a single principle. His conclusion is that they could not and while all objects must be subject to the determining principles of mechanism, mechanistic causality is insufficient to explain the cognition of such natural wholes as trees, whose existence relies on the tele-

ological principle. The existence of such objects does not seem to put the determinative nature of the mechanism in doubt, but rather adds to our scientific pursuits a heuristic principle (Kant, *CPJ* §78). The solution to the antinomy raised by Kant in §70 is that teleological principles do not lie in nature, as those of the mechanism do; the thesis and antithesis of the antinomy are (respectively) 1. that all material objects are produced according to merely mechanistic principles, and 2. that some material objects (i. e.: objects in nature taken as wholes) cannot be generated according to merely mechanistic principles. Given that mechanistic principles are constitutive and teleological principles are regulative, the thesis is a statement about objects of experience, whereas the antithesis is a statement about our cognitive faculties and how we can explain our cognition of certain objects as wholes.

This solution suggests that the principles of mechanistic nature are in the objects themselves, and as already discussed, these must be the principles that are given to the understanding beforehand (e. g.: causality in general). But what are the principles of science? Is science satisfied to conclude, albeit determinatively, that everything stands in a causal relation but one that science itself is completely unable to provide? No, the pursuit of science is precisely to determine those rules of the mechanism of nature that constitute our experience. As Kant himself says:

In natural science we are far from being satisfied with an explanation of the products of nature by means of causality in accordance with ends,<sup>5</sup> ... and not according to the things themselves as is appropriate for the determining power of judgment. (*CPJ* §77)

Through natural science, we are striving to ascertain the constitutive causality of our world, and I even believe that Kant argues successfully (mainly in the first *Critique*, and also in §78) that the world of experience is necessarily constituted by an underlying mechanistic causality; however, because of the nature of our cognitive faculties and piecemeal experience, I believe that scientific research and experimentation can yield only regulative judgments about the mech-

---

<sup>5</sup>i. e.: regulatively, through teleology.

anistic causality of objects, an interpretation supported by the text of Kant himself.

The scientific method is to observe, form a hypothesis, and then through a series of experiments and further observations, attempt to find evidence that contradicts the hypothesis. There is no step in the process that concludes the hypothesis to be true. For example, I observe one billiard ball rolling across the table, then the two billiard balls meet, and finally the second begins to move and the first stays at rest. My hypothesis is that some energy in the first ball is transferred to the second upon impact (according to some mathematically defined rule, perhaps) and I therefore go about testing in order to find evidence that might show my hypothesis to be false. So long as I do not come across any contradictory findings, my hypothesis stands as a possibility. It would be false to conclude that as no contradictory findings have arisen, my hypothesis must be an accurate explanation; no evidence supports my supposition that there is energy being transferred from one ball to the other. After all, I have never observed the energy itself, only its supposed effects.<sup>6</sup> To conclude from the repeatability of the supposed effects that the energy must exist and that my hypothesis was correct would be to affirm the consequent.<sup>7</sup> Rather, we can only say that the hypothesis contains no contradiction to the observed phenomena, and stands open as a possibility.

To discuss the same issue in Kantian terms, we are given the experience piecemeal, at which point reason comes up with a principle that links each individual intuition as an example of this universal principle. This universal principle serves as my hypothesis, until another experience contradicts it, at which point the principle given by reason can of course no longer be considered universal, and thus a new principle is required. No amount of non-contradiction is sufficient to show that the principle yielded by reason is a necessary and constitutive one, and our scientific hypotheses must remain forever contingent.

An understanding that is intuitive, one that is given all experience simultaneously, would be able to make judgments in which “there is no contingency in the combination of the parts” (*CPJ* §77). By

---

<sup>6</sup>This was the original objection raised by Hume.

<sup>7</sup>Given “If P, then Q,” and “Q,” one cannot logically conclude “therefore P.”

contrast, our own understanding cannot eliminate such contingency of hypotheses: through

our discursive, image-dependent understanding (*intellectus ectypus*) and the contingency of such a constitution, we are led to that idea (of an *intellectus archetypus*), and that this does not contain any contradiction. (Kant, *CPJ* §77)

In this passage, Kant is talking about the possibility of ends in nature and about how we are unable to experience them but are unable to rule them out through contradiction; however, this discussion of teleological judgment sounds uncannily like a discussion of scientific investigation and the grounds thereof. The implications of such a parallel are immense, on the philosophy of science, on metaphysics and epistemology generally, and particularly on Kant's refutation of Hume.

As the implications are so serious, let us first explore the validity of the claim (based on Kant's own text) that we cannot have determinate knowledge of the mechanism of nature.

The principle which is to make possible the unifiability of both [mechanism and teleology] in the judging of nature in accordance with them must be placed in what lies outside of both (hence outside of the possible empirical representation of nature) but which still contains the ground of both, i. e.: in the supersensible. (Kant, *CPJ* §78)

In this passage, it seems clear that Kant is putting both mechanism and teleology in a comparable epistemological situation: judgments about each must be regulative, as their grounds lie in the supersensible.

Kant distinguishes explanations from elucidations, the former being an application of the determinative power of judgment, the latter being an application of reflective power of judgment. It follows, then, that constitutive judgments are judgments about the metaphysical underpinnings of our experience, whereas regulative judgments are judgments about our ways of cognising the objects of nature: about how we understand them, rather than about the properties in the

objects themselves. He goes on to say that “to explain means to derive from a principle, which one must therefore cognize distinctly and be able to provide” but that if “[the principle] can be indicated but can never be determinately cognized and distinctly provided for use in actual cases, then from such a principle there can be drawn no explanation” (*CPJ* §78). We must therefore have a principle given to us a priori that explains experience. Kant maintains that we are not given such a principle in experience, claiming that

the common principle of the mechanical derivation on the one side, and the teleological on the other is the supersensible, on which we must base nature as phenomenon. But from a theoretical point of view, we cannot form the least affirmative determinate concept of this. (*CPJ* §78)

Here Kant sets mechanical and teleological judgments on a level plane: each is based on the supersensible, and therefore yields no determining concept of nature itself, but rather only a regulating one that we use to understand nature.

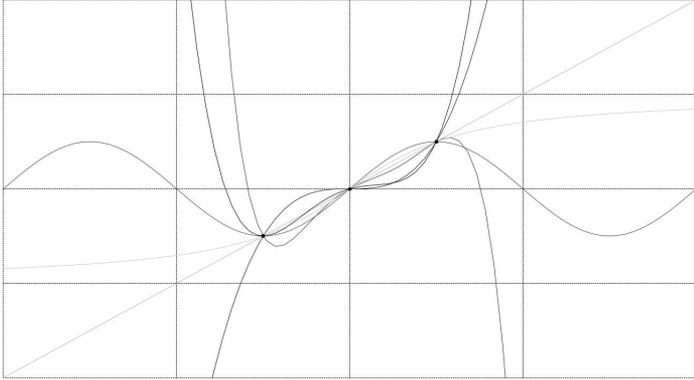
The explanation of experience is derived from causality in general, which is the real metaphysical underpinning of experience. Kant supports this claim in saying that “the principle of mechanism [is that] which always has a claim on any natural being,” and that “the mechanism of natural causes [is] found in nature” (*CPJ* §78). I believe that the case presented by Kant leads us to understand his use of the word “found” in the following way: the mechanism of specific causes *lies in* nature, and *finds itself in* nature, but not that *we find it there*, although we *do* find causality in general. This supports the claim that mechanism is the real cause underlying experience, and not just our cognition of it, as teleology can be.<sup>8</sup> If there were no distinction between the objects of experience and our cognition of them, what would be the difference between regulative and constitutive judgments, as one is supposed to be about objects as they are and the other about objects as we take them to be?

In the *Critique of Pure Reason*, Kant shows causality in general to be an *a priori* constitutive form of experience; it is through such

---

<sup>8</sup>Kant discusses the distinction between real mechanical causes and ideal teleological causes in §65, where the former are constitutive of the objects, and the latter are the matter of regulative judgments. Regulative judgments only constitute our way of cognising certain objects, not of what those objects of experience are.

causality that we cognise the world, as the world is given to us as already causally related. But what those particular relations are we must determine for ourselves through reason, using causality in general as our starting point. In order to elucidate this process, I will present the following analogy, along with a visual depiction.



According to Kant, we are given experiences piecemeal, and the framework of causality in general into which all experiences must fit. Furthermore, he proves that there is a real mechanistic causality that underlies all of our experience. Using the illustration above, the piecemeal experiences correspond to points on the Cartesian plane, specifically the points marked with black dots in this instance.<sup>9</sup> The necessity of causality in general as the form of our experience corresponds to our taking these points to be generated by functions in general. The underlying mechanism of nature corresponds to the function that *in fact* produced these points. However, as the illustration indicates, we can conceive of a multitude (and possibly even an infinite multitude) of functions that encompass the given points. These functions are not equivalent to one another, meaning that they could not each be the true representation of the one objective underlying function. Similarly, we can hypothesise (or at least conceive of) a multitude of causal scientific theories, none of which can be concluded as the determinate causal relation that underlies experience. To reiterate: the mechanical principle is one that is “supersensible, on which we must base nature as phenomenon” (Kant, *CPJ* §78), but as nature is given to us piecemeal, and the true mechanistic principle

<sup>9</sup>The three points: (-1, -1), (0, 0) and (1, 1).

is not given *a priori*, we cannot compare the hypotheses of reason to the metaphysical grounds of causal experience. Similarly, given only the points on the plane, we cannot know conclusively what function produced them.

So what conclusions can we make about the philosophy of science based on Kant's portrayal of the mechanism of nature? First of all, Kant conceives of natural science as a pursuit of the determinations of nature,

far from being satisfied with an explanation ... [based merely on] the generation of nature as is appropriate for our faculty for judging them, i. e.: the power of reflecting judgment, and not according to the things themselves as is appropriate for the determining power of judgment.  
(*CPJ* §77)

Given this conception, Kant must accept that natural science ultimately fails to reach its goal. We cannot know the underlying causal mechanism of nature, but rather are only able to construct models that map the tendencies we experience in nature.

However, the regulative judgments that we make about experience are not to be overlooked, either. Kant's conception of science dooms his pursuit to fail to meet its ends, but other conceptions of science are possible. The regulative judgments that science ultimately makes are useful technical judgments. If one can describe how something operates, one can understand it objectively, even if one cannot explain it.<sup>10</sup> Given an objective understanding of nature, even a merely descriptive one, one can develop techniques and technology that help one to master nature, and achieve one's intended ends. Natural science in this light would be a technology and not an epistemology. This is indeed what Kant maintains; on his view, the distinction is analogous to that between the Greek *techne* and *episteme*—where the former is a study of perfected craft (in the English sense) or art (in the German sense) and the latter is a study of truth (in a deep, metaphysically correspondent sense).

---

<sup>10</sup>That is to say, if one can map out the behaviour of an object, and then derive a function that describes this behaviour, then one achieves an objective understanding of the object, even though it cannot be confirmed that the derived function in fact is responsible for the behaviour witnessed.

The metaphysical foundations that Kant establishes in the first *Critique* are not affected by this clarification of the scientific project as technical. Causality in general is still a constitutive ground of experience; but after this discussion, it seems clear that specific principles of causality are not necessary for experience or constitutive of it. Rather, specific principles of causality are regulative judgments that are contingent elucidations of our experience. The necessary and constitutive mechanistic cause of nature is metaphysically supported, but entirely inaccessible to us, as we are subjects given piecemeal experience. The discursive nature of our understanding precludes the possibility of our access to these objective metaphysical foundations.

While our access to the supersensible metaphysical foundations of experience are severed by this discussion, there is still an interesting epistemological discussion to be had regarding our objective regulative judgments about natural causality. Looking back to the analogy of the Cartesian plane, if we say that our judgments are merely regulative rather than constitutive, we are not forced to choose one mapping function as the true one and negate the truth of the others (as non-identical, and hence contradictory functions). Rather, we can say that each function is a correct description of the points that we are given. In terms of experience, all of our scientific hypotheses are said to be true, insofar as they correctly describe the experiences we are given piecemeal. This shift toward regulative judgments represents a massive change of truth conditions from what Kant laid out in his epistemological description of the goal of natural science. Given that experience is piecemeal, can we ever say that a hypothesis is true? As a constitutive principle, we could never know. As a regulative principle, we can say that certain hypotheses have been true *up to now*, the possibility of disproof always remaining open.

Furthermore, as we are given more and more experience, our discursive understanding refines our judgments to a closer and closer mapping of the actual mechanical causality that truly underlies the causality we seek in nature. This refinement of our regulative principle means that we are more likely to correctly predict future behaviour of objects according to causal mechanical laws. Correctly predicting behaviour, and acting according to these predictions is the

technological power of science. Kant claims that “[p]robabilites count for nothing here, where judgments of pure reason are at stake” (*CPJ* §75). When I first came across this statement, I wondered whether it would be problematic for science given the piecemeal nature of our experience, and the discursive nature of our understanding. In fact, given Kant’s epistemological notion of science, it is. The statement denouncing the determinative validity of probabilities only serves to reinforce the fact that the notion of science as constitutive judgments is untenable. However, it does not at all affect the technical notion of science as regulative: there is no pretension that the derived principles will be determinative, and the usefulness of science thus conceived is not diminished by its reliance on probability and ever-increasing accuracy of approximation.

How does this discussion reflect on Kant’s response to Hume? Hume argues that causality is never experienced in the object itself, but is rather an addition made by the imagination after experiences of constant conjunction (43). One of Kant’s main projects in his *Critiques* is to show that causality is a metaphysical fact about objects of experience and not merely a subjective notion. However, in looking at the conclusions that Kant draws, he has only succeeded partially. It seems that his argument is successful regarding the metaphysical reality of a mechanism of nature. Also, as he shows in the *Critique of Pure Reason*, causality is a constitutive moment of the categories of experience. The success ends here, though, as Kant is forced in the *Critique of the Power of Judgment* to relinquish our access to any causality underlying the objects of experience themselves. He must conclude that we can make only regulative and not constitutive judgments about the causality of objects, and his rehabilitation of a metaphysically-grounded science ultimately miscarries.

I would like to argue that Hume’s own positive project is in fact not far off from the final conclusions that Kant reaches about causality. Hume’s definition has two parts: the first is that causality is “an object, followed by another, and where all objects similar to the first are followed by objects similar to the second” (76). This definition has to do with the piecemeal nature of our experience, and that the objects themselves follow one another in a temporal sequence. It is from this temporal sequence that the notion of their causal connection arises in us. But more importantly, their constant conjunction,

whether due to a causal connection or merely an ordered series of accidents, is still a piece of knowledge that we can use to operate within our environment.

The second definition advanced by Hume, and it is in the combination of these two that he sees the full definition of causality, is that of “an object followed by another, and whose appearance always conveys the thought to that other” (77). This is roughly equivalent to the definition of regulative scientific judgments that Kant is forced to concede, whereby causality becomes merely a way for us, as subjects, to describe and operate within our environment, although the causality itself that we derive from piecemeal experience is not in the objects of experience themselves. The first half of the definition deals with the fact of conjunction, whereas the second half deals with our judgments about this conjunction as causal, and the feeling of necessity that those judgments have for us.

This feeling of necessity really elucidates the distinctions between Hume’s positive project and the technical notion of science that must be derived from the *Critiques*. Hume believes that we are given the constant conjunction and then by some act of the imagination we assume that there is cause involved. As we never experience the cause, we are not justified in concluding that causality is the metaphysical underpinning of experience, and not merely our habitual way of understanding the world. By contrast, Kant argues that causality in general is a foundational building block of experience, and that we cannot have experience without it; it is thus necessary. He also argues that there is a metaphysically real causal mechanism that underlies nature, although we cannot have knowledge of it. Therefore, the causal chains we derive from experience are merely regulative.

Hume argues that we are given individual experiences and the imagination derives causality to describe these experiences, but that such a derivation does not prove that causality is a metaphysical reality. Referring back to my analogy of the Cartesian plane, this is to say that we are given points on the plane, and the imagination derives a comprehensive function to describe these points, but such a derivation does not prove that the points are originally given by a function. By contrast, Kant says that we are given objects of experience and that in their being given it is implicit that they are

causally implicated, but that in making regulative judgments in natural science, we cannot be sure whether our regulative judgments correspond to the determining mechanism of nature. Analogously, we are given the points on the plane, and in their being given it is implicit that they belong to certain functions, but in deriving the functions from these points we cannot be sure that the derivation maps the true function that produces these points. Kant and Hume disagree about whether causality is a constitutive form of experience, but they both agree that the particular causal judgments we make are mere elucidations of nature, not explanations of it. But even in being elucidations, they are still useful technical tools.

## Works Cited

- Hume, David. *Enquiry Concerning Human Understanding. Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Ed. L.A. Selby-Bigge. Oxford: Clarendon, 1975. Print.
- Kant, Immanuel. *Critique of the Power of Judgment*. Trans. Paul Guyer and Eric Matthews. Cambridge, UK: Cambridge UP, 2001. Print.
- . *Critique of Pure Reason*. Trans. Paul Guyer and Allen W. Wood. New York: Cambridge UP, 1998. Print.