

Proof and content in Sidgwick and Mill

David Brax, Department of Philosophy, Lund University
david@brax.nu

The subject of this paper is the relationship between the proof for and the content of the utilitarianism of John Stuart Mill and Henry Sidgwick. It may be advisable to start of with some cautionary notes, though: the term ‘proof’ may be somewhat ill chosen in order to treat the accounts given by Mill and Sidgwick. Mill, indeed, goes to some length in order for us to understand that strict ‘proof’ cannot be given when dealing with fundamental moral matters and Sidgwick, in turn, does not even speak of ‘proof’, but prefer to talk about the *methods* of ethics. In both cases, though, the objective is to establish the truth of ethical utilitarianism, and the procedures by which Mill and Sidgwick arrive by it should be taken as proof, understood as the kind of proof that ethical propositions are susceptible of. The question, then, is what the variants of ‘proofs’ used can establish about the possible truth and, if actual, the content of utilitarianism. Where Mill and Sidgwick right in taking themselves to have established the kind of utilitarianism they ended up defending? For the sake of convenience only, I will start of with Mill and close with Sidgwick.

1. Mill

John Stuart Mill’s short but influential book “Utilitarianism” is probably (I’m not in possession of the exact numbers) among the most widely commented, criticized and interpreted philosophical texts in the language. It is, if I may be allowed a coquettish remark, partly surprising that this should be so, considering that it is a remarkably clear text¹. Throughout the years, Mill has been accused of making almost every conceivable philosophical blunder; and parts of “Utilitarianism” have even been used as textbook examples of philosophical mistakes. Admittedly, parts of “Utilitarianism”, if taken out of context, seem to exhibit these mistakes, and in a fairly enlightening way at that. But Mill did not commit all, or even most of them. Still, the manner in which he does not commit the mistakes he is accused of is a fitting entry to the argument he presents. The utter delicacy of Mill’s argument would, come to think of it, perhaps have been lost if it weren’t for the misguided critique. This is not to say that “Utilitarianism” walks free from objections, far from it (indeed, I add some of my own), but only that these objections does not point to *mistakes* on Mill’s behalf, but on weaknesses in the kind of evidence that can be given when it comes to the subject of morality. Being subject to objections of this sort, Mill seems to say, is an occupational hazard.

As early as on the fourth page Mill strikes the note that needs to be kept in mind throughout the reading of his text. Speaking of the proof that he is about to give he says that “it is evident that this cannot be proof in the ordinary and popular meaning of the term². Questions of ultimate ends are not amenable to direct proof. Whatever can be proved to be good, must be so by being shown to be a means to something admitted to be good without proof”. And then:

¹ It may sound trivial, but actually isn’t, that in order to understand Mill properly you should really read Mill himself. Not that Mill hasn’t got his fair share of brilliant readings (Millgram and Sayre-McCord, to mention two), but it remains clear that the most reliable account of the argument presented in “Utilitarianism” to date is really “Utilitarianism”.

² See Elijah Millgram’s excellent “Mill’s proof of the Principle of Utility” for a thorough treatment of Mill’s different conceptions of ‘proof’.

considerations may be presented capable of determining the intellect either to give or withhold its assent to the doctrine; and this is equivalent to proof (my emphasis).

This, Mill writes, is as true of first principles for morality, as well as for first principles in any area of knowledge. Strict proof in Mill's sense is *logical* proof, but nothing follows from logic alone. Explanation comes to an end somewhere, preferably at first principles. In his "general remarks" Mill does acknowledge what he takes to be an important difference between "the sciences" and the "practical arts" (to which latter category morality belongs): in the former particular truths precede the general theory, but the contrary might be expected in the latter case. This feature is one of the assumptions for which Mill does not provide evidence, but only appeal to our intuitions about morality. In effect, particularism is ruled out from the start. The specific "proof" which Mill gives for his theory, and which we will arrive at in due time, does not support this particular feature³.

Mill sets out to establish the "Greatest Happiness Principle", the principle that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. "By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure."⁴ It should also be noted that this proposition concerns the rightness of actions. Mill does not explicitly establish the link between the goodness of happiness and the rightness of actions being determined by their conduciveness toward goodness.

Mill supports his moral utilitarianism by claiming psychological hedonism to be true. "...the theory of life on which this theory of morality is grounded – namely, that pleasure, and freedom from pain, are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain." We will return to this "grounding" relation in a minute, but let's first stop and consider what it is that is thus said to ground moral utilitarianism: The psychological hedonism of Mill, as is well known, involves the claim that not "any old pleasure" would do as the ultimate goal of human action. As more or less sophisticated human beings, we would not be content with living the life of a pig, however pleasant. It would not satisfy us as it would the pig. This part of Mill's utilitarianism has received a lot of critique on account of being incompatible with the greatest happiness principle, but that critique was successfully met already in the original text: "It is quite compatible with the principle of utility to recognize the fact, that some *kinds* of pleasure are more desirable and more valuable than others. It would be absurd that while, in estimating all other things, quality is considered as well as quantity, the estimation of pleasures should be supposed to depend on quantity alone⁵." The quality of the pleasure is not

³ It is worth to notice here that the "considerations capable of determining the intellect either to give or withhold its assent" in most cases does not concern the final end (happiness) itself, but rather what is argued to be compatible with, indeed to follow from, happiness being the final end. That is: Mill does not, exactly, give considerations to provide support for happiness (or "utility") being the normative measure in isolation, but in a roundabout way, appealing to intuitions about the good, now argued to be instrumental goods. The things that were to follow by inference from happiness being the final end (and thus "proved" in the proper sense), are actually invoked in order to support that thesis. This is far from being a substantial objection, of course, but it is still noteworthy. Personally, I am much taken in by the elegance of this method – by making our everyday intuitions about morality compatible with utilitarianism Mill reduces their validity to this compatibility which, in turn, makes it possible for him to argue that under different circumstances, we are obliged, in order to stick by our values, to abandon them!

⁴ As we shall see when we arrive at the proof, it is important for the argument that the absence of pain is included in the concept of happiness, as well as pleasure. Both are, namely, the object of intrinsic desires.

⁵ The point here is beautifully made: it is easy to think that all "quality" differences in pleasure are quantity differences, or translatable to such. But Mill is here stating the precise opposite: pleasure is not a simple property, variable on the quantity scale alone. The quality difference is just as important when it comes to

by definition determined by the order of the mental faculty to which it is satisfying, but by the quality-ordering inherent in the desire for it. The quality difference is, in other words, not just a difference in the quality of the faculty involved, but a difference in the quality of happiness. Mill makes his case by arguing that the better of two pleasures is the one that “almost all who have experience of both give a decided preference, *irrespective of any feeling of moral obligation to prefer it*” (my emphasis). But note that this is a different argument from the one given directly above. The latter is a method to establish which one is the better pleasure, whereas the former is a statement about the psychological conditions for the betterness relation⁷.

Mill recognises that the capability to enjoy higher pleasures is not really reducible to the knowledge of them, a person may very well “lose his taste” for higher pleasure by losing his ability to enjoy them. Admittedly, this is fearfully close to begging the question: Whenever Mill comes across an example of someone preferring a baser pleasure, he can explain it away with this device. But again: it is important to note that the higher pleasures are higher, not necessarily but only contingently by being connected to the higher faculties. The *necessary* relation by which a pleasure is higher holds between the pleasure and the nature of the desire involved. There is really nothing (if not psychological limitations) which stops some particular instance of the happiness of push-pin from being higher than some instance of the happiness stemming from poetry; it just happens to be a fact that it is usually the other way around. And, also, preference is not always a reliable guide to the best of our knowledge. It’s more of a statistical indicative device. My point is that Mill is here collecting circumstantial evidence, not strict proof.

“This [happiness], being, according to the utilitarian opinion, the end of human action, is necessarily also the standard of morality; which may accordingly be defined, the rules and precepts for human conduct, by the observance of which and existence such as has been described might be, to the greatest extent possible, secured to all mankind; and not to them only, but, so far as the nature of things admits, to the whole sentient creation”. This leap, from being the end of human action to being the standard of morality, is necessary in order to account for the fact that rightness can be ascribed to actions of self-sacrifice⁸. But not sufficient, though: we also need to establish that the happiness principle is concerned with anybody’s happiness, not just the happiness of the agent. Mill, of course, arrives at this point by the generality requirement (built into the claim above that practical arts begins with general principles). But this requirement, as noted, is only supported intuitively.

Mill makes an important and quite intriguing point right here: in order to be the right action, the *reason* for the action need not be to achieve the greatest amount of happiness possible. That would be “to mistake the very meaning of a standard of morals, and confound the rule of

pleasure’s status as an end, and the support for this statement is inherent in the “proof” for the utilitarian principle, the desire as proof of desirability thesis.

⁶ Mill does not say, but does not deny either, that the feeling of moral obligation may very well be one of the higher capacities which give us “higher pleasures”.

⁷ There is, though, the problem of interpersonal comparisons when one is the more sophisticated. Mill can be used in a defence for elitism if understood in one way, and “strategic ignorance” could be the result if he is interpreted the other. Does the human really know what it is like to be a satisfied pig? And even if he does, what does that say about the value of the pleasure as experienced by the pig? Even if it would be better for the pig to be able to experience higher pleasure, that does not mean that the higher pleasure as experienced by the human is better than the lower as experienced by the pig. The only thing Mill can establish here, in other words, is what he did above: that hedonism does not require our behaving like pigs

⁸ Rightness is, thus, never an intrinsic property of actions. If not, of course, the action is defined as including its relevant consequences.

action with the worth of the agent". This, of course, is far from being an obvious mistake and is rather a substantial ethical question in itself, but Mill does not pretend to have given conclusive evidence for his view, only to show how it goes with ordinary intuitions about the rightness of actions. In a footnote, Mill elaborates this point further: "The morality of the action depends entirely upon the intention –that is, upon what the agent *wills to do*. But the motive – that is, the feeling which makes him will so to do – when it makes no difference in the act, makes none in the morality; though it makes a great difference in our moral estimation of the agent, especially if it indicates a good or bad habitual *disposition* – a bent of character from which useful, or from which hurtful, actions are likely to arise." This is rather peculiar: why should the *intention* be thought to matter, and not rather the actual effect of the action? The former seems only to be important as part of the indication of character, or as an indication of what is usually good effects. And is there really a clear line between intention and motive?

Mill then argues that morality cannot end by establishing the ultimate goal, but needs to provide guidance towards that end. "Whatever we adopt as the fundamental principle of morality, we require subordinate principles to apply it by".

Both these features come in handy when Mill explains that the best moral character may not be the "utilitarian" character. But there is a "natural basis of sentiment for utilitarian morality" if there were not, utilitarianism would surely be wrong. "But there *is* this basis of powerful natural sentiment; and this it is which, when once the general happiness is recognised as the ethical standard, will constitute the strength of the utilitarian morality."

Note here that all the explanations given are given in order to account for normal moral intuitions that exist side by side with the intuitive desire-proof, and to show that they are not incompatible with the greatest happiness principle. Naturally, Mill here makes an assumption about what intuition "accepted to be true without proof" is to be understood as *fundamental*. And *then* the other intuitions can be "proved" to follow, in a way⁹. He does not consider whether some or other of the other intuitions could be taken as fundamental, and the value of pleasure be understood derivatively.

But let's now turn to what is supposed to be the main subject of this paper: the concept of proof. Mill introduces his discussion with the following statement:

To be incapable of proof by reasoning is common to all first principles; to the first premises of our knowledge, as well as to those of our conduct. But the former, being matters of fact, may be subject of a direct appeal to the faculties which judge of fact – namely, our senses, and our internal consciousness. Can an appeal be made to the same faculties on questions of practical ends? Or by what other faculty is cognizance taken of them?"

What other faculty, indeed? As we have seen, Mill makes the leap from the question about ends to the question about desirability, and the utilitarian doctrine, of course, is that happiness is the only thing desirable in itself. Then, of course, comes the part where Mill makes the analogy with the senses. The only proof of visibility being that the thing is actually seen and the only proof of desirability is that the thing is actually desired. Moore, famously, said that Mill here mistakes the meaning of "desirability". "Desirable", he noted, does not mean "capable of being desired" but rather "worthy of being desired". It would indeed be a mistake to confound these two meanings. But it is a mistake which Mill did not commit. Admittedly, the analogy drawn suggests that he did, but the analogy does not stand alone in Mill's

⁹ Actually, they are rather shown not to be ruled out by the happiness principle

argument. The analogy to visibility is made in order to show that they both are *final* evidences for the property in question. Directly after the analogy, Mill writes “If the end which the utilitarian doctrine proposes to itself were not, in theory and in practice, acknowledged to be an end, nothing could ever convince any person that it was so¹⁰.” The “only proof” is thus not that pleasure is desired, but that it is notoriously desired *as an end*. Actually, if Mill did commit a mistake here, it was in supposing that being desired was the *sole* evidence needed for establishing that happiness is desirable. It is a necessary, but not a sufficient condition. To see that Mill was aware of this, just note that he acknowledged that we sometimes desire, indeed, that we sometimes *should* desire, other things than pleasure for themselves. “The desire of virtue is not as universal, but it is as authentic a fact, as the desire of happiness”. Mill is very quick to admit that the utilitarian doctrine maintains that virtue is to be desired disinterestedly, for itself. There are *conditions* for this, utilitarian conditions, but that does not change the fact that virtue could and should be desired for itself, namely *as part of happiness*. This seems to be a contradiction, but look at the following: “The ingredients of happiness are very various, and each of them is desirable in itself, and not merely when considered as swelling an aggregate”. It is important here to note that while being desirable in themselves, these “ingredients” would not be desirable outside happiness. Virtue, or whatever gets desired in itself through being desired as conducive to happiness becomes *part* of happiness when desired for itself. Happiness as an end is thus, to wit, instrumental to the instrumental values becoming part of the end itself.

“It results from the preceding considerations, that there is in reality nothing desired except happiness. Whatever is desired otherwise than as a means to some end beyond itself, and ultimately to happiness, is desired as itself a part of happiness, and is not desired for itself until it has become so.” It is not very clear whether Mill is presenting something that is true conceptually, or psychologically. It is clear that he intended it as a psychological statement and that *if* it is true, we have all evidence we can require for the truth of the greatest happiness principle. The problem is, of course, that it is not psychologically true. If, that is, “desire” is not given a very specific reading, making it necessarily true (in which case it follows, conceptually, that it is true psychologically). What *might* be psychologically true is that happiness is the only thing that is *notoriously* desired for itself, whereas the desirability of other things are always dependent on their consequences. And this, indeed, is the best argument for hedonism.

Mill’s argument is that the sole *evidence* that something is desirable is that it is in fact desired. He does not say that to *be* desirable just is to be desired (see Hall (1949), see also Sayre-McCord (2001)). Granted, but did he *mean* it or not? Mill writes that ‘pleasurable’ and ‘desirable’ “refer to the same psychological fact”, but does that mean that they share *meaning*? Actually, there is a way out for Mill without having to bring in conceptual truth: If it can be shown that a certain form of desire is only present in connection to something intelligibly understood as happiness, Mill can say that it is *this* desire we are looking for, and that it is psychologically (contingently) true that this relation holds. But his argument becomes much weaker if he does. And it is most likely that if asked to choose, Mill would have preferred the conceptual thesis. Some indication to this effect is present in the following: “I believe that these sources of evidence [self-consciousness and self-observation, assisted by observation of others], impartially consulted, will declare that desiring a thing and finding it pleasant, aversion to it and to think of it as painful, are phenomena entirely inseparable, or rather two parts of the same phenomenon; in strictness of language, two different modes of

¹⁰ But, as we saw above, it is not enough that this end should be acknowledged to be such, but it must also be compatible with most of our everyday intuitions about right actions, and Mill does use this device in order to provide “considerations capable of determining the intellect either to give or withhold its assent to the doctrine”.

naming the same psychological fact: that to think of an object as desirable (unless for the sake of its consequences), and to think of it as pleasant, are one and the same thing; and that to desire¹¹ anything, except in proportion as the idea of it is pleasant, is a *physical and metaphysical impossibility*¹².” (my emphasis). Naturally, in conjunction with the “desire as proof for desirability” thesis, this means that hedonism is true¹³. But which of these is then the “first principle” for which no strict proof can be given? Is it that desirability follows from desire, or that desiring something and finding it pleasant is the same thing? Mill does not say. But if the latter is the case “in strictness of language” and Mill would like to say something substantial, it is probably the former statement that should be taken as first principle.

Here is another problem for Mill: even if “desiring” something and finding it pleasant refers to the same psychological fact, this does not make happiness the only candidate for desirableness. It’s pleasure itself, not the thing found pleasant, which is supposed to be desirable. The desired thing could as readily be something external, and thus not a state of consciousness at all, and it is clear from what Mill says that only states of consciousness can be considered as being part of happiness. Indeed there are no clear indications in Mill’s text as to how the relationship between desire and happiness is to be understood. There is an ambiguity, a tension, in the psychological and conceptual characteristic of this relation. Naturally, the strength of the “proof” and, to an even higher extent, the content of the utilitarianism thus “proved” is much dependent on what interpretation is given at this point. Probably, Mill should be thought to argue that nothing is really pleasant but pleasure itself, but then, again, the statement depends on the success of the argument presented above. And, besides, it depends on a stretch of language.

What have been shown in the above? It seems to me that precious little of the content of Mill’s utilitarianism really *follows* from any “first principles”. Every part of it: generality, quality-ordering, happiness as being the only object for intrinsic desires etc. are *intuitively*, not inferentially, argued for. There is, then, some tension between the argument Mill set out to make and the argument he actually made. “Utilitarianism” presents, to repeat what I take to be the single most important quotation of the text, *considerations capable of determining the intellect either to give or withhold its assent to the doctrine; and this is equivalent to proof*. But to what doctrine?

Sidgwick

“The methods of ethics” being Sidgwick’s main work in moral philosophy defines it’s subject as “any rational procedure by which we determine what individual human beings “ought” – or what it is “right” for them – to do, or to seek to realise by voluntary action”. In a sense, Sidgwick does the exact opposite from Mill: the strongest intuitional evidence accrues not to ends of actions but to principles and maxims. Indeed, he can not rely on Mill’s kind of evidence since he refutes it. Psychological hedonism is not correct, and even if it were true that we desire nothing but happiness, that would still not mean that we always desire our own *greatest* happiness, which would have to be the case if something like Millian utilitarianism

¹¹ Mill keeps forgetting to insert “in itself” after “desiring”, but it is clear from the argument he is making that it is intrinsic desires he is writing about.

¹² If this is true, the meaning of “desirable” indeed should be taken as an analogue to “visible”. Mill does seem here to claim that happiness is the only thing *capable* of being intrinsically desired. The “worthy of being desired” meaning could then be taken to apply to *instrumental* values, and the worthiness be determined by whether they are instrumental to happiness or not. Is the intrinsic desire thus picked out the “restricted” sense of desire asked for above?

¹³ *Hedonism*, nota bene, not utilitarianism.

should follow. Sidgwick stresses the *multiple* methods available to ethics. Utilitarianism is surely one of them, since happiness is one of preciously few plausible candidates for bearing ultimate value (perfection/excellence is the other), but intuitionism is another. But intuitionism support maxims and knowledge of moral principles, not knowledge of final ends. Utilitarianism and intuitionism, then, supplement each other.

Normativity, Sidgwick claims, is probably not resolvable into simpler notions, “it can only be made clearer by determining as precisely as possible its relation to other notions with which it is connected in ordinary thought, especially to those with which it is liable to be confounded” (p 33).

He starts out by noticing some things about ethical judgments. Emotivist accounts can not catch their meaning since “the peculiar emotion of moral approbation is, in my experience, inseparable bound up with the conviction, implicit or explicit, that the conduct approved is ‘really’ right – *i.e.* that it cannot, without error, be disapproved by any other mind”.

Mill, as we remember, claimed that to desire a thing, and to find it pleasant where to modes of stating the same psychological fact. Sidgwick refutes this, but concedes that “pleasure is a kind of feeling which stimulates the will to actions tending to sustain or produce it, - to sustain it, if actually present, and to produce it, if it be only represented in idea” and that it “seems convenient to call the felt volitional stimulus in the two cases respectively Desire and Aversion”. (42-3) *But –granting that pleasures normally excite desire – it still does not seem to me that I judge pleasures to be greater and less exactly in proportion as they stimulate the will to actions tending to sustain them.*” (p 125-6, my emphasis). Here again, an example of how Sidgwick demonstrate that the evidence given by Mill is not enough to establish the *utilitarianism* Mill wanted to defend. Sidgwick is a bit hazy on this point, though, and occasionally seem to say that when a kind of pleasure is preferred, though less pleasant, something other than the state of pleasure is considered. The following quote is enlightening in this respect: “...if we take the definition of pleasure just given – that it is the kind of feeling which we apprehend to be desirable or preferable – it seems to be a contradiction in terms to say that the less pleasant feeling can ever be thought preferable to the more pleasant.” (p 128) This, of course, does not follow. It is *not* a contradiction in terms given *that* definition. The problem with Sidgwick is that he allows preference for a pleasure to vary with the external conditions for that pleasure, but not with internal variation. And I do not see why our preferences should be restricted in scope in this way. Especially not when (egoistic) hedonism is thought to be a *substantial* view.

“The aim of Ethics is to systemise and free from error the apparent cognitions that most men have of the rightness or reasonableness of conduct, whether the conduct be considered as right in itself, or as the means to some end commonly conceived as ultimately reasonable. These cognitions are normally accompanied by emotions of various kinds, known as “moral sentiments”: but an ethical judgment cannot be explained as affirming merely the existence of such a sentiment: indeed it is an essential characteristic of a moral that it is bound up with an apparent cognition of something more than mere feeling.” (p 77) Namely “dictates” or “imperatives”. The intuitive concept of feeling involved and recognized by most is a mere transitory state on the way to what is ethically more interesting: volition.

Sidgwick does make a distinction that Mill “fails” to make: “In the recognition of conduct as ‘right’ is involved an authoritative prescription to do it: but when we have judged conduct to be good, it is not yet clear that we ought to prefer this kind of good to all other good things: some standard for estimating the relative values of different ‘goods’ has still to be sought.” (p 106)

“the general admission that things which are called ‘good’ are productive of pleasure, and that the former quality is inseparable in thought from the latter, does not involve the inference that the common estimates of the goodness of conduct may be fairly taken as estimates of the amount of pleasure resulting from it”. This is, in effect, what I said was wrong with Mill’s account above. And here comes the good part:

“It would seem then, that if we interpret the notion ‘good’ in relation to ‘desire’, we must identify it not with the actually *desired* but rather with the *desirable*: - meaning by ‘desirable’ not necessarily ‘what *ought* to be desired’ but what would be desired, with strength proportioned to the degree of desirability, if it were judged attainable by voluntary action, supposing the desirer to possess a perfect forecast, emotional as well as intellectual, of the state of attainment or fruition”. (p 112) (But this is obviously not what we *mean* with ‘good’, it is what is good in the *substantive*, not the *adjective* sense. And this should be included in this text, so help me god).

Sidgwick notes that the hedonistic paradox is an objection to the *method* of egoistic hedonism, not to the doctrine as such. (I just came to me that the hedonistic paradox is really about the “wrong kinds of reasons”). Actually, it is even hard to conceive of how a pure pursuit for *pleasure* would go, if not by way of some instrument. “Pleasure” might be the answer to the question “why do you seek too realize X?” Pleasure is a permanent “for the sake of which”, and that is why it is so hard to conceive of separate from its conditions. There may be a problem for egoistical hedonism, though. It seems that the “deputy-goals” must be motivated by pleasure, and that the necessity to hide the true goal must be explicitly known in order for this view to be an unambiguous version of *egoism*.

“the question then remains, whether any general theory can be attained of the causes of pleasure and pain so certain and practically applicable that we may by its aid rise above the ambiguities and inconsistencies of common or sectarian opinion, no less than the shortcomings of the empirical-reflective method, and establish the Hedonistic art of life on a thoroughly scientific basis.” (p 160) It does indeed. And this may very well be where Intuitionism kicks in as a way to supplement utilitarianism.

Intuitionism

“(…) we are accustomed to expect from Morality clear and decisive precepts or counsels: and such rules as can be laid down for seeking the individual’s greatest happiness cannot but appear wanting in these qualities.” (p 199) Here we have the outline for the government house utilitarianism of Sidgwick. This is also, mind you, Mill’s standpoint, as we saw above in the statement that morality does not end at establishing final ends, but must provide ways of arriving at them. Sidgwick echoes Butler in that conscience, the moral faculty, may have practical supremacy over self-love, but that self-love is “theoretically” prior to conscience. Intuitionism, here, is concerned with right conduct, not with the good.

Let’s consider another feature in Sidgwick’s reasoning: “...while a man can resolve to aim at any end which he conceives as a possible result of his voluntary action, he cannot simultaneously resolve *not* to aim at any other end which he believes will be promoted by the same action; and if that other end be an object of desire to him, he cannot, while aiming at it, refuse to act from his desire.” (p 203) This doesn’t sound right. If I aim at a consequence perceived to be good and realise that my action will also bring about some other end that I find desirable but that I don’t perceive to be good, I should not necessarily be thought to aim at *that* consequence. Why should I? In this particular action, my desire for that consequence might have no causal relevance whatsoever.

S. wants conscience and self-love to yield the same, or at least consistent, results (he refers to Butler). He then makes a distinction between “formal” and “material” rightness, where formal rightness concerns the agents willingness to do right while material rightness is aimed at when the agent has the right particular effects. It is with material rightness that Sidgwick is mainly concerned.

The question arises if it is better that a man does right believing it to be wrong or wrong believing it to be right. The latter is probably “formally better” while the former is “materially better”.

“I cannot see how the mere ascertainment that certain apparently self-evident judgements have been caused in known and determinate ways, can be in itself a valid ground for distrusting this class of apparent cognitions.” (p 212) Naturally, if I have been tortured, Orwell-style, into believing something (big brother) to be right, this belief should be suspected. But it doesn’t prove it to be wrong.

A rather sneaky argument for utilitarianism surfaces in S. discussion of *justice*. What, he asks, is the relevant class to which justice accrues? What are reasonable claims as to justice? It cannot be based on explicit contracts alone, but must also, but only to some extent, concern tacit understandings¹⁴. But reasonable claims do not arise from any expectations that results from certain behaviour. To determine this one are left floating by the intuitionistic method, and left is only unsupported dogmatism or utilitarianism. And, clearly...

But expectations is not the only foundation of justice, is it?

“...from one point of view, we are disposed to think that the *customary* distributions of rights, goods, and privileges, as well as burdens and pains, is natural and just, and that this ought to be maintained by law, as it usually is: while, from another point of view, we seem to recognise an ideal system of rules of distribution which ought to exist, but perhaps have never yet existed, and we consider laws to be just in proportion as they conform to this ideal. It is the reconciliation between these two views which is the chief problem of political Justice.” (p 273) Justice does seem to be different conceived according to what you are looking for.

“...perhaps...Fitness should rather be regarded as a utilitarian principle of distribution, inevitably limiting the realisation of what is abstractly just, than as a part of the interpretation of Justice proper: and it is with the latter that we are at present concerned.” (p 283)

[“Any strong sentiment, however purely subjective, is apt to transform itself into the semblance of and intuition”]

Intuitionism, as supporting moral common sense, does not live up to the requirements of “self-evidence”. They are found lacking in consistence and in determination. The latter is, of course, where utilitarianism can help. But S. thinks that he establishes inconsistency just by showing moral intuitions to yield *conflicting* results. This, of course, is not really honest, but the point is taken. And articulated “Nothing that I have said even tends to show that we have not distinct moral impulses, claiming authority over all others, and prescribing or forbidding kinds of conduct as to which there is a rough general agreement, at least among educated persons of the same age and country. It is only maintained that the objects of these impulses do not admit of being scientifically determined by any reflective analysis of common sense.” (p 360)

His point about the list of motives (p 372) is that there is ultimately a “master value” (not his term) with reference to which we decide between, not just actions, but motives to.

¹⁴ 120 years later, Scanlon makes a similar argument, but without noticing the difficulties

S. criticizes Mill for not having established the greatest general happiness as desirable, but only, if that, the desirability of individual happiness. "...there is a gap in the expressed argument, which can, I think, only be filled by some such proposition as that which I have above tried to exhibit as the intuition of Rational Benevolence." (p 388) Mill, as we have seen, did try to close this gap, but not with the 'desire-proof'.

"Shall we then say that Ultimate Good is Good or Desirable conscious or sentient Life – of which Virtuous action is one element, but not the sole constituent? This seems in harmony with Common Sense; and the fact that particular virtues and talents and gifts are largely valued as means to ulterior good does not necessarily prevent us from regarding their exercise as also an element of Ultimate Good". (p 395-6)

"But this is not because the mere existence of human organisms, even if prolonged to eternity, appears to me in any way desirable; it is only assumed to be so because it is supposed to be accompanied by Consciousness on the whole desirable; it is therefore this Desirable Consciousness which we must regard as ultimate Good." (p 397)

Desirable consciousness involves the consciousness of virtue (compare with Mill). *Is it happiness/pleasure?* S. argues that it is, understood as desirable *feeling*.

Truth, as being in the relation between a conscious subject and an object is not, S. argues, intrinsically desirable. It's only the relation's and the object's conduciveness to happiness that matters. This Sidgwick takes to be obvious on reflection. The "Ultimate Good" chapter wonderfully begs the question. No other candidate remains standing.

Utilitarianism.

Utilitarianism, as an ethical doctrine, is not necessarily connected with the psychological theory that the moral sentiments are in some way derived from experiences of non-moral pleasures and pains. In general, Sidgwick notes that the genealogy of morality is not really relevant for its content or validity.

The *formal principle* of equality allows for Sidgwickian utilitarian to have opinions on the distribution of a given amount of happiness. Maximising Hedonism, as we have seen, is not the final word when it comes to right conduct. .

"if Utilitarianism is to be *proved* to a man who already holds some other moral principles (...) it would seem that the process must be one which establishes a conclusion actually *superior* in validity to the premises from which it starts." (p 419)

"...it must be borne in mind that Utilitarianism is not concerned to prove the absolute coincidence in results of the Intuitional and Utilitarian methods." P 425 It's enough to point out that they tend to co-refer, and that a natural transition from common sense morality to U. is possible.

"..when the same rule is interpreted somewhat differently by different persons, each naturally supports his view by urging its Utility, however strongly he may maintain the rule to be self-evident and known *a priori*" (p 426, important section, I believe).

S. accounts for praiseworthiness by appeal to the utility of praise, not the utility of that which is praised (but, of course, these things tend to be connected).

"We may now observe that this hypothesis of "unconscious Utilitarianism" explains the different relative importance attached to particular virtues by different classes of human beings, and the different emphasis with which the same virtue is inculcated on these different classes by mankind generally." (p 454)

“Indeed from the considerations that we have just surveyed it is but a short and easy step to the conclusion that in the Morality of Common Sense we have ready to hand a body of Utilitarian doctrine; that the “rules of morality for the multitude” are to be regarded as “positive beliefs of mankind as to the effects of actions on their happiness,” so that the apparent first principles of Common Sense may be accepted as the “middle axioms” of Utilitarian method; direct reference being only made to utilitarian considerations, in order to settle points upon which the verdict of Common Sense is found to be obscure and conflicting. On this view the traditional controversy between the advocates of Virtue and the advocates of Happiness would seem to be at length harmoniously settled.” (p 461)

The “Government House” variety of utilitarianism is being introduced in chapter V, p 475
Motivated in part by the fact that to establish a new rule, more conducive to happiness, often implies the breaking of other rules generally, leading to worse results on the whole. Calls desperately for speculation, but this is how he argues.

“...the opinion that secrecy may render an action right which would not otherwise be so should itself be kept comparatively secret; and similarly it seems expedient that the doctrine that esoteric morality is expedient should itself be kept esoteric.” (p 490) Utilitarian motivation of moral disagreements.

Now, the mutual relations of the three methods:

Intuitionism seems (I’m not altogether sure about this one) to provide a transition from egotistical to general utilitarianism. The transition, S. admits, is only by a stretch of language properly called a “proof”. There is no logical transition, in other words. If there is a tug-of-war between egoistic goals and rational benevolence, S. bets is on egoism. He does take seriously the distinction between persons (and thus answers Rawls before Rawls was even born). “The inseparable connexion between Utilitarian Duty and the greatest happiness of the individual that confirms to it cannot be satisfactorily demonstrated on empirical grounds” (p 503)