

Can a Consequentialist Be a Good Friend?¹

Tomasz Zyglicz, University of Warsaw

Consequentialism is often held to be self-defeating due to its incompatibility with intimate relationships. This objection is especially vivid with respect to friendship, highly voluntary character of which is believed to be irreconcilable with impersonal, teleological and maximizing attitudes. There seems to be hardly any place for the least necessary of loves in die-hard consequentialist's motivational structure. Another problem arises from the fact that both consequentialist and her friend might feel alienated from their relationship in the face of its immediate termination upon realization that it no longer provides maximal good.

In the first three sections of this essay, I am trying to clarify the notion of friendship. Next, I sketch a problem modern ethical theories face in accommodating friendship and then turn to discussion devoted specifically to consequentialism. Frank Jackson's and Peter Railton's takes on this objection are presented in the fifth section. In the final section, I put forth a way of looking at friendship which I believe best reconciles it with the demands of consequentialism. It revolves around the idea that friendships are constitutive to agent's personal identity, thus limiting the scope of possible actions even before she engages into consequentialist deliberation. A much overlooked fact that we represent certain persons as our friends exclusively in virtue of common past experiences is underlined in order to dodge the charge of partiality.

But friendship, as between our heroes, can't really be: for we've outgrown old prejudice; all men are zeros, the units are ourselves alone. Napoleon's our sole inspiration; the millions of two-legged creation for us are instruments and tools; feeling is quaint, and fit for fools. More tolerant in his conception than most. Evgeny, though he knew and scorned his fellows through and through, yet, as each rule has its exception, people there were he glorified, feelings he valued—from outside.

S. Pushkin, *Eugene Onegin*

1 I would like to thank the two anonymous referees and the audience of Zlot Filozoficzny 2015 for their valuable comments on the draft versions of this paper.

Friendship is certainly a relationship of a special kind. Unlike being a mother, a husband, a teacher, a student or a doctor, being a friend hardly seems to be restricted by social conventions. One is believed to be completely free to choose whether or not to engage in friendships at all. Moreover, each particular friendship exhibits a unique character, at least to the extent it emerges from a deeply personal interaction of two irreplaceable human beings.² Such diversity significantly undermines any prospective yearnings for a common denominator. Yet, the quest is worth its price, for there is a strong intuition that friends can be counted among the few things which make our lives worthwhile.

Having acknowledged the elusive nature of friendship, it would be at best naïve to strive for a complete analysis in terms of necessary and sufficient conditions. In this essay I aim at something more modest—to provide a plausible case for the compatibility of consequentialism and friendship. I begin by locating friendship within the historical outline of the Western moral thought. In the third section, I try to establish by means of conceptual analysis what we intuitively count as friendship.³ The fourth section is an exposition of the objection of our interest. In the fifth section, I sketch and criticize Frank Jackson's and Peter Railton's responses to the objection. The final section is a presentation of my solution, resting on an idea that friends, in virtue of being constitutive of an agent's identity, pre-theoretically reduce the scope of executable actions in any particular situation.

A word of clarification about the eponymous question is needed—for a reference to *good* friendship hints at the existence of a *bad* one. Why such a choice, given that a less loaded alternative, i.e. *real* friendship,⁴ is available? I follow Jonathan Schaffer's remarks that such a decision might lead to undesired ambiguities:

*But "real" is used flexibly in ordinary English to mark a multitude of distinctions. For instance, it can be used to mark the existent/non-existent distinction, the objective/subjective distinction and the basic/derivative distinction, inter alia.*⁵

2 Of course any relation involving two people has this formal feature. It is therefore important to emphasize friendship's being an essentially personal relationship, which seems not to be true of other relations mentioned in this paragraph. This is not to say that exemplars of the latter cannot exhibit deeply personal involvement of the parties, for often they do. The point is that it is not constitutive of most of them. Family bonds seem to be the likeliest counterexample to this claim; nevertheless, I would still urge that it is clearer to conceive friendship as a paradigm of personal relationships which may, but do not have to, underlie other, more conventionalized relations.

3 Methodology applied in this article strongly leans on F. Jackson, *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford, Oxford University Press 1998. Those suspicious of the conceptually analytic methods instantiated in this article, I would like to refer to this book.

4 Witness e.g. E. Mason, *Can an Indirect Consequentialist be a Real Friend*, 'Ethics' 1998, Vol. 108, No. 2, pp. 386–393.

5 J. Schaffer, *On What Grounds What*, [in:] *Metametaphysics: New Essays on the Foundations of Ontology*,

Moreover, it shall be noted that this essay analyses a phenomenon deeply entrenched in our everyday lives, rather than positing a new one. It is acknowledged that friendship comes in multiple variants, and the author feels in no position to judge which of them is the real one. In particular, it is not my intention to argue for any kind of error theory, trying to prove that our ordinary talk of friendship is misguided in a systematic way.⁶ Instead, I will take familiar intuitions about friendship at face value and do my best to expose what is crucial about them.

In order to see why I decided to use a prescriptively loaded adjective *good*, consider Neera Kapur Badhwar's⁷ distinction between instrumental friendship and ends friendship. Roughly, the former treats the relationship as a means to achieving some ulterior goal, e.g. the pursuit of pleasure or advance in one's career. The latter perceives friendship *per se* as an ultimate goal of friendly actions; in other words, one should always act *for the friend's own sake*. Unless noted otherwise, I will be referring to ends friendship whenever the term *friendship* occurs. Having settled that, the two intuitions underlying my terminological choice are:

- (a) Friendship as an experience of mutual care and trust is an irreducible source of happiness, and therefore something desired from a moral point of view.
- (b) Alienation resulting from a realization that one's friendship is merely an instrumental one may lead to suffering which is likely to outweigh the happiness such a relationship is capable of bringing about. Therefore, it is morally condemnable to create appearances of an ends friendship.

The following section is an attempt to tell a coherent story of why friendship had been almost completely ignored by philosophers for about two millennia. By no means do I claim to have succeeded in providing an exhaustive historical reconstruction of the concept. Rather, my exposition is meant to serve as an introduction to the topic by loosely tying it to roughly outlined general directions of the Western history of moral considerations.

eds. D. Manley, D. Chalmers & R. Wasserman, Oxford, Oxford University Press 2009, p. 360.

6 Should that be the case, one could no longer successfully hold it against consequentialism that it fails to accommodate friendship; unless such an error theorist managed to show that the friendship as understood under her new interpretation is likewise beyond consequentialist's reach. I am rather skeptical of the latter scenario, for the tendency is rather to criticize the notion for being too idealistic. Thus majority of such error theories would likely claim friendship to be something quite mundane in comparison to our ordinary and literary concepts thereof.

7 N. K. Badhwar, *Friends as Ends in Themselves*, 'Philosophy and Phenomenological Research' 1987, Vol. 48, No. 1, pp. 1–3, N. K. Badhwar, *Why is it Bad to be Always Guided by the Best: Consequentialism and Friendship*, 'Ethics' 1991, Vol. 101, No. 3, p. 483.

1. A fall from grace

It shall be no surprise that first things that come to one's mind when talking about friendship within a philosophical context are (books VIII and IX of) Aristotle's *Nicomachean Ethics*, Cicero's *De Amitia* and Plato's *Lysis*. We may thus fall victim to a false impression that the topic has received extensive philosophical attention. Yet, even an ethicist may find it difficult to name an exhaustive account of friendship which is not a product of the antiquity. It should be admitted that the Stagirite pretty much got it right when he explained what the relationship in question is about, for even today hardly any of the substantial claims made by him is being questioned. Nevertheless this is not a satisfactory explanation of why philosophers for around two millennia remained almost silent about a phenomenon not only present in our everyday lives but also counted among the few things which make them worth living.

It was the rise of Christianity that marked the removal of friendship from the moral realm. R. M. Adams invites us to compare two following passages from the Bible:⁸

Hear, O Israel: The Lord our God, the Lord is one; and you shall love the Lord your God with all your heart, and with all your soul, and with all your might. (Deut. 6:4)

Yet,

You shall love your neighbor as yourself. (Lev. 19:18, Matt. 22:39)

The tension between those commandments is often referred to as the problem of total devotion. For if one loves God with all her heart, soul, and might, it seems that not much is left for one's neighbor. In tackling this question, Augustine distinguishes between using something and enjoying it. A good believer shall thus not enjoy the love she has for another person, for this implies directing one's efforts at achieving a pleasure other than the one arising from the contact with God or from the pursuit thereof. Quite clearly a relationship originating from such love cannot be counted as an ends friendship. Certainly, the problem of total devotion could and already has been addressed in ways other than Augustine's. This particular solution was brought up, for it exhibits a long dominant way of conceiving relationships as means to achieving, rather than as a part of, the absolute good.

⁸ R. M. Adams, *The Problem of Total Devotion*, [in:] *Friendship: A Philosophical Reader*, ed. N.K. Badhwar, Ithaca, Cornell University Press 1993, pp. 108–132.

The tendency translated easily into secular ethics. A central notion for this shift would be Spinoza's *sub species aeternitatis*, from the perspective of eternity. It is a postulated point of view for an agent who is expected to exclude particular interests from moral deliberation. Impartiality and, closely related, impersonality become prerequisites for morality. This is hard to deny once we consider notions dominant in the modern ethical debate, like "the *universal* law" or "the *total* sum of happiness." A shift from cherishing personal relationships to a general directive of benevolence towards fellow human beings can be observed in modern ethical theories. Some may be inclined to justify this phenomenon by saying that morality is exclusively a matter of individual's relation to the society. Hence, intimate relationships are morally significant only inasmuch as they ground some wider-scale processes. Thus conceived, they may be neatly couched in terms of social conventions and thus easily accommodated within a more universal theory. However, their purely personal aspect, involving affection for its own sake, had been declared as *ex definitione* non-moral.

2. Friendship rediscovered

It was only in the second half of the twentieth century that friendship was brought again into ethical discussions. It all started from an observation that even though intimate relationships do not seem to translate directly into the well-being of a society as a whole, life lacking in them is self-evidently less happy than such life which involves genuine friendship. And this is often true even when the latter means lesser accomplishment in the other spheres of one's life. Friendship has been rediscovered as an intrinsic value. I take Railton's⁹ suggestion that a plausible way of deciding whether a value in question is an intrinsic or an extrinsic one is to conduct thought experiments. The idea is to compare two lives identical in all respects but the presence of the value in question.

For instance, one may be inclined to deny that doing sport has an intrinsic value, given that adrenaline rush, fitness, better health, and toughness of character sport helps to achieve may be gotten in some other way. Sport is thus instrumental for them, while this seems not to be the case for friendship. Consider life of a person enjoying a sense of security, high self-esteem, amusement, company, personal accomplishment, living in a society of mutual trust and general welfare. Still, there is a strong intuitive feeling that being devoid of genuine friendships significantly deteriorates such a life.

9 P. Railton, *Alienation, Consequentialism, and the Demands of Morality*, 'Philosophy and Public Affairs' 1984, Vol. 13, No. 2, pp. 145–146.

Keep in mind, however, that enjoying measurable profits from a particular person does not preclude one from valuing that person for her own sake. In fact, the two phenomena coexist in most cases and thus the temptation to reduce friendship to its less mysterious instrumental function, which better fits into universal moral theories. Once again, I shall apply the method of counterfactual cases in order to support an anti-reductionist claim. Consider the following sentence about a putative friend X. For any good G:

(A) *Were I not to receive G (which I actually receive), I would still care for her and do my best to maintain our special relationship.*

There is a strong intuition that should we find any such G for which (A) is false, then the relationship in question turns out to be an instrumental one. Notice that one should be possibly liberal with respect to what is allowed into the domain of G, for

(B) *You're hanging out with her only because of her looks/brains/wit.*

Sounds just as accusing as:

(C) *You're hanging out with him only for the sex/for the sake of your career.*

On the other hand, modal theorizing of this sort must be conducted with caution. Consider the conjunction of all G's possible within a context of particular friendship. It raises the question whether it is possible to sensibly speak of one's friend as a mere personal substratum, abstracted away from all her properties. Were the answer positive, ends friendship would probably turn out to be an unachievable idealization, irreconcilable not only with modern ethical theories, but also with the actual life. Still, I believe there is a way we can make sense of this extreme reading. I am going to get back to this issue in the next section.

3. What counts as friendship?

We are now ready to fill our theoretical considerations with some intuitive content about which I tried to remain neutral up until now. Recall that so far only one feature of the subject of our inquiry was made explicit, namely that ends friendship embraces disinterested care for the other person for her own sake. Clearly, it is by no means a sufficient characterization, for it may well be argued that this feature is present in other

types of intimate relationships as well. For instance, as Helm¹⁰ felicitously points out, what distinguishes friendship from love is that it is *ex definitione* mutual, whereas the latter appears to be more of an evaluative attitude. While the idea of unreciprocated love is perfectly plausible, this seems not to be the case for friendship.

Moreover, friendship seems to presuppose, at least relative, permanence, which again does not perforce come to our mind when discussing love. Witness how often one gets to hear about a *love at the first sight*. The adverb phrase in this example does not seem to be necessarily diminishing the affection's value, sometimes even to the contrary. Some may claim that genuine love, just like genuine friendship is a long-term venture and should be clearly distinguished from momentary affection. But maybe what this person *really* means is that genuine love must be underlain with friendship, which, as has just been settled, is a long-distance endeavor. This move has the systematic advantage of keeping the term *love* for a domain of biologically induced relationships, which encompasses both lovers' *eros* and a mother's tender feelings for her baby. We are thus not committed to refusing certain strongly intuitive cases of love, e.g. that of a mother whose baby died shortly after birth.

Let us now address the worry mentioned while discussing (A). The criterion clearly conflicts with a fact identified by C. S. Lewis:

*That is why those pathetic people who simply "want friends" can never make any. The very condition of having friends is that we should want something else besides Friends. Where the truthful answer to the question "Do you see the same truth?" would be "I see nothing and I don't care about the truth; I only want a friend," no friendship can arise—though affection of course may. There would be nothing for the friendship to be about, **and friendship must be about something**, even if it were only an enthusiasm for dominoes or white mice. Those who have nothing can share nothing; those who are going nowhere can have no fellow-travelers.¹¹*

This may leave us with an impression that friendship always amounts to whatever we use to fill in the ellipsis, as in 'about ...'. However, such a conclusion rests on the false assumption that friendships and *a fortiori* their relata, i.e. people, are static. *Persistence* does not imply *constancy* in this case. I would even risk defining persistence of friendship in terms of surviving changes of the subject of its *aboutness*. This

10 B. Helm, *Friendship*, [in:] *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), ed. Edward N. Zalta, <http://plato.stanford.edu/archives/fall2013/entries/friendship/> [accessed on 28.08.2015].

11 C. S. Lewis, *The Four Loves*, London, Geoffrey Bles 1960, p. 79 [emphasis mine].

seems not to lie too far from the common sense, for one important feature of friends is that they not only bring out each other's traits, but also exert influence on each other's systems of beliefs and values. We may take Aristotle's idea of friends as one's *another self* as a hint at how tightly interrelated psychological unity, personal identity, and friendship are in practice.

It should now be a mere formality to say that this sort of intimacy requires mutual trust, which can only be developed if friends are honest to one another, thus allowing for an accurate representation of who they are in one another's mind. For if one misrepresents her friend, then it seems questionable whether one is really *able* to care for the friend. The problem here is not her lack of affection, but rather that it is directed at a non-existent subject (a false representation of who the friend is).

Let us end this section by recapitulating the features of the ends friendship:

- a. Mutuality
- b. Relative permanence
- c. Aboutness
- d. Mutual trust
- e. Honesty

4. Problem of consequentialist friendship

Since we have specified the domain of our current inquiry, let us proceed to the discussion of a particular type of arguments against modern ethical theories concerning their failure to accommodate intimate relationships, as put forward by philosophers like Bernard Williams, Michael Stocker, Derek Parfit, and Neera Kapur Badhwar.¹² The needle of those criticisms is directed against accounts which associate morality with an impersonal perspective. Those include, most importantly, deontology and consequentialism. Roughly speaking, the former is mainly concerned with duties and obligations, while the latter evaluates actions with respect to their consequences. Alternatively, the difference may be perceived in the relation of priority between the right and the good. Deontology takes rightness to be primary to goodness, consequentialism—the other way round.

The coarse structure of the arguments against the above-mentioned accounts runs as follows:

¹² See: J. J. C. Smart, B. Williams, *Utilitarianism: For and Against*, Cambridge, Cambridge University Press 1973, M. Stocker, *The Schizophrenia of Modern Ethical Theories*, 'The Journal of Philosophy' 1976, Vol. 73, No. 14, pp. 453–466, D. Parfit, *Reasons and Persons*, New York, Oxford University Press 1984, N. K. Badhwar, *Why is it Bad to be Always Guided by the Best: Consequentialism and Friendship*.

- P1. Ethical theories are supposed to promote human well-being.
- P2. Modern ethical theories require us to take an impersonal perspective.
- P3. Personal relationships are one of the major sources of happiness.
- C. Modern ethical theories fail to accommodate important elements of our lives.

Many take C to be a sufficient reason for rejecting modern ethical theories. For instance, Stocker identifies them as a cause of what he calls moral schizophrenia: they require us to adopt a psychology divided between universal reasons and personal motives.

Nevertheless, the argument must not be seen as an outright refutation of, but rather as a challenge for impartial theories. In order to get a better grasp of the problem, let us make a quick excursion to an apparently distant debate in the philosophy of language/mind, i.e. the one over the essential character of first person statements. Despite an almost complete defeat of logical positivism, a core element of its program remains popular up until today. It is the idea that scientific knowledge should serve as a paradigm of our notion of knowledge in general. An important consequence of such an approach was the belief that all or at least all the large and most interesting parts of what we say and think can be expressed in terms of propositions; by propositions I mean the entities traditionally conceived of as the primary bearers of truth. Ontological difficulties aside,¹³ they are typically believed to express universal truths, i.e. such that are either true or false independently of who, when, and where utters them; even of whether they are uttered at all.

Thus conceived, propositions seem to have enough gravitas for philosophers to believe that anything worth saying or thinking can be expressed in terms of them. In other words: for every piece of information, conceivable or not, that may be of any interest for us, there is a proposition that conveys it. Nevertheless, a serious case has been made against the idea that propositions are the best candidate to play the role of the content of our mental states:¹⁴

Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient.

13 F. Kawczyński, *Czy filozofia języka potrzebuje pojęcia sądu logicznego?* [in:] *Współczesna Filozofia Języka. Inspiracje i Kierunki Rozwoju*, red. P. Stalmaszczyk, PRIMUM VERBUM, Łódź 2013, pp. 111–135 is a good overview of the arguments for and against propositions.

14 Two seminal works on this topic are J. Perry, *The Problem of the Essential Indexical*, 'Noûs' 1979 Vol. 13, No. 1, pp. 3–21 and D. Lewis, *Attitudes de Dicto and De Se*, 'The Philosophical Review' 1979, Vol. 88, No. 4, pp. 513–543. Reference to similar arguments by Hector-Neri Castañeda, Arthur Prior, and Peter Geach, the earliest dating back to late 1950s, can be found in these papers.

*Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts.*¹⁵

What the gods in this example clearly lack is the information about who they are. No traditionally conceived proposition can convey it, for in order to do so they would need to include a particular object which is contrary to their nature. Similar cases can be made for information concerning time and place. Thus what we are after here is typically dubbed the locating belief. The imagery of a little marker pointing at a point in the logical space, saying “you are here now” is sometimes used to elucidate this notion.

I believe considerations along the same lines underlie our worry about impartial ethical theories.¹⁶ For impartiality may be plausibly understood as a presupposition that moral statements do not make outward reference to any particular person. Instead they limit the scope of free variables by imposing conditions on them. For instance in

(S) “If a person is one’s friend and that person is sick and one does not carry a contagious disease, then one should go visit her and try to cheer her up.”¹⁷

The antecedent involves conditions on both the agent and on the situation. These allow the agents to classify the circumstances they find themselves in, while the consequent maps classes of circumstances to actions morally commendable in them. For instance, given that I am perfectly healthy and my friend John underwent a serious treatment two days ago, the application of the general statement to our particular situation would lead me to go visit him.¹⁸ Roughly, the idea is that moral statements

15 D. Lewis, *op. cit.*, pp. 520–521. A number of much more mundane cases were presented in the extensive literature as well. I chose the gods scenario as the most explicit one.

16 What exactly the relation between propositions and moral judgements is, if any, is a separate metaethical issue, which need not concern us here. My point is rather to point at the structural resemblance of the two problems: objectivity of scientific knowledge versus irreducibly personal belief on one hand, impartiality of ethical theories versus irreducibly personal relationships on the other. Much in a similar manner Perry draws an analogy between philosophy of belief and philosophy of visual perception (J. Perry, *op. cit.*, p. 20).

17 I am aware of the “should” having a deontologist sound to it. This is an ordinary way people express their moral judgements. Yet, there is no bar to modifying the sentence so that it sounded in a more consequentialist fashion, e.g. by replacing “one should” with “the best thing for one to do is to.”

18 Of course there is a number of difficulties here, concerning, *inter alia*, weakness of will, lack of other overriding judgements, prescriptiveness of moral judgements or identifying the situation on the basis of description. However, a proper treatment thereof would significantly distract us from the problem in question.

are universal, while any non-universal statement is *ex definitione* non-moral.¹⁹ It is a matter of practical application to subsume singular statements, like “John is sick” and “John is my friend,” under them. Now compare (S) with the following statement:

(S’) “I should go visit John and cheer him up.”

On the one hand, only (S) fulfills the requirement of impartiality, for a hard-nosed consequentialist/deontologist who cites (S’) as a justification of her action can be charged with egoistic motives. On the other hand, if I explain to John why I visited him by citing (S) and that I am perfectly healthy and that he is my friend who is in hospital, both he and I are likely to feel alienated from our relationship. John—due to the impression that should the circumstances be slightly different, I would not care to go visit him; me—upon the realization that this deed is not special in any respect in comparison to any other thing I do because of my moral convictions (for instance not crossing the street on a red light). And since friendship is a relation of mutual care, it is hardly possible to draw a clear line between *mine* and *his* feelings, alienation thus growing even further when we come to realize each other’s feelings.

Let us now focus specifically on consequentialism,²⁰ which I already tentatively characterized as a moral theory taking goodness to be prior to rightness. The stance in question can be plausibly reconstructed as one involving a two-tier evaluation.²¹ In order to compare the moral values of alternative actions, a consequentialist has to firstly look at their consequences and secondly to ascribe values to them, typically on the basis of the extent to which they contribute to the increase in the total sum of happiness. Consequentialism can be distinguished from other impartial theories

19 This is not to say that any universal statement is moral.

20 I suspect that much of what will be said carries over to other impartial theories, deontology in particular. The major divergence, as I see it, is that the latter view seems to essentially include obligations-talk which, partly due to civilistic associations (or rule-worship as J. J. C. Smart used to call it), seems extremely difficult to square with the demands of friendship, or at least with my reconstruction thereof.

21 J. J. C. Smart, B. Williams, op. cit., pp. 13–14. For the sake of simplicity, the following two paragraphs are stated in terms of act consequentialism. Nevertheless, consequentialism remains an umbrella term for a number of views, with respect to which I try to be as neutral as possible. Thus when asking consequences of what are to be evaluated, some of the answers are *acts*, *traits*, or *rules*. *Indirect* consequentialist allows exceptions from time to time provided they lead to an increase of overall happiness; her *direct* counterpart finds such a practice intolerable. There remains a question whether the agents should be judged by the *objective* consequences of their actions or rather by their *subjective* cognitive capacities. Those classifications cut across each other, giving numerous variants of consequentialism, which often turn out to promote extremely divergent actions when tested against particular cases.

Moreover, some philosophers believe maximization to be a far too strong requirement, hence argue that a consequentialist is only committed to satisficing. I will ignore this as a watered-down version of the view, consideration of which would only introduce complexities irrelevant to my argument.

by its adoption of a conjunction of two features, each of which corresponds to one of the aforementioned levels of evaluation.

- 1) Its adoption of *teleology*, in that the deeds are evaluated in terms of their efficacy in bringing about the desired consequences.
- 2) Its *maximizing* character which implies that, given a range of executable actions, the theory commends one with the best consequences, as opposed to the merely good ones.

It is important to notice that adoption of consequentialism does not entail that one always acts on a consequentialist basis. All in all, consequentialism is primarily a theory of moral evaluation. Some additional assumptions are needed in order to establish a link between moral deliberation and action.²² With this in mind, one may be tempted to reply to the problem of consequentialist friendship in the following way. What difference does it make—one would argue—what happens in one's head. All in all, it is perfectly conceivable that even the most diehard consequentialist can be a good friend, given that she does not apply her moral judgements to actions too often. Moreover, she may consider, as a result of her consequentialist considerations, that it is best not to confess to her friends that she is a consequentialist.

This response is not satisfying for at least two reasons. Firstly, as already mentioned, the alleged alienation is mutual. Secondly, and more importantly, friendship requires honesty. This feature may be conceived of as not telling that which is not true, but also as telling everything that is relevant. The first part is fairly uncontroversial and rules out the premeditated lie scenario outright. The second has to do with the Rawlsian publicity principle and needs some argument. Although it may be reasonably questioned as a general metaethical view, I believe it to be uncontroversial with respect to friendship. For, as I have already argued in the end of the third section, it seems that only communicating sufficiently extensive parts of our motivational structure to our friends, directly or not, enables them to create a satisfying representation of us; i.e. one which makes an apt basis for them to act for *our* own sake. This sort of *personalization* of friendly acts distinguishes them from those stemming from general benevolence towards human beings.

22 Many authors concerned about the *close relationships* objection assume that any plausible moral theory has to account for such a connection. E.g. Stocker, Railton, Jackson. I agree with them. Nevertheless, as I hope to show in the following paragraph, the sort of honesty required of a friend suffices to give rise to the problem of our concern, even if one never acts on her deepest consequentialist beliefs (assuming in the first place that it is possible for a person with such a fractured psychology to make friends at all).

5. Jackson's and Railton's responses

If the problem described above is a genuine one, then there are two major ways to justify a positive answer to the eponymous question of this article. One may either try to weaken the notion of friendship or that of consequentialism. I take the proposals of Frank Jackson²³ and Peter Railton²⁴ as representative of the respective alternatives. I will now sketch their views and point out their shortcomings.

Jackson argues that our question is best addressed by conceiving consequentialism from a decision-theoretic perspective, which he uses to establish that there is a number of situations in which best effects are achieved if agents act within a smallish group they are well acquainted with.²⁵ In a nutshell, his point is that a consequentialist can account for friendly actions by appealing to extensive knowledge of her closest ones, which allows her to best satisfy their needs (or—speaking less naturally—increase their happiness). The problem with this account is its being too *technocratic*.²⁶ It seems to mistake friendship for a cluster of empathy, charisma, personal charm, and possibly something else of a kind. One may claim this to be a close enough approximation, with friendship being nothing more but relativized empathy, charisma, personal charm etc. I disagree for two reasons. Firstly, such a stance has nothing to say about the intuition that we value friends for their own sake (For to say “John is my friend because he knows me best, can thus comfort me best, and vice versa” is still not to transcend the instrumental talk).

Secondly, consider the following setting. John, Jill, and Jake are a pack of friends since kindergarten. All being devoted consequentialists, they agreed it to be the best thing to do to assign one of them to spend more time with the lonely Jim. The underlying consideration is that due to the friendly dispositions each of them has developed thanks to the their long-lasting relationship, by using this friendly disposition on Jim one of them would be able to produce incommensurable happiness of Jim, when compared to the mere consumption of the existing relationship (i.e. the relationship between John, Jill, and Jake). Jackson's account seems incapable of explaining our intuitive reluctance to such unnatural schemes of friends distribution.²⁷

23 F. Jackson, *Decision-Theoretic Consequentialism and the Nearest and Dearest Objection*, 'Ethics' 1991, Vol. 101, No. 3, pp. 461–482.

24 P. Railton, op. cit.

25 These are the circumstances which require (a) getting to know certain individuals; (b) coordinating a series of actions; (c) mutual trust, respect, and understanding. Plus, those in which there is (d) a risk of agents' actions nullifying one another; (e) an obvious way to assign agents to tasks corresponding to their natural inclinations and enthusiasms. F. Jackson, *Decision-Theoretic...*, p. 474.

26 Interestingly enough, Railton, op. cit., whom Jackson extensively discusses, identifies this problem. Nevertheless, Jackson does not address it.

27 F. Jackson, *Decision-Theoretic...*, p. 478 distinguishes two *nearest and dearest* objections: 1. “How can consequentialists make sense of the fact that there is a relatively small group of people whose welfare plays a special role

Railton argues that the tension between friendship and consequentialism could be resolved by the adoption of an objective variant of the latter. “It is the view that the criterion of the rightness of an act or course of action is whether it *in fact* would most promote the good of those acts available to the agent.”²⁸ The objective consequentialism’s advantage over subjective consequentialism is that it endows agents with some liberty not to perform acts that appear maximally good-producing in given circumstances, if they believe doing so to be beneficial to pursuing one or more of their ground projects. This pertains to the notion that personal integrity is at least as important for generating happiness as even the most supererogatory but disconnected deeds. Railton cites personal relationships as a paradigm of non-moral values, pursuit of which could justify not abiding by consequentialist two-tier evaluation at particular times. Unfortunately, the generality of this position makes it prone to slippery-slope worries of the following kind:

P1. Sometimes exceptions can be made to what results from consequentialist deliberation.

P2. If sometimes exceptions can be made to what results from consequentialist deliberation, then always exceptions can be made to what results from consequentialist deliberation.

C. Always exceptions can be made to what results from consequentialist deliberation.

If the conclusion is correct, then is it really plausible to claim to not have abandoned the consequentialist framework? After all the entire point of positing the impartial point of view in general, and consequentialism in particular, is the desire to act without the arbitrariness of the sort presented in (C). This is not to say that the above argument refutes Railton’s response flat-out. The point is rather that his position is in a desperate need of a principled basis for telling in which circumstances exceptions can be made to what results from a consequentialist deliberation (due to friendship).

The two authors tackled our question from slightly different perspectives, those of subjective and objective consequentialism respectively. As we have seen Jackson had to water down our intuitive notion of friendship, while Railton’s solution raises doubts as to whether it remains faithful to the consequentialist spirit. Nonetheless, their contributions cannot be overestimated. The goal of the concluding section of this

in our lives, given the agent-neutral nature of consequentialism’s value function?” 2. “How can consequentialists make sense of it being the *particular* small group of people that it mostly is?” His ambitions in that article are limited to answering the first question. Concerning the second, he argues it is to be explained “in terms of empirical facts about human character and psychology.”

28 P. Railton, *op. cit.*, p. 152 [emphasis mine].

article is to put forward a fairly intuitive way of looking at consequentialist friendship which strongly leans on the ideas of two discussed authors.

6. The solution

Let us begin with two fairly uncontroversial observations. Firstly, that friendships do not come into existence in an instant, but rather require quite some time and effort to flourish. Secondly, that all we think about our friends *qua* our friends—i.e. in terms of person-specific qualities and not as a result of projection of some general belief on an individual—can be traced back to (a memory of) some common experience we shared.²⁹ Should that be the case, justification of any action friendlier than that dictated by general benevolence can be criticized for being motivated by mere sentiments. In consequence the agent may be charged with negative responsibility for not doing what is required by rational deliberation. But this is so only if there is an alternative executable action capable of bringing about more good.

One of the crucial factors determining the scope of executable actions are the characteristics of a given agent. For instance, it is generally commendable to administer first aid when there is need for it. However, one is likely to do additional harm if she does not know how to properly perform it. This observation applies to traits as well. A consequentialist would recognize going to medical school as a waste of time, if she was aware of her tendency to faint on every occasion involving the sight of blood. Any coherent project of a consequentialist society requires a division of work, based on each member being aware of her dispositions, skills, traits etc. This is but another way of saying that everyone should be able to immediately identify what she is good at or fit for. Otherwise consequentialism has to face a familiar objection that too much deliberation, even if it eventually results in the choice of the very best action available, is often worse than a merely good but immediate reaction. Let us posit the notion of *default self-representations* to play the theoretical role of the agent's cognitive capability of quickly identifying her properties relevant to decision-making. Assuming this kind of thoughts can take a linguistic form,³⁰ they would look more or less like the following:

29 I ignore the tricky cases which involve indirectly learning something about someone, e.g. via a testimony of a third party. Uncontroversially, these may supplement one's knowledge of her friend. However, the idea that a friendship may arise without any sort of direct contact between individuals strikes me as flat-out impossible.

This is not to deny the possibility of an epistolary friendship.

30 This assumption is made for the sake of presentation only. Nothing really depends on that.

- a. I am a well-trained surgeon.
- b. I carry a contagious disease.
- c. I am ill-tempered.

Because they help to determine what was called the conditions on an agent in the fourth section, they either broaden or narrow down the set of executable actions in any given circumstances. It goes without saying that the agent's set of default self-representations has to be of a limited size in order to successfully do its designated job in moral deliberation. Moreover, such a set has to be dynamic, so as to accurately reflect the agent's current state. Clearly, some of the set's items, like a and c, will be anchored in it deeper than the others, thus constituting *core default self-representations*.

Now, my claim is that our tension arises from considering friendly acts as those entering into the scope of possible actions primarily via conditions on situations, rather than via core default self-representations, and thereby via conditions on an agent.³¹ The fact that this is so almost automatically makes friendly acts seem bleak in comparison to a number of more spectacular alternatives, for almost always there is a strong sense that one could have done better (consider the conceivability of the claim that the tallest man in the world might have been taller than he actually is³²). But we have already established that a scrupulous examination of every such distant possibility is likely to have disastrous effects. That is why we need to evoke default self-representations in order to reduce the set of alternatives to choose from. There is no bar to claiming that beliefs of the sort "John is my friend" belong to the core part of our sets, since we have established that friendships play a constitutive role in one's personal identity. This need not contravene impartiality, for we have established that representations of friends amount to nothing but past experiences. All in all, there is nothing subjective about the past spatiotemporally located events, for no one can influence them anymore. It would have been better if I had taken years to prepare my army for today's battle. But since I had not, the best thing to do is to surrender in order not to send my soldiers to certain death. Both subjective and objective consequentialism are committed to evaluating from the perspective of a particular time, place, and person, with the entire world's (including that of an agent in question) history as a given. For if they do not, they fall into a *regressus ad infinitum*, all the way up to the Leibnizian question, whether the world we live in is the best among the possible ones.

31 This is my interpretation of Railton's postulate of "adopting a non-alienated starting point—that of situated rather than prosocial individuals." Railton, *op. cit.*, p. 171.

32 This example was given by J. Divers, *An Inconvenient Modal Truth*, 'Analysis' 2014, Vol. 74, No. 4, p. 575.

By way of conclusion let us summarize the advantages of the account just presented. Firstly, it nullifies the charge of partiality in that motivation to act friendly is conceived as a part of an agent's personal identity. As we have seen, it could not be sensibly raised against a consequentialist that she ought to abstract away from her personal identity, for doing so would significantly reduce her chances to achieve best results. Secondly, in that a particular friendship is considered a given at the moment of evaluation, we secure Railton's idea that friendship be treated as a non-moral value. I am nice to John, because I am his friend. Much in the same way as I go to work and heal people, because I am a surgeon. One cannot say that I am a surgeon in order to heal people, but rather that I became a surgeon in order to heal people. But the latter is objective past, already beyond my, or anyone else's, control. How we make friends need not concern us here. Moreover, the vividness³³ of representations of a person *qua* one's friend can serve as a principled basis for justifying exceptions to what a consequentialist calculation commands us to do. Thirdly, by linking default self-representations to available actions, we adopt a Jacksonian idea that the maximizing character of friendly deeds is plausibly explained by our knowledge of the closest ones. The crucial difference is that in my account this knowledge goes piggyback on objective spatiotemporal facts, thus not allowing for deviant redistribution schemes mentioned in the fifth section. Finally, my account is compatible with both objective and subjective consequentialism as modes of decision-making. The objective version, more or less along the Railtonian lines, is available, since a principled basis has been provided for telling when it is acceptable not to choose the best possible alternative and when it is not. My account also works with subjective consequentialism, for the default representations of oneself as linked by a special bond to a particular person secure that one automatically chooses to act for her friend's sake on a regular basis. Conservatism of interpersonal relationships secures that one abstains from doing so only if either the consequentialist value of the alternative is evidently and significantly greater or if the feelings for the closest ones fade away. The first option should not be feared, but rather welcomed, for, as we have established, friendship has to be about something. And what makes a better candidate for its content than what one party so deeply cares about? As for the second one—friendships sometimes end.

33 Worries may be raised concerning the essentially phenomenal character of such representations. On the one hand, I will not object, for their phenomenal nature helps us account for the uniqueness of each friendship. On the other hand, all such experiences must have something in common—otherwise no one would treat the friendship objection as a serious problem. We may plausibly approximate the degree of vividness as a function taking the subjective importance of common experiences and the temporal distance from them as its arguments.

Abstrakt

Czy konsekwencjalista może być dobrym przyjacielem?

Niejednokrotnie podnoszono jakoby konsekwencjalizm miał być samoucnestwiającym się stanowiskiem, ze względu na jego niekompatybilność z bliskimi relacjami. Zarzut ten staje się szczególnie wyrazisty na przykładzie przyjaźni, której wysoce woluntarystyczny charakter zdaje się być nie do pogodzenia z bezosobowymi, teleologicznymi i maksymalizującymi postawami. Wątpliwym jest czy ta najmniej konieczna ze wszystkich miłości w ogóle figurować może w strukturze motywacyjnej zagorzalej konsekwencjalistki. Kolejny problem wynika z faktu, że zarówno konsekwencjalistka, jak i jej przyjaciółka mogą doznać uczucia alienacji od łączącej ich więzi, gdy uświadomią sobie, że ta ostatnia zniknąć musi z chwilą, gdy przestanie generować najwięcej dobra.

W pierwszych trzech sekcjach tego eseju staram się doprecyzować pojęcie przyjaźni. Następnie zarysowuję problem, z którym zmierzyć się muszą nowoczesne teorie etyczne, by później zawęzić dyskusję do konsekwencjalizmu. W piątej części prezentuję Franka Jacksona i Petera Railtona odpowiadzi na omawiany zarzut. W ostatniej sekcji przedstawiam ujęcie przyjaźni, które moim zdaniem najlepiej godzi ją z wymogami konsekwencjalizmu. Zasadza się ono na spostrzeżeniu konstytutywnej roli, jaką przyjaźnie odgrywają w tożsamości osobowej aktorki, ograniczając tym samym jej zakres możliwych działań, jeszcze zanim podejmie ona konsekwencjalistyczne rozważania. By uniknąć zarzutu ze stronniczości, zwracam przy tym uwagę na często przeaczany fakt, że konkretne osoby przedstawiamy sobie jako naszych przyjaciół wyłącznie w świetle łączących nas przeszłych doświadczeń.