Veritas

VARIA

Conversion of Consciousness as Principle of Morality

A Conversão da Consciência como princípio da moralidade

* Konrad Utz

Abstract: Kant shows that a fundamental theory of normativity and morality can give neither an explanation nor an explication of normativity, but can only articulate and render explicit its origin. It can do so by indicating the place or *topos* and the turn or trope of its originating. According to Kant, the *topos* of normativity is the will *qua* practical reason and its trope is the general, typically instrumental use of this reason, i.e. reflection. The trope of the origin of morality is autonomy, i.e. practical reason turning on itself and thus becoming pure practical reason, thereby establishing its own form as law: the categorical imperative. Consequently, fundamental ethics serves two functions (if it is indeed successful): formally, it provides evidence for the originality or authenticity of morality; materially, it gives a criteriological principle for the content of morality. The article argues that Kant was right in his view of the foundation of ethics, but was wrong in the manner in which he met the requirements thus established. The *topos* of normativity and, consequently, of morality cannot be reason, but must be consciousness or, more precisely, (actual) knowledge *de se*; and its fundamental trope cannot be reflection (and then autonomy), but must be what can be described as a *conversion of consciousness*. This conversion can be identified with *philein* in the sense of Aristotle. It has four different aspects: desire, cognition, benevolence, and recognition. When *philein* is reciprocal and lived, *philia* ensues. Friendship (in the broad sense of Aristotle) is then described as the *topos* of the origin of Normativity and Morality.

Keywords: Aristotle. Friendship. Kant. Normativity. Morality.

Resumo: Kant mostra que uma teoria fundamental da normatividade e da moralidade não pode dar nem uma explanação nem uma prova da normatividade, mas apenas pode articular e explicitar sua origem.

_____
* Doutor, Professor da Universidade Federal do Ceará, Fortaleza. <konrad.utz@gmx.net>.

Ela pode fazer isso indicando o lugar ou o topos e a virada ou a trope de seu originar. Conforme Kant, o topos da normatividade é a vontade enquanto razão prática e sua trope é o uso geral desta razão que tipicamente é instrumental, no sentido da reflexão. A trope da origem da moralidade é a autonomia, i.e., a virada da razão prática sobre si mesma, tornando-se pura neste ato. Nisso, a razão prática estabelece sua própria forma como lei para si mesma, na forma do imperativo categórico. Em consequência disso, a ética fundamental serve duas funções (quando for bem sucedida): formalmente, ela fornece evidência da originalidade e autenticidade da moralidade; materialmente, ela fornece um princípio criteriológico para o conteúdo da moralidade. O artigo argumenta que Kant estava certo em sua visão da fundamentação da ética, mas estava errado com relação à maneira como ele tentou cumprir as exigências estabelecidas. O topos da normatividade e, em consequência disso, da moralidade, não pode ser a razão, mas precisa ser a consciência ou, mais exatamente, o saber de se (atual); e sua trope fundamental não pode ser reflexão e, depois, autonomia, mas precisa ser o que pode ser descrito como conversão da consciência. Essa conversão pode ser identificada com o philein no sentido de Aristóteles. Este "amar amigável" tem quatro aspectos diferentes: desejo, cognição, benevolência e reconhecimento. Quando este philein for recíproco e estiver continuamente vivido, nasce a philia, a amizade (no sentido amplo de Aristóteles). Essa é descrita, consequentemente, como o topos da origem de normatividade e moralidade.

Palavras-chave: Aristóteles, amizade, Kant, normatividade, moralidade

Kant on the Foundation of Normativity

G.E. Moore has argued that normative sentences cannot be reduced to descriptive sentences,[1] and despite much debate, this has been widely accepted by the philosophic academic community. In contrast, Immanuel Kant has articulated the even deeper insight that consciousness[2] of being obliged cannot be reduced to consciousness of sentences (at least not *simpliciter*), no matter if they are normative or

---

[1] To be exact, this is how Moore's "naturalistic fallacy" is commonly understood. Moore himself did not discuss sentences, but the meaning of the word "good", cf. idem, *Principia Etica* (Cambridge: Cambridge University Press, 1903). But this is not important for what I want to explain.

[2] I use the term consciousness in a broad sense, roughly as that, which we have when we are awake or dreaming and which we do not have when being in deep sleep or in coma. In Kant, consciousness in this sense is the basic feature of "Vorstellungen" (representations), as in: "The representation of an objective principle, inasmuch as it is compelling for some will, is called *commandment* (of reason) …" *Grundlegung zur Metaphysik der Sitten* (GMS – *Groundwork of the Metaphysics of Morals*; cited according to the Akademie Ausgabe, Berlin 1900ff), 413. For a more detailed account of my views on consciousness cf. K. Utz, *Bewusstsein. Eine philosophische Theorie*. Paderborn: Schöningh, 2015.

descriptive.[3] The problem is that a sentence *per se* can inform a subject only about *something*, that is, about things, including other subjects and attitudes (whether propositional or not) of other subjects towards something. Therefore, any normative sentence can inform a subject only of the fact that there is an ought or that somebody means that somebody ought … . A subject may even learn that somebody means that *she* ought. But that does not *by itself* imply that this subject *herself* consequently means that she ought, that she is *conscious* of being obliged (or that she has not only an objective, but also and more importantly subjective ought-consciousness). This is evident from the fact that contradicting claims can be understood by a subject at (more or less) the same time.

Consciousness of being obliged implies knowledge *de se*. If I am conscious of being obliged, I am conscious *of myself* being obliged. I have practical self-consciousness; I know *myself* under an obligation (I have "subjectual" knowledge *de me*, different from "objectual"[4] knowledge *de me*). Knowledge *de se* cannot be reduced to other kinds of knowledge (*de re*, *de dicto*), as Hector-Neri Castañeda[5] and others have argued. M. Frank has recently presented the discussion in a way that seems conclusive to me.[6] Nothing can convey knowledge *de me* to me; it is not information that can be passed on to me – or that I could pass on to someone. Therefore, a sentence can only be an *articulation* of my self-knowledge: I can never arrive at self-knowledge by acknowledging a sentence and then, by means of this, get knowledge *de me*.[7]

---

[3] Cf. GMS. The following exposition is based mainly on the *Second Section*. Cf. also Kritik der praktischen Vernunft, 64 (KPV – *Critique of Practical Reason,* also cited according to the Akademie Ausgabe).

[4] It is necessary to talk in this way, because knowledge *de se* cannot be object-knowledge in the intentional sense – if it was, a problem of identification would arise.

[5] Cf., e.g., idem. He. A study in the logic of self-consciousness. In: Andrew Brook, Richard C. DeVidi, ed., *Self-reference and Self-awareness*, Amsterdam, Philadelphia: John Bejamin's Publishing Company, 2001, 51-80.

[6] Manfred Frank, *Ansichten der Subjektivität*, Frankfurt a.M.: Suhrkamp, 2012, esp. 359-368, see also 99-126.

[7] Of course, I can receive or accept information „from outside" into my knowledge de me. I can identify myself in the mirror; and I can identify myself as the one someone else is talking about, especially when she is directing her attention unto me and uses the word "you". But this presupposes that I know how mirrors and how the language in question function; and it presupposes that I already have knowledge de me beforehand – information received from outside can only inform me about myself if I, myself, link it to knowledge de me which I already have. I cannot learn by information I receive that I am myself. And I can only receive information as about myself if I already have some contentful knowledge de me to which I can link this information – and not just abstract, transcendental self-consciousness. I must know, e.g., that I have a body to identify myself in the mirror (though I may be mistaken, of course, as to the concrete specifications of my body). I must know (at least implicitly) that I am listening, to understand that by using the word "you" the other person is referring to me as the (intended) listener. And if I am to understand that, e.g., "everybody still!" should include myself, I must *beforehand* (at least implicitly) know myself to be somebody: to be an individual that may fall under the extension of "everybody".

The primary form of practical self-knowledge is, of course, volition or volitive consciousness. Prescriptive sentences can only have normative meaning in the subjective sense for a subject directly if they articulate a volition of the same subject. If I want the sun to shine, I mean that the sun should shine, and I myself enforce this ought. If I want to go to the theater, I mean that it should happen that I go to the theater (but not that I *should* go to the theater in the sense that I have the obligation to do so – *it* should happen, not *I* ought to go).

Consequently, all other forms of (subjective) normative consciousness must be derivative of volition. All ought-consciousness must formally be a want-consciousness turned against itself because volition is the original mode of *understanding* normative sentences. I cannot understand normativity if I cannot understand myself "in a practical mode", that is, as willing. So there must be some mediation, some turning or some "trope" by way of which my own volition is turned "against me". This trope is typically the expectation of sanctions. We accept prescription as something we (subjectively) ought to do, because we expect to be rewarded if we do as it prescribes and/or to be punished if we do not do so. It does not matter who will effect these consequences nor by which causality: whether it is another subject (a human being, a community, a god etc.) by means of her will and powers or whether it is nature by means of natural causality. Of course, "turning" my own volition in this way "against me" implies the use of practical reason. I conclude that I should do *this* because I want *that*, understanding (the premise that) I only can obtain *that* if I do *this*.[8]

That is, my own practical reason tells me to do something I *do not want* to do (e.g. renounce sweet food and drink) in order to obtain something I *want* (e.g. to be slim). The resulting consciousness is ought-consciousness precisely because it prescribes me to do what I *do not* want. Consequently, noncompliance is possible: I cannot, e.g., resist the cake offered to me despite knowing that I *should not* eat it. It is obvious to Kant that all external prescription can be turned into ought-consciousness *only* in this way, i.e. by instrumental practical reasoning.[9] Ought-consciousness can only arise *within* volitional consciousness by

---

[8] Cf. GMS 444: "*I should do something because a want something else*".

[9] This is obvious from the fact that, according to Kant, there are only three forms of imperatives (Kant evidently uses this term not in the grammatical sense, but in the phenomenal: as that which a subject, within her consciousness, accepts a practically binding. Of these three, the only the first two present prescription which do not origin from the subject herself (from her pure reason). However, these prescriptions (the first premises in the practical syllogism) do not bind by themselves, but only under the condition of the second premise which is subjective and articulates an immediate volition or an "inclination", in Kant's terminology. Cf. GMS 413-419, 444; KPV 20.

turning a subject's volition against itself. To put it differently, volitional consciousness is the *topos* of ought-consciousness;[10] and the reflection of volition "upon itself" through practical reason is the *trope* of ought-consciousness. As such, a fundamental theory of normativity can have no other function than that of articulating or rendering explicit the *topos* and trope of volitional consciousness, i.e., of the will (in rational beings). There is no *explanation* of ought-consciousness – because it is essentially a form of knowledge *de se*, and such knowledge cannot be reduced to anything else (i.e. to some *explanans*, no matter what specific character this reduction may have); and there is no *explication* of ought-consciousness in the sense of conceptually constructing it out of something different than volition and its self-reflection, for instance, by analyzing parts that constitute it or by systematizing it into a larger whole which determines it (language games, social practices, e.g.) – for knowledge *de se* can only be self-evident inasmuch as nothing but itself can inform it of being of the subject in question.[11]

Kant on the Foundation of Morality

Therefore, any fundamental *moral* theory can only be an articulation or rendering explicit of some special form of the *topos* of volition and its trope – of some special way of volition *itself* articulating itself in consciousness.[12]

This is exactly what Kant's moral theory is doing. My contention is that most interpreters have failed to understand the most crucial aspect of his theory because they expected him (or moral theory in general) to do more – and some because they expected him to do less. But one just cannot do more (or less) if one wants to develop a fundamental moral theory.[13] Any theory that does more, that tries to explain or explicate moral consciousness, will destroy moral consciousness because if morality is reduced to something else or if it is (re-)constructed from something else (be it mereologically or holistically) then it ceases to exist.

---

[10] CF. Allen W. Wood, Kantian Ethics, New York 2008, 116: „… the nature of rational will turns out to be the location" „where we should expect to find the sole source of normativity".

[11] Cf. K. Utz, Quid mihi? Zur Methode der Grundlegung der Ethik bei Kant. In: *Deutsche Zeitschrift für Philosophie* 64/2 (2016), p. 218-223.

[12] Cf. K. Utz, loc. cit., p. 213-227, 217-218.

[13] A fundamental moral theory can be described as a normative theory which aims at basing morality as such on a principle (or principles) which it justifies by some form of rational argument and which allows to develop moral content or at least to judge the content of existing moral systems; or in simpler terms: a theory which argues why morality is, in fact, valid and which expounds some way of determining what it obliges to.

Kant's rendering explicit of morality starts by claiming that there is another possible way or trope of a subject's volition turning against itself by practical reason than by *instrumental* reasoning. Instrumental reasoning starts from some concrete, primary volition or from some "inclination" as Kant would call it. It then searches for some law-like nexus (of natural causality or social sanction) by means of which it can realize this volition. That is, practical reason takes a *representation of some law(s)* and *prescribes to act* under that law – since the minor premise, the volition for the goal, is given. Kant reserves the term "will" for this *capacity to act under the representation of laws*, which he identifies with "practical reason" (GMS 412).[14]

Yet, according to Kant there is another way in which practical reason can prescribe to act: it can be non-instrumental. Practical reason *turns* non-instrumental if it does not accept input from outside, i.e. if it does not operate under a premise given by inclination.[15] But by what means can such a turning point take hold? What could function as the trope of non-instrumental volition? According to Kant, it can only be the turning of practical reason to itself.[16] But what does practical reason find if it turns to itself, searching for resources to make itself into non-instrumental volition? Well, it finds its own form: the form to "act under the representation of some law". Practical reason thereby "imposes its own form on itself"; it establishes its own form as a law for itself, a law under which to act. And this means that the aspect by which practical reason was instrumental, i.e. the subjective or motivational principle (borrowed from inclination) should take the form of a *representation of a law*, that is, it should be such that it could be a norm which is universally and necessarily binding (for a *representation of a law* is something which either *is* or *could become* a law).[17] This, thus, is autonomy for Kant: practical reason establishing itself as a law – and since for Kant the only possible normative prescription for myself, as shown above, is practical reason, and since practical reason by itself represents only the *form* to act under the representation of (some) law(s), autonomy is nothing but

---

[14] To be precise, Kant in GMS 412 uses the definite article "… under the representation of the laws", because he talked about laws immediately before. However, the parallel GMS 427 ("… some/certain laws") makes it clear that Kant does not speak about *the* law of morality or any other specific law here, but about laws in general. Cf. also KPV 32. For exegetical discussion of the passage see Pierre Laberge, 'La définition de la volonté comme faculté d'agir selon la représentation des lois (GMS: 412)', in O. Höffe, *Kants Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar*, 4th edn (Frankfurt a.M.: Klostermann, 2010), 83-96. I basically agree with his exegetical conclusion (90f).

[15] Cf. GMS 419-421.

[16] Cf. GMS 444.

[17] Cf. KPV 31: "The will is being thought as independent of empirical conditions, that is as pure will, if it is thought as determined by the pure form of the law …"

practical reason's establishing its own form as a law for itself.[18] The outcome of this is articulated by the categorical imperative: I should act in such a way "that I a can also want that my maxim [the subjective principle of my will] should become a universal law".[19] And this is the principle of what we call "morality". To summarize, according to Kant the *topos* of the origin of morality is volition in the form of practical reason; and the trope from which it originates is autonomy in the literal sense of the word in Greek.[20]

Note that this does not explain why we should be moral.[21] Such an *explanation* would lead into an infinite regress of oughts, since it would invite the question why we *ought* to accept the duty to be moral – and so forth. It should also be highlighted that this also is not an *explication* of how moral sentences are constructed in our mind (or, even worse, in language). Such mental content would never, in itself, constitute obligation, since, as mere sentences, it still would have to be connected with knowledge *de me* which cannot be imparted by sentences (see above). I could say: "O look, there is sentential mental content originating in my mind which demands me to obey it unconditionally. That is very interesting! Kant explains me, how this happens. That is more interesting still! – But what has this got to do with me?! After all, there is a lot coming up and going on in my mind. So let us just go on with that sweet immoral life I enjoy so much!" A fundamental moral theory can do no more then to articulate the origin or the originating of moral obligation which *in fact* is happening. It is a rendering explicit of a fact or of the originating of a fact. Kant calls this the fact of reason.[22] *We know ourselves as morally obliged*. No moral theory can go beyond that fact.

But that does not mean that a fundamental moral theory has no function, that it only informs us what of what we already know. Quite to the contrary: in the first place, the rendering explicit of the originating of moral consciousness provides evidence for the fact of this moral consciousness

---

[18] Cf., e.g., GMS 444: "The absolutely good will … will therefore … contain only the *form of willing* as such, and it will contain it, namely, as autonomy". I have explained this interpretation of Kant in detail in: K. Utz, Quid mihi? Zur Methode der Grundlegung der Ethik bei Kant. *Deutsche Zeitschrift für Philosophie* 64/2, p. 213-227, 2016.

[19] I cite the simplest version of the categorical imperative, GMS 402.

[20] Cf. KPV 28f; 33: "The autonomy of the will is the sole principle of all moral laws." I have argued for the exegetical correctness of this interpretation in: Utz, Konrad. Praktische Vernunft in der „Grundlegung zur Metaphysik der Sitten". In: *Zeitschrift für philosophische Forschung*, v. 69/4 (2015), p. 474-501.

[21] Cf. GMS 463.

[22] Cf. Critique of Practical Reason, 31. I cannot enter here into the vast discussion on the interpretation of this topic, which is one of the most controversial among scholars of Kant. For an interpretation that is close to mine cf. Klaus Steigleder, *Kants Moralphilosophie: Die Selbstbezüglichkeit reiner praktischer Vernunft*, Stuttgart, Weimar: Metzler, 2002, 102-108.

as an *original* fact. That is, it establishes my knowledge *de me* of being non-instrumentally or of being "simply" and thus "unconditionally" obliged as *authentic* knowledge (well, to be exact: as at least *possibly* authentic knowledge – but that suffices totally[23]). For it could be, and many people, beginning with ancient sophists, in fact held that moral consciousness is not *authentic* but *really* an instrumental ought the sanctional premise of which has been forgotten or has been established so firmly by education or cultural conditioning that it subsided into the subconscious. There is, in each of us, a natural tendency (which stems from inclination) to *doubt* the validity, the purity and rigor of morality, i.e. *to doubt its authenticity*.[24] Therefore, it is of utmost importance to provide evidence that authentic non-instrumental ought-consciousness is possible; and such evidence represents a real *foundation* of morality – not because it bases morality on something else, but precisely because it shows how it *can* arise based on nothing but "itself", i.e. on nothing but its own *topos* (in Kant: the will as practical reason) and its own trope (in Kant: autonomy). In the second place, such a fundamental theory may lead us to a formal, criteriological principle of the content of moral norms (in Kant: the categorical imperative).[25]

Philein

In my view Kant is right in establishing what a fundamental moral theory is viz. what it can accomplish. But he is in a crucial manner wrong in the way he renders explicit the *topos* and the trope of moral obligation. This is where friendship comes in.

The problem is that from Descartes to Kant and even beyond philosophers have not clearly distinguished between consciousness and thought (or reason). It is indeed correct that practical reasoning prescribes to do something else than volition (or, in Kant's terminology: than inclination) originally wanted. But it is not correct that this prescription

---

[23] Kant is not very clear about the question whether (apparent) moral consciousness can be non-authentic (there are passages which indicates that he accepts this, others that he does not). However, it seems obvious that this can be the case. But if, in case of reasonable doubt about the authenticity of my apparent moral consciousness, I can, by means of moral *reflection*, establish that this *would* be what my moral consciousness would really say if it *was* authentic, then I have *no reason* not to obey it (by the way, this is one *pragmatic* function of moral reasoning and, consequently, of moral theory), i.e. no reason to expel it from my consciousness *de me*.

[24] Cf. GMS 404f.

[25] Kant formulates the two functions of foundation of morality, viz. articulating its origin (and thus its formal aspect, its authenticity) and rendering explicit a criteriological principle for its content, in GMS 405: practical philosophy informs common moral consciousness about the "fountain of its principle and the right determination of the same".

always enters consciousness as an ought. In fact, it often does not enter consciousness as anything distinct at all – but as just the thing I want to do. If I drink wine, part of that action is obviously instrumental: I lift the glass to my lips *in order to* drink from it. I do not *want* to lift the glass to my lips *simpliciter*. But it does not even come to my mind that, in lifting the glass to my lips, I am doing something different from what I (directly) want. In my (conscious) mind, lifting the glass is just part of the action that I want and enjoy. Even in cases where there is a distinct consciousness of realizing the means, very often this is not an ought-consciousness. Take a girl that is spotting her friend outside the window. She wants to go to her. But to do so, she has to do exactly the opposite: she has to go away from her since the door of her room is opposite to the window. But the girl is not conscious of doing something she *should* do – but does not want to do – *in order* to get to her friend, of having to do something *different* than she actually wants to do. No, she is conscious of running to the door as *part of* the action of running to her friend. The situation changes when she is far away from her friend and the journey to see her is strenuous. In that case she actually will regard the journey as something she *has* to undertake, something *imposed* upon her by her wish to see her friend. Consequently, there certainly *is* a difference between having knowledge *de me* or self-consciousness as wanting to do something and having knowledge *de me* as being obliged which *does* rest on the *difference of ends and means*. But it is not *reason* that makes the difference because in many cases practical reason is active in prescribing means, but this prescription does not enter consciousness *as an ought*, i.e. as a prescription *turned against* the primary volition. It just amalgamates with that volition. To draw some of the consequences of the above discussion, the first point is that the *topos* of the origin of ought-consciousness is really practical knowledge *de se* or volitive consciousness – and not practical reason. Practical reason *does* make a difference and that difference *may* enter consciousness as an obligation turned against primary volition. But it *need not* do so. Therefore, there must be something else that *makes* the difference. It is important to remember that normativity is not just a different content that emerges, but a different way of being conscious of some content: as something I *should* do. This latter difference in the *way of being conscious* cannot have its origins in reason, since it does not necessarily emerge with the use of reason. But it is properly speaking the difference which is decisive.

Once this is clear, it is immediately evident that even in those cases where the prescription of practical reason *does* turn against the primary volition as an ought, this turn is but the second turn. It is the *reflective* turning *back* on myself in my practical knowledge *de me*. Since I do not

always and necessarily have ought-consciousness, but very often have want-consciousness instead, this reflective turning "back on myself" is as such a mere possibility – not a necessity. Consequently, we have to ask for its condition of possibility. That condition evidently consists in me *turning away* from my primary volition. This turning-away may have various causes, the most typical being that somehow, I am hindered to do what I want, or what I do does not have the results I want, or I do not know how to realize what I want. In these cases, I ask myself "what should I do?". I am asking for advice from practical reason (be it my own reasoning capacity or someone else's). I am asking for an instrumental prescription to realize my primary volition. This turning-away from my primary volition, i.e. its suspensio, is the condition of the possibility of the second turn, the prescriptive turn of practical reason against primary volition. But first, volition must somehow "by itself" or "out of itself" turn to something different from itself – in this case to practical reason. "By itself" does not mean that this is its free decision. But it does mean that it is volition *itself* which turns, thereby taking a different direction than it was originally headed, but which thus continues to be itself in the very act of turning to something else. If this entire process was not enacted by the primary volition *itself*, this volition could not serve as the bases of the instrumental prescription (as a premise in the syllogism of practical reason). But then this prescription could not bind me. If the primary volition (in view of an obstacle, a failure, etc.) just *stopped* instead of turning, the result would be frustration – but not an ought-consciousness. So the turning away, the turning into a new direction, must really be an act of volition itself – even though it need not be (and cannot be, at this level) a deliberate or free action.

We therefore have two tasks to accomplish in order to reestablish a fundamental ethics after Kant's failure: we have to go back to consciousness as the *topos* of the origin of moral obligation; and we have to search for the basic trope which renders explicit the originating of this obligation. For even though reason sometimes *is* the addressee of this turning, it is not so *necessarily*. Consequently, this turn cannot be *defined* as a "turn unto reason".

If, for the moment, we cannot be sure what or who is the addressee of this fundamental trope, we should look at the other side. From what is this trope turning me *away*? It is turning me away from the actuality of my primary volition, i.e. from my practical knowledge *de me*. It is "me turning away from myself". Of course, this cannot just mean turning to some (other, new) object; for if I am turning to some object, I am turning *myself* towards it – either cognitively or volitionally. No object as such can make me turn away from me, for objects are by definition given,

and if they enter my consciousness they are given *for me*. One object may make me turn away from another object, but not from myself, from my centeredness on myself. It will not change the centeredness of my consciousness on *my self*.

This centeredness can only be "turned" away form *itself* if it turns to *another self*. And this is what, according to Aristotle, happens in *philein*, in the act of loving.[26] The immediate implication of this trope is that my volition is *oriented onto* this Other Self – not only in the sense of an object, but in the sense that the volition itself is turned around: If I love a friend, I do not only want *for me*, I also want *for her*. I wish the good not only for me, but also for her. And I wish it for her not instrumentally, but *for her own sake*.[27] My volition in fact turns its orientation onto the Other Self[28], the self that I love. I want *for* this Other Self as a subject, subjecting *myself*, my volitive consciousness, my practical knowing *de me* to her, i.e. to her well-being. This is *benevolence*.[29]

Once we understand that benevolence originates from the trope of volition turning to an Other Self, we immediately see that it is originally quantified and qualified. For the specification of *what* good I want for the other person can only be the specification of my own volition which is turned unto her. But this specification of the first or basic volition *in view of* the other person I love is just the good I desire for myself from her (according to Aristotle, this good can be: 1. the other person's goodness *in itself*, her excellence or *virtuosity*; 2. her doing good *to me*, i.e. her in some way having some positive *effect* on me, i.e. her being *pleasurable*

---

[26] Cf. NE 1155b30f. Cf. Solokowski, Robert, Phénoménologie de l'amitié. In: J.-C. Merle, B.N. Schumacher, ed. L'amitié, Paris 115-135. – In some way, Kantian reason is another self within myself – or there are two Selfs within myself, two ego-centers: sensibility (including inclination) and reason. Of course, this does not work. But it explains why Kant's theory seems plausible at first sight.

[27] Cf. NE ibid. Cf, John Cooper, Aristotle on the Forms of Friendship. In: *Review of Metaphysics* 30 (1976), 619-648.

[28] Cf. NE 1169b7.

[29] The originating of ought-consciousness out of the use of practical instrumental reasoning as described above is then explained as a case of self-benevolence or self-care: instrumental prescriptions of practical reason enter my consciousness *as* oughts exactly to the degree in which I enter into a "distance to myself" (typically over time) in which e.g. the "myself" who will enjoy the fruits of my action isn't immediately "I myself" at the moment I have to take up the strenuous work. So when I am saving money for my future pension I (typically) cannot identify fully and directly with the "myself" I am working for – different from the "myself" to whose lips I am raising the glass of wine so that I may drink. On the other hand, I can, of course, be selfish "against myself" in this sense – and many people often are: "I do not care how much I'll suffer tomorrow from headache: I'll get drunk tonight!" Of course, it would be necessary to go into more detail about self-care and its various forms to present a full moral theory. But the short remarks given here may suffice to show in which way the "ought of practical reason" analyzed by Kant can be incorporated into the strategy of foundation outlined above.

for me, be it intentionally or not; 3. her being *useful* for me to attain other goods; and accordingly, there are friendships based on virtue, on pleasure and on utility). The volition of this good with its generic, qualitative and quantitative specification (and thus limitation) I turn unto the other person. Therefore, automatically and without further regulation, original benevolence has an inherent measure or criterion: "How I want for myself, thus I want for you." This, of course, is the golden rule or at least one version of it (however, not yet in the form of a *prescription*). Consequently, the trope of volition into benevolence inherently generates the golden rule (very much in the same way as the trope of practical reason into *pure* practical reason *simpliciter* generates the categorical imperative). Just to give it a name, we can call this the "principle of justice", viz. in the Aristotelian sense of the term as a criterion of balance or equality.[30] With this, the second function we may expect a fundamental moral theory to fulfill is guaranteed *in principle* (only in principle, since we have not got yet to the level of Morality; but we already have a criteriological principle we will be able to draw upon once we reach there): the function of articulating at least some formal or criteriological principle for the content of Morality.

Now that we have gotten so far – understanding that the original trope which opens up egocentric consciousness for normativity is turning knowingly away from the ego centeredness of my consciousness and understanding that it is concrete and actual practical knowing *de me* which is turned away *de me* unto an Other Self – it seems natural that there should be a trope of *cognitive* consciousness *de me* which corresponds to benevolence as the trope of *volitional* consciousness *de me*. And it is not difficult to se what this should be: as I *know* myself as "my self", this trope should make me *know* the other person as an "other self". It should make me know her "in terms of self" – as I originally, non-inferentially, non-intentionally and non-discursively know myself (and *only* myself) "in terms of self", i.e. by knowledge *de me*. Now, being a conscious self is something I cannot *cognize* in an object. I can have no *cognition* of another consciousness and thus of Other Selves in the sense of "*de se* knowings" or "egocenteredness".[31] To know an Other Self is not to *cognize* her as such, but to *recognize* her as such. If I recognize an Other Self, I turn onto her my own mode of being-in-cognition, i.e. of knowing myself *implicitly and non-discursively* as the center of my cognizing, as that unto which my cognizing is directed. In *philein*, I *implicitly and non-discursively, non-inferentially and non-constructually recognize*, i.e. I *originally* recognize

---

[30] Cf. NE V.
[31] I cannot even have objective cognition of my *own* consciousness.

an Other Self. The "cognitional content" of this re-cognition is her being equal to me – not in her objective characteristics which I *cognize*, but in her subjectivity, in her being a self like I am myself, which I *recognize*. In English, we could say: I recognize the Other Self as a *fellow* – as a fellow citizen, a fellow human being, a fellow living being.

It is obvious that with recognition we have one more original principle which our foundational theory of morality may draw upon, that is, once we get there. And it also seems obvious that benevolence and its inherent principle of justice need the principle of recognition as its counterpart. To modify a well-known Kantian formula: benevolence (or justice) without recognition is blind, recognition without benevolence is empty. Consequently, it is characteristic of *philein* that both volitional and cognitive consciousness or that consciousness in both its volitional and cognitive aspect is *at once* "turned around", that *philein* is the trope of consciousness as a whole, in its integrity.

Once we understood that it is consciousness in both its volitional and its cognitive aspect, which is turned around in the basic, original trope of *philein*, we can coin a term for the structural aspect of this trope. I suggest we call it the "conversion of consciousness".[32]

It is very important to see that the conversion of consciousness is not its inversion. The egocentric orientation of consciousness is not substituted by an orientation unto the Other Self. This would destroy benevolence and recognition because at the moment my knowledge *de me* is abolished I cannot recognize somebody else as an Other Self – for I can only know by means of myself, thanks to my own knowledge *de me,* what it is to be a self. Likewise, I cannot wish somebody well if my own will is abolished. It must be *my own will* which is conversed unto the Other Self. Consequently, egocentric volition must persist, as well as my egocentric cognition. Therefore, the original conversion of consciousness has four aspects: (egocentric) cognition of an Other Self, (egocentric) desire for him (as a *phileton*), (converse) recognition as an Other Self, and (converse) benevolence directed unto him. The whole of this we can call "*philein*" in the tradition of Aristotle[33] or (with some caution) "love" in English. Different from charity or *eros*, it is the balance of desire, cognizing, benevolence and recognition that characterizes *philein*.

---

[32] I have explained this principle in more detail in: Utz, Konrad. *Freundschaft. Eine philosophische Theorie*. Paderborn: Schöningh, 2012, especially 63-105.

[33] Cf. once again his definition of *philia* in NE1155b30f. He hasn't got cognition in his list, but it is obvious that I must cognize someone to love him. And in place of recognition Aristotle understands (mutual) unconcealment as the fundamental cognitive aspect of *philein*. But this, in my view, is something much more complex and cannot be the result of the basic conversion of the cognitive aspect of consciousness.

And this is what conversion of consciousness *originally* is – though one or another of its aspects may become predominant afterwards. Note also that "love" used in the sense of the conversion of consciousness becomes a very broad term. *Wherever* a relation between subjects is characterized by *some* (even weak) desire, cognizing, benevolence, and recognition, there is *love*.[34] In the following, we will call "friendship" love which is reciprocated. Consequently, "friendship" also becomes a very broad term, as in Aristotle – different from its usage in English and other modern languages of European origin. There is or can be "friendship" between lovers, family members, citizens, members of communities of all sorts and even of humankind. Therefore, the principle of Conversion of Consciousness is not restricted to close, personal relations, but applies, possibly, to all human relations and even to relations to non-human living beings. At the same time – in Aristotle as in the theory developed here – exclusive, intimate, personal and free friendship (i.e. in the sense of the modern usage of the term) is *paradigmatic*, since here the different moments of love are most developed, and most balanced.

A Theory of *philein* as Fundamental Ethics

The theory of the conversion of consciousness is a *foundational* theory exactly in the sense described above: it articulates or renders explicit how we can love others. It does not prove that we *should* love others by way of some normative principle; and it does not show that it *necessarily so happens* that we love others by way of some causality. But it does show how love, benevolence and recognition can be *original* or *authentic*, i.e. non-derivative and non-constructive, non-reducible to egocentric volition and cognition or to the external force of natural causation (some circuits in our brain, mirror-neurons or such things). And the theory of the conversion of consciousness offers formal, criteriological principles for developing moral content: the principle of justice and the principle of "recognizing as equal" or of fellowship.

Of course, we can indicate what *occasions* love. But these occasions do not explain or explicate love. Beauty is a typical occasion for love. But I do not love the Mona Lisa even though I find her beautiful. Sexual attraction is also a typical occasion for love. But, very unfortunately, sexual attraction is not always the origin of benevolence and respect. Occasions for love are not explanations or explications of love. Love is

---

[34] Of course, in this usage, love is not a sentiment (though, naturally, it may imply sentiments), nor is it selfless.

"original". Benevolence and recognition are not reducible to desire and (merely descriptive) knowledge.

As in the case of Kant's reflection of practical reason upon itself, there can be no explanation and no explication (in the sense given above) for the conversion of consciousness.[35] There can be neither descriptive nor prescriptive law under which this conversion occurs and which make this occurrence necessary or obligatory. Like reasons turning pure in Kant, the conversion of consciousness is a fact the occurrence of which is irreducible and non-(re-)constructible (out of other things). It is what could be called an *original* fact – an "originally originating" fact. In this respect, the conversion of consciousness is no "better" nor "worse" off than pure practical reason.

However, the occurrence of the conversion of consciousness or of *philein* is different from Kant's rational-reflexive turning in that it is *clearly* an occurrence in space and time. It occurs in view of another subject (typically a human being, but possibly also some other living being). The conversion of consciousness has an *occasion* (not a cause) *in space and time*. It is not clear where Kant's reflection of practical reason upon itself originates and what occasions it. Since autonomy or freedom belongs to the intellectual world which is abstract, non-temporal and non-spatial, it seems that its occurrence should take place in this world and, therefore, is not occasioned, but simply given (as an original, irreducible fact). Yet this appears somehow odd, because practical reason clearly *can* be instrumental or heteronomous, it *need* not necessarily be autonomous or free. But how should we understand this: is reason autonomous and heteronomous, free and unfree at the same time? Or are there two reasons within one subject one of which is free and the other is not? Does reason speak in two tongues, that of morality and that of instrumental deliberation? Or are there two speakers within the subject, pure reason on the one hand and inclination on the other? But if this is the case, why does Reason speak, on the one hand, on its own, and on the other helps inclination to articulate its goals and to find the means to reach them? How can reason be, on the one hand, a *subject*, giving commands; and on the other a mere objectual *instrument* used by an inclined Subject? Or are neither inclination nor reason the *real* subject, but the real subject

---

[35] In the following, I'll compare the theory I want do develop only to that of Kant which I exposed in the beginning. This restriction is first due to limitation of space. But second there are in my view very few theories to which the "theory of the conversion of consciousness" could be (*directly*) compared as a foundational theory of ethics, since very few theories follow that notion of what Foundation of Ethics is and what its functions are which Kant exposed and which I regard as essential. And those few who do fundamental ethics (more or less) in the vein of Kant (e.g. Apel, Habermas, Rawls) stick firmly to the Kantian *topos* of rationality (even if in different forms), which I am criticizing.

is somehow behind or above both and then *identifies* with either one or the other? But then, what governs or motivates this identification? How could it be that I *should* identify, i.e. that I have the *duty* to identify with pure Reason and *not* with inclination, if duty is something that can have its place and origin (its *topos* and trope) *only* in reason? – All this seems absurd. In my view, it indicates that something is fundamentally wrong with Kant's conception of the *topos* and trope of moral obligation.

But even if it is not, we still do not know what occasions reason's turning pure in the intellectual world. Note, also, that the *form* of pure reason is aprioristic and hence *necessary*: Reason's purity is a *necessary* possibility. But its being *in fact* pure is *not*; it is a *mere fact*, the fact of reason. Reason *can* be instrumental, and it *can* be non-instrumental; so it *cannot* be an aprioric necessity that it is non-instrumental. We cannot deduce or (re-)construct this fact, not even apriori, for this would destroy morality. Because of this, it is by no means clear that every rational subject should in every instant have morality. We can never be sure anyway (not even about *ourselves*), as Kant points out, if we are hearing the voice of pure reason or the voice of self-interest in the disguise of morality, using *conventional* morality as a means (to avoid punishment, to gain recognition, etc.). The *intellectuality* of the fact of pure reason does not guarantee in itself that a rational subject in the *empirical* world should always be a *moral subject*.[36] As a result, it does not seem that Kant's intellectual trope of reason turning on itself can guarantee the fact of morality in any better way than the conversion of consciousness can.[37] Consequently, *formally* the conversion of consciousness gives the same "security" to the fact of morality as pure reason would give (if "pure reason" was a viable option, which I think it is not). At first sight, it may seem that it offers less *stability* once it has, in fact, occurred. One is inclined to assume that, somehow, reason cannot forget about the *fact* of its being pure. But, of course, this fact only persists as long as reason *remembers* this fact or as long as it continues to abide by its purity. It seems *obvious* that we cannot forget that we are free once we have reached the knowledge (*de se*) of ourselves being free (which, biographically speaking, we clearly have *acquired* at some time). But in the same sense it seems obvious that I cannot *forget* my *philia* for another

---

[36] Naturally, this has nothing to do with the question of subjects actually *acting* morally. Even if a subject is conscious of its moral obligation, it can act against it (not by use of its freedom, but of its *liberum arbitrium* or "*Willkür*" – it is well known that Kant, like many other philosophers, distinguishes between the two).

[37] I haven't shown yet how *morality* originates out of *philia*; but since *philia will* be the origin of morality according to the theory I propose, the question of the "security" of this origin can be discussed already at this stage.

person (as long as I do not forget that person herself). Naturally, I can lose or renounce my friendship to her. But that is an altogether different question.

Again, this does not only apply to close, personal friendship. It also applies to friendship in the wider sense, to friendship towards family members, citizens, humankind and even towards the community of living, sentient beings. Once I have, in fact, realized the Conversion of Consciousness towards these groups, i.e. once I have established a relation of recognition and benevolence towards them, understanding them to be my fellows in at least some fundamental sense, this act is normatively irreversible. Of course, I can, *in fact*, disregard them and harm them afterwards. But I know that I should not do this.

Of course, the conversion of consciousness is altogether different from pure reason with regard to its *content*: it is directed toward the actual, concrete person I am encountering. Therefore, any obligation that may spring from *philia in the first instance* can only be obligation toward this individual person. What is more, not only *philia* as such but even the content of benevolence involved in it is actual and concrete; it is not aprioristically and universally determined: that good which I desire for me from the beloved, I desire for her. That good which I, in fact, receive from her, I want her to receive, in fact, from me.[38] The obligation of *philia turns* more and more universal or general in the course of one's individual life as also in the course of the "life" or the history of a community and of mankind as a whole. There is *in fact* something like "universal" or better "general" friendship – nowadays at least for humankind, if not for all sentient living beings. Unfortunately, given space constraints for this piece I am unable to demonstrate how this generalization of *philia* works.[39]

But even without this, it is evident that benevolence has something like an "aprioristic minimum". If I wish somebody "the good for her own sake", I necessarily imply the wish that she is able to receive from me. And this of course implies that she should *exist* – as a *receptive*, i.e. as a *conscious* being. Therefore, if I wish another person *well*, I necessarily, with this, wish at least that she should *live* – since life is the empirical condition of being conscious. And respecting an Other Self aprioristically implies the minimal content of this self having interests and opinions of

---

[38] This is the first, immediate specification of benevolence. In the next step, reflection modifies *equality* of what I want to give and to receive to *equivalence*. This is one of the processes which develop the structures of friendship beyond its first, most basic form, which I cannot describe here.

[39] For more details, cf. Utz, Konrad. *Freundschaft. Eine philosophische Theorie*. Paderborn: Schöningh, 2012, 228-231, 320-326.

her own, of her having her own egocentrically determined "good for her", and it moreover implies that she is in some fundamental sense equal to me as my fellow. And this in turn implies that once I have developed a *conceptually* articulated regard for *myself* as what I fundamentally am – a family member, a clan member, a member of my people, a human being, a living being[40] – I locate myself within a minimal or basic fellowship with all those who are *the same* as me, i.e. who are equal to me in this fundamental regard. Consequently, the theory of the conversion of consciousness or of "friendship" is in fact capable of articulating "universal" or "general" content, even if the criteriological principles it offers are even weaker than the categorical imperative.

But even in this respect, the "conversion of consciousness" is not much worse off. Where does the Categorical Imperative get content to turn morality concrete? From maxims. Where do maxims come from? Kant does not tell, but most plausibly from inclinations, education, cultural conditioning etc. – How do the abstract forms of benevolence and recognition turn concrete? By the actual encounter with the Other Self, by my desire for her and my cognition of her, by living together with her, by being conditioned by the community I form with her and others.

Friendship

As explained from the outset, the conversion of consciousness is the rendering explicit of the *first* trope of consciousness. As such, it corresponds to volition turning to reason in Kant, i.e. to volition's turning away from its immediate self-execution in simply doing what it wants (drinking, eating, walking, scratching, etc. – normally because some impediment emerges). Because of this, the primary conversion of consciousness does not establish yet the realm of normativity. In Kant, it is only the second trope (by my count, he himself does not count turning to reason as a distinct act) that constitutes ought-consciousness: the *reflection* of (instrumental) reason back on volition, telling it what it *should* do in order to realize what it *wants* to do. And it is only the *third* trope which constitutes moral consciousness, i.e. consciousness of being unconditionally obliged: the reflection of reason upon itself, its turning pure.

---

[40] Note that *emancipation* from my family, clan, etc. implies that I *do not* regard my self (any more) as fundamentally determined by that community; and this implies *directly* that fellowship to *this* community cannot be for me the ultimate level or the ultimate boundary of recognition any more. Those who are equals to me, i.e., equal to what I fundamentally am (by my own self-regard), cannot be only the members of my respective community any more.

The simple, basic conversion of consciousness or simple love is not conscious yet of being obliged. In basic benevolence I *simply and directly* want "the good for the beloved"; it is my own volition, not something which is elicited from me, not a prescription turned against me. My own volition has turned away from me; but it has not yet been reflected back on me as an ought.

However, there already is a first trace or a foreshadowing of normativity in the basic consciousness of love: in executing my benevolence, in doing my friend good (for her own sake), I am conscious of doing what I want. But I am also conscious of this being right. I have consciousness (literally con-sciousness, accompanying awareness) of my action being not only subjectively good, but also *objectively* good. I do not do this action *because* it is objectively good or right, I do it because I want, because I want to do the other person well. But that very volition of mine, in my own consciousness, has the character of being (objectively) right. Note: the good I wish for the other person is *subjectively* good – not for me but for her. But to wish well not only for me but for her, i.e. the *orientation* of my volition on her seems objectively good or right to me. And it seems so immediately, "on its own evidence", not because I thereby obey some norm. Given the theory presented above, it is obvious why this should be so: since, in the trope of benevolence, volition is turned away from its ego centeredness, it is not validated by one's own ego any more (the subject does not want something "*because it is good for me*" – but rather "because it is good for you"). At the same time, it is not yet validated by the other subject insofar as this subject becomes the center of volition only by the conversion of volition, so it cannot validate this conversion. But the benevolent volition *is* valid (in the consciousness of the subject, in its *de se* knowledge: she knows herself wishing the other well). It "continues" its validity "into" conversion – which is just to say that it continues volition or that it its volition itself which turns to the Other Self. Consequently, since this benevolent volition is not validated by the subject any more, its validity stands on its own. And this means: it is objective (in the subject's own consciousness).

There is no other way in which volition could be *originally* right but in its turning unto an Other Self. It could be *derivatively* right by fulfilling some norm. But in that case this norm has to be right. And, as has been shown above, nothing can be right for me if it does not originate (directly or indirectly) from my knowledge *de me* – by some trope or other. What is more, since benevolence is the basic (volitional) trope of consciousness, normativity can only have its origins in this trope, that is, its objective validity must spring from that of benevolence. As such, the conversion of consciousness or love already displays an aspect of "right". This

develops into consciousness of ought when love turns into *friendship*. This happens – according to Aristotle, and I follow him again – when love is reciprocated by the beloved and when those loving one another *live* their love by spending time together and by doing things for one another or together.[41] Friendship, the reciprocation of love, is the *topos* of the second turning or trope by which benevolence turns into ought-consciousness.[42]

In the course of friendship friends (get to) know each other, they desire each other and things from each other, they wish each other well and recognize each other. Therefore, if I have a friend, it naturally happens that the one I wish well articulates some wish of hers to me. (If this does not happen, it means that she either has no desire for me or she does not live her friendship with me; in both cases, there is no *real*, "realized" friendship.) Since in my own loving consciousness I am "wishing for her", and since I become conscious of her wish *within* this consciousness (or within this "orientedness" of my volitional consciousness), *eo ipso* this wish constitutes an ought for me in my consciousness. I do not have to ask myself, as in the cases Kant refers to, if I want to accept this prescription (which the wish of the other person represents to me) into my volitional knowledge *de me* – because of consequences I want or do not want for me. Since I love the other person, her wish immediately articulates itself in my consciousness as an ought.

This ought is distinct from my primary volition and even from my benevolence in that it really comes "from the other", in that it represents a prescription directed at me. It is not any more what I *directly* and *spontaneously* wish for the other in my own benevolence; it is *not* what I want. Maybe it is not contrary to my own desires. But it does not *spring* from my own volition and thus is, at least in the first place, not what I myself want, but what is elicited from me. And this is why it is in my own consciousness an ought – but not only an objective one (as Kant would have it), but a subjective one, an ought which I know myself to "stand within" or to be bound by.

This ought, however, is not unconditioned. It is not conditioned by my self interest, for the basis of its validity is my benevolence. But

---

[41] This is the *original* way of entering into friendship. However, there also is a derivative way – derivative in the logical sense, not in the temporal, for most of our first friendships are derivative: we are born into friendships or fellowships that have been established by other people, like our family, our people, our religious community etc. And we may enter into friendships or fellowships concomitantly when the person with whom we're establishing a friendship already has other friendships which, somehow, bear on the friendship we're about to contract (e.g., when one person marries another person, she normally not only enters into a new, specific form of fellowship with her, but also with her family).

[42] Something analogous happens with recognition. But since my space is limited, I'll restrict myself to benevolence.

since primary benevolence is always generically, qualitatively and quantitatively specified and thus limited, the demand of the other person may exceed my benevolence. Everyone knows these cases: typically, I first feel awkward, then indignant. This feeling would be inappropriate if the other would just have presented me a hypothetical imperative which I can obey according to my own preferences, as Kant would have it. But this is not the case: I have been wishing the other well *beforehand*; with this I (implicitly) was ready to accept her wishes into my consciousness as oughts; consequently, I somehow have "given myself away" to her. My indignation is explained by the other person abusing (be it intentionally or not) this "license" I have given her beforehand: my readiness to do her well. But this means that this readiness was in fact a readiness to receive her wishes *directly* as oughts; for if it was only a readiness to accept them on the deliberation if I *want* to accept them, I should not feel offended. I would do nothing else but state that I have come to the conclusion that I do not want to accept them.

Such conflict between my benevolence and the desire of my friend very often is the *occasion* to effect the third trope: the turning onto friendship itself. In the situation of such a conflict, I typically ask myself: "What do I really *owe* my friend?" Evidently, what I owe her is not simply what she desires from me. Just as evidently, it also is not simply what I spontaneously desire to give her. To determine what I owe my friend, it is not enough to look at her or at me. I have to look at the *friendship* that binds both of us. As explicated (or rendered explicit) above, benevolence *originally* has an objective aspect. It is *right* – independently of my own, primary volition being right or of my friend's volition being right. In the original trope of love, benevolence was actual and situational. But once friendship has been established, benevolence and recognition have been *established* between the friends in question. The general structures of this establishment we may call the (specific) *ethos* of (a specific) friendship. The objective aspect, which was *present* in the original act of benevolence and recognition, is in friendship instituted and developed beyond mere momentary presence. It is not objectivity as such which is established by friendship and revealed by the "third trope", but its independent subsistence beyond a mere aspect or momentum of my (conscious) actuality.

In asking what I owe my friend, I turn to the objective reality of our friendship and its *ethos*, the objective reality of the concrete, specific benevolence and recognition we established between us. This objective reality may not and need not be real and objective beyond my and my friend's consciousness of it. We need not stipulate a *metaphysical* reality of friendship. But for the friends, in their consciousness, in their

knowledge *de se*, their friendship has objective reality: it actually binds them (both in the descriptive and normative sense), it is an objective bond between them. Friends may dissolve their friendship. But as long as it lasts, they (in their own consciousness) objectively owe certain things to their friends. As a friend, I know myself to be obliged not only by my friend, but by our friendship.

With friendship, we have reached the level of objective, non-conditional (that is: not dependent on subjective conditions) normativity. Have we also reached morality? That depends on what we take morality to be. According to the theory developed here, there is no absolute, aprioristic, universal system or doctrine of morality beyond the *ethē* of concrete friendships – including, as indicated, wider friendships in the form of families, peoples, states, religious and other communities, and, finally, the fellowship of all human beings or even all living beings. Morality is nothing else but the objective, ethical aspect of the friendships and fellowships I know myself bound by. This aspect often differs from what I egocentrically desire from my friend or what I desire to do together with her. But it also sometimes differs from what my spontaneous benevolence or sympathy makes me want to do.[43]

Since we normally have many friendships on different levels[44], and since all of them have an aspect of morality, there is the possibility of genuine moral conflict. According to the theory developed here we must accept the tragedy of moral conflict as part of the contingency of our human existence. However, it rarely happens. The conflicts which may arise between demands of different *ethē* very seldom are, in fact, aporetic. Normally we know what we owe to whom and which ethos is the ultimately binding one in which context – and by no means have the claims of closer friendships always priority over those of wider ones. E.g., I owe to my children to care for their school education – and I do not owe this care to other children. But I may not use the power I got as a public office holder to promote my children's professional career. This would be against what I owe to the community in question. And since the public office and its powers are given to me by that community, I may not use them to foster other friendships I have. *As* an office holder, I am bound by the ethos of the community which bestowed this office on me. And, of course, once again, what friendship demands from me may not coincide with what my friend demands from me. My friend may demand that, as an office holder, I should give a public contract to her firm. But I

---

[43] E.g., I would very much like to give my friend a glass of whiskey, since he feels so miserable without it. But I know that I should not do so, because he is an alcoholic.

[44] Once again, I remember that I am using the term friendship in the Aristotelian sense, which doesn't limit it to friendship in the narrow, modern sense.

know very well that the ethos of our friendship has not got the power to support this demand: that it has, so to speak, no authority in the field of my duties as an office holder. And, again, what I owe to my friend may not coincide with my spontaneous benevolence. E.g., I would very much like to give the public contract to my friend (since she desperately needs it), but I know I should not do so.

There is no absolute, transcendental super-ethos beyond concrete friendships. There is the ethos of our most general, most encompassing and most basic friendship or fellowship: the fellowship of all human beings or even of all sentient living beings. Some people like to restrict the use of the term "morality" to the ethos of this minimal friendship. As long as that is a mere terminological decision, nothing depends on it. But according to the theory developed here, the ethos of all-encompassing friendship, as any other friendship, has been brought forth and has been developed in human history, it is not substantially different from them. Throughout the centuries, individuals and peoples have broadened their comprehension of who belongs to "all of us", i.e. of the extent of (basic, fundamental) recognition; and they have deepened their understanding of what goods we owe to all, i.e. of the content of basic, fundamental benevolence.

This latter point shows, however, that we do not have to accept the *ethē* of our friendships simply as given, that we do not have to accept anything as "moral" for a given community that the ethos of that community says and that we must recognize only those as subjects (for its members) who that community regards as fellows. There are transcendental, universal and unconditional criteria to judge the "morality" of *ethē* themselves: benevolence and recognition, justice and equal fellowship, mutual care and mutual respect for individual freedom. If the ethos of a given community lacks in one of these or even violates one of these, it is objectively deficient or "bad" and objectively *should* be changed. We may or even should go beyond a given ethos by exactly the same fundamental principle which – in *principle* – gives *ethē* their validity: the principle of the conversion of consciousness. This means, on the other hand, that the development of our *ethē* under the transcendental criteria of this conversion is normatively irreversible. We cannot decide to go back, e.g., to slave holder society without incurring moral guilt: because this modification of our actual ethos (of western societies) would not be supported by the principles of recognition and benevolence, but, on the contrary, would violate these principles and the standard of their concrete realization we already achieved. However, the principle of *philia* does not give us or lead us to any *really* universal, unconditional, definitive and complete ethos which *somewhere*, *somehow*

eternally exists – maybe in the intellectual world. There is no morality beyond friendship. It is not possible to derive or deduce a concrete ethos out of the criteriological principles of the conversion of consciousness – as it is not possible to deduce actual friendships out of it. Love has to take place and friendships have to grow and to form *ethē* in real life. But the principle of *philia* is sufficiently strong to take us beyond the mere givenness of some (ethical) reality toward its *validation* or *critique*. And this is all we need.

According to the theory of friendship presented here, what is called morality in the traditional and the Kantian sense is a project. There is no end to this project and there is not even a definition of its end, i.e. of what would be its perfection if we ever reached there. But, nevertheless, the development of this project is not arbitrary. It is guided or *oriented* by the principles of *philia*, the principles of desire, knowledge, benevolence and recognition, the principles of equal fellowship and individual freedom. The inherent tendencies of this project are towards universality and individuality: the tendency to make benevolence wider and wider and to make recognition deeper and deeper. These tendencies stem from love to friendship itself, i.e. from the desire to "augment" friendship by making it wider and deeper. *Formally* and *abstractly* we can describe the goal of the development of friendship and thus of morality as all-embracing (all-benevolent), thoroughly respectful love. But this formula is not even clear and distinct enough to serve as a regulative idea. There is no other way but to go back from theory to life, to strenuous, fulfilling, sorrowful and joyful everyday life, trying to make friends, to make ourselves and our friends better, to make our lives better lives, to make us better friends and to make our friendships better. This implies that we will have to fight violence, disrespect and discrimination, sometimes by argument, sometimes even forcefully. A foundational theory of ethics is not superfluous; it can give us orientation in this. But it cannot substitute life – not even in the achievement of morality.

References

CASTAÑEDA, Hector-Neri. 'He': A study in the logic of self-consciousness. In: BROOK, Andrew; DeVIDI Richard C. (Eds.). *Self-reference and Self-awareness*. Amsterdam; Philadelphia: John Bejamins Publishing Company, 2001. p. 51-80.

COOPER, John. "Aristotle on the Forms of Friendship". *Review of Metaphysics*, 30 (1976), p. 619-648.

FRANK, Manfred. *Ansichten der Subjektivität*. Frankfurt a.M.: Suhrkamp, 2012.

KANT, Immanuel. Grundlegung zur Metaphysik der Sitten [1785]. In: _____. *Gesammelte Schriften*. Berlin: Königlich Preußische Akademie der Wissenschaften, 1911. v. 4, p. 385-463.

KANT, Immanuel. Kritik der praktischen Vernunft [1788]. In: _____. *Gesammelte Schriften*. Berlin: Königlich Preußische Akademie der Wissenschaften, 1913. v. 5, p. 1-163.

LABERGE, Pierre. La définition de la volonté comme faculté d'agir selon la représentation des lois (GMS 412). In: HÖFFE, Ottfried. *Kants Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar*. 4. ed. Frankfurt a.M.: Klostermann, 2010. p. 83-96.

MOORE, George Edward. *Principia Etica*. Cambridge: Cambridge University Press, 1903.

SOLOKOWSKI, Robert. Phénoménologie de l'amitié. In: MERLE, J.-C.; SCHUMACHER, B.N. (Eds.). *L'amitié*. Paris, 2006. p. 115-135.

STEIGLEDER, Klaus. Kants Moralphilosophie: Die Selbstbezüglichkeit reiner praktischer Vernunft.Stuttgart. Weimar: Metzler, 2002.

UTZ, Konrad. *Freundschaft. Eine philosophische Theorie*. Paderborn: Schöningh, 2012.

UTZ, Konrad. *Bewusstsein. Eine philosophische Theorie*. Paderborn: Schöningh, 2015.

UTZ, Konrad. "Praktische Vernunft in der 'Grundlegung zur Metaphysik der Sitten'". *Zeitschrift für philosophische Forschung*, 69, 4 (2015), p. 474-501.

UTZ, Konrad. "Quid mihi? Zur Methode der Grundlegung der Ethik bei Kant". *Deutsche Zeitschrift für Philosophie,* 64, 2 (2016), p. 213-227.

Endereço postal:
Programa de Pós-Graduação em Filosofia
Universidade Federal do Ceará – UFC
Av. da Universidade, 2995
Fortaleza, CE, Brasil