Bentham's Binary Form of Maximizing Utilitarianism*

Johan E. Gustafsson[†]

ABSTRACT. Jeremy Bentham is often interpreted as defending a satisficing, rather than maximizing, version of utilitarianism, where an act is right as long as it produces more pleasure than pain. This lack of maximization is surprising given Bentham's maximizing slogan 'the greatest happiness of the greatest number'. Against the satisficing interpretation, I argue that Bentham consistently defends a maximizing version of utilitarianism, where an act's consequences are compared to those of not performing the act. I show that following this version of utilitarianism requires that one realizes the greatest happiness for all affected individuals.

A puzzling feature of Jeremy Bentham's statement of utilitarianism is its apparent lack of maximization. This is puzzling because, as is well-known, Bentham popularized the maximizing slogan 'the greatest happiness of the greatest number' and he coined the terms 'maximize' and 'maximization'.[1] Marcus G. Singer, for example, writes:

> There is no hint [in Bentham's statement of utilitarianism] that one *ought*, or is obligated, to do that act which, of all the available alternatives, would produce the absolutely best consequences. (Singer, 1977, p. 68)

Singer (1977, pp. 67–8) and Anthony Quinton (1973, pp. 2–3) both argue that Bentham defends a satisficing criterion of rightness, rather than a maximizing one.[2] Others, such as Michael Slote (1984, pp. 153–4) and Ger- p. 88

1 Bentham (1977, p. 393) popularized the slogan but was not the first to use it; see footnote 32. For his first use of 'maximize' and 'maximization', see OED2, vol. 9, p. 497.

2 Rosen (2003, pp. 226–7) does not accept the satisficing interpretation but takes Bentham's talk about maximizing happiness to merely make the distributive point 'that

ald J. Postema (2002, p. xv), interpret Bentham as defending satisficing in *An Introduction to the Principles of Morals and Legislation* but as defending maximizing elsewhere. In this paper, I shall argue that Bentham consistently defends maximizing rather than satisficing but his maximizing has a peculiar, binary form.

We shall see that Bentham's version of utilitarianism differs structurally in many ways from the more standard version of utilitarianism presented in G. E. Moore's *Ethics*. According to Moore's maximizing version of utilitarianism,

> A voluntary action is right, whenever and only when no other action possible to the agent under the circumstances would have caused more pleasure; in all other cases, it is wrong. (Moore, 1912, p. 31)

Here, on what we can call *Moorean maximization*, an act is right if and only if it has at least as good consequences as every act the agent could have performed instead in the situation. Bentham, I shall argue, does

---

happiness should be extended to the greatest possible number of people'. Postema (1998) objects to Rosen's argument, yet favours an equality sensitive interpretation based on the following paragraph:

> I recognize, as the *all-comprehensive*, and only right and proper end of Government, the greatest happiness of the members of the community in question: the greatest happiness—of all of them, without exception, in so far as possible: the greatest happiness of the greatest number of them, on every occasion on which the nature of the case renders the provision of an equal quantity of happiness for every one of them impossible: it being rendered so, by its being matter of necessity, to make sacrifice of a portion of the happiness of a few, to the greater happiness of the rest. (Bentham, 1831, p. 7)

This paragraph says, according to Postema (1998, p. 157), that we should aim for an equal distribution of happiness and that we should only aim for the greatest happiness if an equal distribution cannot be achieved. The paragraph says, I think, the opposite, namely, that we should aim for the greatest happiness even if, in order to realize the greatest happiness, the happiness may sometimes need to be unequally distributed. The key is the relation between (i) the part starting 'the greatest happiness of the greatest…' and (ii) the part starting 'it being…' Part (ii) clarifies that the equal distribution in (i) is 'impossible' only in the sense that it cannot be achieved if one is to achieve the overall greatest happiness. With this in mind, (i) does not say that one should aim for the greatest happiness if an equal distribution cannot be achieved. Rather (i) clarifies that the greatest happiness should be our end even when it rules out an unequal distribution of happiness. This clarification need not have been obvious to Bentham's contemporaries, because early utilitarians typically favoured egalitarian policies.

not defend Moorean maximization; he defends a version of utilitarianism with a non-standard, binary form of maximization, where an act is only compared to the negative act of not doing that act—rather than to every act the agent could have performed instead. He holds that an act ought to be done if and only if it produces a greater balance of pleasure over pain than the negative act of not doing that act. As I will explain, this binary act utilitarianism still requires that, in order to avoid wrongdoing, one must follow the course of action that would produce the greatest happiness among all feasible courses of action.[3] Hence Bentham does defend a maximizing, rather than satisficing, version of utilitarianism.

1. The principle of utility and its connection to right and wrong

Why would anyone think that Bentham defends satisficing? To understand this, we need to explore how rightness in Bentham's ethics is connected to pain and pleasure via his principle of utility. In the opening chapter of *An Introduction*, Bentham presents the principle as follows:[4]

(B1)   An action then may be said to be conformable to the
         principle of utility, or, for shortness sake, to utility,
         (meaning with respect to the community at large) when
         the tendency it has to augment the happiness of the
         community is greater than any it has to diminish it.
         (Bentham, 1970, pp. 12–13)

The principle of utility is Bentham's standard of right and wrong. He writes:

(B2)   Of an action that is conformable to the principle of utility,
         one may always say either that it is one that ought to be
         done, or at least that it is not one that ought not to be done.
         One may say also, that it is right it should be done; at least
         that it is not wrong it should be done: that it is a right
         action; at least that it is not a wrong action. When thus
         interpreted, the words *ought*, and *right* and *wrong*, and
         others of that stamp, have a meaning: when otherwise, they
         have none.   (Bentham, 1970, p. 13)

---

[3] The two approaches will, however, differ at times about the rightness and wrongness of individual acts as will be explained in Section 8.

[4] The numbers prefaced by 'B' refer to Bentham quotes.

Yet Bentham's criterion in (B2) is somewhat obscured by an unfortunate use of disjunctions (which we will return to in Section 5). He presents the principle of utility and its relation to right and wrong a little bit clearer in *Rationale of Judicial Evidence*:

(B3)   In morals, as in legislation, the *principle of utility* is that which holds up to view as the only sources and tests of right and wrong, human suffering and enjoyment—pain and pleasure. It is by experience, and by that alone, that the tendency of human conduct, in all its modifications, to give birth to pain and pleasure, is brought to view: it is by reference to experience, and to that standard alone, that the tendency of any such modifications to produce more pleasure than pain, and consequently to be *right*, or more pain than pleasure, and consequently to be *wrong*, is made known and demonstrated.   (Bentham, 1827, vol. 1, p. 120; 1838–1843, vol. 6, p. 238)

From (B2) and (B3), it seems that Bentham holds the following:

(4)   An act is right if it conforms to the principle of utility.

(5)   An act is wrong if it deviates from the principle of utility.

This reading is also supported by Bentham's more general view, put forward in *Of the Limits of the Penal Branch of Jurisprudence*, that

(B6)   right is the conformity to a rule, wrong the deviation from it   (Bentham, 2010, p. 186)

From (B1), we have that an act conforms to the principle of utility if it augments the happiness of the community. In (B1), Bentham does not say when an act deviates from the principle of utility. But (B1), (B3), and (B6) combined suggest that Bentham holds that an act deviates from the principle of utility if it lessens the happiness of the community. Here, happiness is simply what is measured by Bentham's felicific calculus.[5] Understanding the felicific calculus is, I think, the key to understanding the structure of Bentham's utilitarianism. In *An Introduction*, Bentham provides the following algorithm for this calculus—which, since it will be crucial for our discussion, is worth quoting in full:[6]

---

[5] Bentham provides a somewhat ambiguous definition of happiness in (B49).
[6] The value of an individual pleasure or pain depends, according to Bentham (1970, p. 38), on four circumstances: intensity, duration, certainty (i.e. probability), and

(B7)   Begin with any one person of those whose interests seem most immediately to be affected by it: and take an account,

  1. Of the value of each distinguishable *pleasure* which appears to be produced by it in the *first* instance.

  2. Of the value of each *pain* which appears to be produced by it in the *first* instance.

  3. Of the value of each pleasure which appears to be produced by it *after* the first. This constitutes the *fecundity* of the first *pleasure* and the *impurity* of the first *pain*.

  4. Of the value of each *pain* which appears to be produced by it after the first. This constitutes the *fecundity* of the first *pain*, and the *impurity* of the first pleasure.

  5. Sum up all the values of all the *pleasures* on the one side, and those of all the *pains* on the other. The balance, if it be on the side of pleasure, will give the *good* tendency of the act upon the whole, with respect to the interests of that *individual* person; if on the side of pain, the *bad* tendency of it upon the whole.

  6. Take an account of the *number* of persons whose interests appear to be concerned; and repeat the above process with respect to each. *Sum up* the numbers expressive of the degrees of *good* tendency, which the act has, with respect to each individual, in regard to whom the tendency of it is *good* upon the whole: do this again with respect to each individual, in regard to whom the tendency of it is *bad* upon the whole. Take the *balance*; which, if on the side of *pleasure*, will give the general *good tendency* of the act, with respect to the total number or community of individuals concerned; if on the side of pain, the general

---

propinquity. The first three circumstances seems fairly straightforward. Bentham's (1998, p. 251–2) discussion in *Codification Proposal* suggests that we calculate the value of a pain or pleasure x by multiplying these along with a discounting function, representing propinquity:

$$value(x) = intensity(x) * duration(x) * probability(x) * e^{-rt}$$

where t is the number of years into the future the pain or pleasure would occur and r is the discount rate. Bentham (1998, p. 251) explains propinquity with an example of continuous compounding with an interest rate at 5 %. Yet Bentham has not told us what the right discount rate is, except that it is non-zero. The 5 % rate seems to be a merely illustrative example. It was probably picked because, in 1821 when Bentham wrote that passage, the discount rate of the Bank of England had been at 5 % for over a hundred years; see Clapham (1944, vol. 1, p. 299).

> *evil tendency*, with respect to the same community.
> (Bentham, 1970, pp. 39–40 with corrected text)

The basis of the satisficing interpretation is the following straightforward reading of the felicific calculus in (B7):[7]

(8*)   An act conforms to the principle of utility if the sum total of the values of the pleasures it appears to produce is greater than the sum total of the values of the pains it appears to produce.

(9*)   An act deviates from the principle of utility if the sum total of the values of the pains it appears to produce is greater than the sum total of the values of the pleasures it appears to produce.

Accordingly, Quinton, Singer, Slote and Postema take Bentham to hold the following satisficing version of utilitarianism, combining (8*) and (9*) with (4) and (5) respectively:[8]

(10*) An act is right if the sum total of the values of the pleasures it appears to produce is greater than the sum total of the values of the pains it appears to produce.

(11*) An act is wrong if the sum total of the values of the pains it appears to produce is greater than the sum total of the values of the pleasures it appears to produce.

To see that this is a satisficing version of utilitarianism, consider the following choice about whether or not to do an act a:

Case 1

Act	Pleasures	Pains
a	2	1
not-a	3	0

In this table, the 'Pleasures' column represents the sum total of the values of the pleasures the acts appear to produce, and the 'Pains' column represents the sum total of the values of the pains that the acts appear to produce. Even though a yields less pleasure and more pain than not-a,

---

[7] The numbered claims with asterisks are not part of my interpretation of Bentham.

[8] Quinton (1973, pp. 2–3), Singer (1977, pp. 67–8), Slote (1984, p. 154) and Postema (2002, p. xv).

we still have that *a* is right according to (10*).[9] And, according to (8*), act *a* conforms to the principle of utility, because the balance of pleasures and pains for *a* is on the side of pleasures, and thus *a* has an overall good tendency. Still, not-*a* has an even better tendency.

In Case 1, not-*a* is an example of a negative act, which consists in forbearing from doing *a*. According to Bentham, such forbearances are also acts. He explains the distinction between negative and positive acts in *An Introduction*:

(B12)   By positive are meant such as consist in motion or exertion: by negative, such as consist in keeping at rest; that is, in forbearing to move or exert one's self in such and such circumstances. Thus, to strike is a positive act: not to strike on a certain occasion, a negative one. Positive acts are styled also of commission; negative, acts of omission or forbearance.[10]   (Bentham, 1970, p. 75)

We shall return to the role of negative acts in Bentham's utilitarianism in Section 8.

2. Prohibition dilemmas and the logic of the will

To see why the satisficing interpretation does not add up, we need to get clearer on Bentham's views on the logic and meaning of deontic concepts. Deontic concepts are, as we shall see, closely linked to commands and prohibitions through their shared connection to the approval and disapproval of the legislator. This link allows us to infer parts of Bentham's deontic logic from his logic of commands, that is, from his logic of the will. Writing in *A Comment on the Commentaries* about the meaning of 'right' and 'wrong', Bentham claims that

---

[9] Shaver (2013, p. 295) similarly points out that, in a case where one can produce one unit of pleasure or produce a million units of pleasure, (10*) yields that it is right to just produce one unit. He therefore thinks that the (10*) interpretation attributes an overly silly view to Bentham.

[10] In *Constitutional Code*, Bentham (1838–1843, vol. 9., p. 480) provides the following definition:

a NEGATIVE act, consisting in *abstinence*, from the performance of some positive act.

As Bentham (1970, p. 76; 2010, p. 253) points out, whether an act is positive or negative depends on how it is described: 'get drunk'/'not stay sober' or 'not get drunk'/'stay sober'. A negative act is merely an act described as not doing some act. Every positive act has a 'correspondent' negative act; see (B17).

(B13)   Were I to be asked what it is I mean when I call an action a
          right one, I should answer very readily: neither more nor
          less than, an action I *approve* of: and so of a *wrong* action,
          an action I *disapprove* of.   (Bentham, 1977, p. 53)

Hence, according to Bentham, rightness and wrongness correspond to approval and disapproval respectively. Similarly, Bentham (1983, p. 207) claims in *Deontology* that, like 'right' and 'wrong', 'ought' and 'ought not' correspond to the approval or disapproval—that is, to the decided will— of the legislator. In the chapter 'Aspects of a Law' of *Of the Limits*, Bentham explains, in turn, how the will relates to commands:

(B14)   To begin with the case where the will is decided. The wish
          must either be that the act should be performed, or that it
          should not be performed: in the former case, the mandate
          is what, in the most confined sense of the word, is termed a
          command: in the latter, a *prohibition*. The will being
          undecided, there is no wish, that can be operative or
          efficient, to express: all that the mandate can express is the
          negation of one or other of the two operative kinds of
          mandates: it is accordingly either what may be termed a
          non-command, or else a *non-prohibition*, that is, to use the
          common language, a *permission*.   (Bentham, 2010, p. 252)

So an act that is right or ought to be done is an act that is approved of and, in turn, commanded. And an act that is wrong or ought not to be done is an act that is prohibited. It might seem strange to treat both acts that are right and acts that ought to be done as commanded acts.[11] But, as we shall see later in (B20), Bentham regards an act's being right as equivalent to its being obligatory. Note also that—in addition to rightness, corresponding to approval, and wrongness, corresponding to disapproval—the second half of (B14) describes a third deontic category corresponding to the un- decided, indifferent will. This third category of permitted acts that are neither obligatory nor wrong would, in contemporary usage, be consid- ered to be right acts, but Bentham treats them as an additional, alterna- tive deontic category. We shall return to the category of indifferent acts in Section 6.

[11] This conflicts with Moore's (1912, p. 36) more standard view, that 'to say that an action is right is *not* to imply that it ought to be done'.

8

Through this link between deontic terms on the one hand and commands and prohibitions on the other, we can learn something about Bentham's views on deontic logic from the following part of his logic of the will, presented later in 'Aspects of a Law':

(B15)   1. The act may be commanded: it is, then, left unprohibited: and it is not prohibited, nor left uncommanded. 2. It may be prohibited: it is, then left uncommanded: and it is not commanded, nor permitted (that is left unprohibited). 3. It may be left uncommanded: it is, then, not commanded: but it may be either prohibited or permitted: yet so as that, if it be in the one case, it is not in the other. 4. It may be permitted: it, then, is not prohibited: but it may be either commanded or left uncommanded: yet so as that, if it be in the one case, it is not in the other.[12]   (Bentham, 2010, pp. 139–40)

Given the link we found, point 1 in (B15) supports that Bentham holds that

(16)   If an act x ought to be done, then x is not wrong.

This reading follows Georg Henrik von Wright (1981, p. 5), who dubbed (16) 'Bentham's Law'.

In 'Aspects of a Law', Bentham also explains the relation between commands of negative and positive acts:

(B17)   A negative aspect towards a positive act is equipollent to an affirmative aspect towards the correspondent negative act. To will during a given period the act of not carrying arms shall be performed is to will that throughout that period the *continued act* or *habit* of not carrying arms should be persevered in. It is in this way that a command may wear the form of a prohibition: and a prohibition that of a command. The law which prohibits the mother from starving her child commands her to take care that it be fed. The one may be at pleasure translated or *converted* into the other.   (Bentham, 2010, pp. 252–3)

In other words, a prohibition of an act translates into a command of the corresponding negative act, and vice versa, which in deontic terms means that

---

[12] See also Bentham's (2010, p. 254) almost identical later draft of this passage.

(18) An act x ought to be done if and only if not-x is wrong.

Which entails that

(19) If an act x is wrong, then not-x ought to be done.

This reading is backed up by a passage from *The Elements of the Art of Packing*, where Bentham complains about

(B20) propositions, of the cast termed by logicians *identical*, fit only for mouths and pens of drivellers: propositions, which neither conveying instruction not imposing obligation, leave every thing exactly as they find it: propositions declaring that what is *right ought* to be done, and what is *wrong* ought *not* to be done, and so forth. (Bentham, 1821, p. 247; 1838–1843, vol. 5, p. 176)

Since Bentham takes it to be logically true that 'what is *wrong* ought *not* to be done', he must accept (19).

Consider then a situation with the following available acts:

Case 2

Act	Pleasures	Pains
a	0	1
not-a	0	2

Since the balance is on the side of pains for both a and not-a, we have that both a and not-a are wrong according to (11*). Hence prohibition dilemmas, where all available acts are wrong, would be possible on the satisficing interpretation of Bentham's utilitarianism. Regarding cases like Case 2, Quinton claims that

it is possible that all the available alternatives would detract from the general happiness to some extent. It would seem congruous with the spirit of the general happiness principle to choose that action which detracts least from it. But this yields a paradox: as detracting from the general happiness it ought not to be done, as detracting less than any other possibility it ought to be done. (Quinton, 1973, p. 3)

While this seems paradoxical, there is a more direct problem for the satisficing interpretation. According to (11*), act a is wrong. But, if a is wrong, then, by (19), we have that not-a ought to be done. And, if not-a ought to

10

be done, then, by (16), we have that not-*a* is not wrong. Yet, according to (11*), not-*a* is wrong. Hence we have a contradiction. Moreover, the result of (11*) in Case 2, that one is bound to deviate from the principle of utility in that case, does not fit with Bentham's remark in *An Introduction* that

(B21)   The principle of utility is capable of being consistently pursued; and it is but tautology to say, that the more consistently it is pursued, the better it must ever be for human-kind.   (Bentham, 1970, p. 21)

On top of that, (B21) does not fit with the implications of (10*) in Case 1. In that case, (10*) yields that it is right to do *a*, yet doing *a* would presumably be worse rather than better for humankind than doing not-*a*. Since these problems are both obvious and serious, it would be too uncharitable to interpret Bentham in this manner if there is a plausible interpretation that avoids these problems.

3. Forbearing pains and pleasures

There is a further problem for the satisficing interpretation in Case 2. The following two passages strongly suggests that, in Case 2, act *a* would, according to Bentham, be right rather than wrong, contradicting (11*). In *An Introduction*, Bentham claims that

(B22)   **1**. The general object which all laws have, or ought to have, in common, is to augment the total happiness of the community; and therefore, in the first place, to exclude, as far as may be, every thing that tends to subtract from that happiness: in other words, to exclude mischief.
2. But all punishment is mischief: all punishment in itself is evil. Upon the principle of utility, if it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil.[13]   (Bentham, 1970, p. 158)

It is hard to make sense of this passage given (10*) and (11*). If the relevant comparison for whether an act is right or wrong is whether the act produces more pleasure than pain, it is hard to understand how the fact that an act excludes some evil—that is, some pain—could be relevant for

---

[13] See also Bentham's (1998, p. 251na) medicine example in *Codification Proposal*.

whether an act is right or wrong.[14] Consider Case 2, which is structured like the examples Bentham describes in (B22), where the likely effect of act *a*—which could, for example, be to take one's medicine—is mainly that it averts some pain. In Case 2, (11*) yields that *a* is wrong even though, if *a* were performed, only a little pain would occur rather than the more significant pain that would occur if not-*a* were performed.[15]

4. Averted pains count as pleasures; averted pleasures count as pains

Bentham's comments in (B22) on cases like Case 2 make clear that the principle of utility is supposed to take into account preventions of pain and pleasure. He also makes this clear in his discussion of motives, where he puts forward a general account of the tendency of things that are themselves neither pains nor pleasures. This account depends, among other things, on preventions of pain and pleasure:

(B23)   With respect to goodness and badness, as it is with every
           thing else that is not itself either pain or pleasure, so is it
           with motives. If they are good or bad, it is only on account
           of their effects: good, on account of their tendency to
           produce pleasure, or avert pain: bad, on account of their
           tendency to produce pain, or avert pleasure.   (Bentham,
           1970, p. 100)

---

[14] Hence we need to reject Schofield's (2006, p. 40) interpretation that whether an act is right or wrong is determined by 'the quantity of pleasure and pain which had been brought into existence by the action in question.' Unless, of course, Schofield follows Bentham's non-standard usage.

[15] One might object that, if Slote and Postema only interpret Bentham as defending satisficing in *An Introduction*, it is unfair to rely on textual evidence from other works.

But, first, note that the logical objection based on (16) and (19) in Case 2 was mainly based on passages from *Of the Limits*, which was conceived as a continuation of *An Introduction* and written before the printing of *An Introduction*; see Bentham (2010, p. xi). Hence it seems that these passages are relevant for an interpretation of *An Introduction*.

Second, note that (B21) on its own is inconsistent with the satisficing view in Case 2. And, for the objection about forbearing pains and pleasures, we need only rely on (B22). Finally, (B45) seems to directly contradict the satisficing view. Yet (B21), (B22), and (B45) are all from *An Introduction*.

Third, it seems to me that Bentham did not change his fundamental ethical views during his publishing career. The changes between different epochs that Bentham (1983, pp. 326–8) reports in the short version of 'Article on Utilitarianism' seem to mainly concern changes in presentation rather than any substantial changes to his ethical theory.

The felicific calculus purports to calculate the good or bad tendency of acts. Since acts are not themselves pains or pleasures, their value, according to (B23), partly depends on avoided pains and pleasures. So the felicific calculus should also take preventions into account.

That the principle of utility takes preventions of pain and pleasure into account suggests that there is a comparative, counterfactual element involved in the principle that is not present in (8*) and (9*). To get clearer on this comparative element of the principle of utility, we need to pay attention to what Bentham thinks should be taken into account when the principle is applied. In *Deontology*, he makes clear that, when one applies the principle of utility, one should compare the balance of the pains and pleasures if the act were done with the balance of the pains and pleasures if the act were not done:

(B24)   On the occasion of any proposed act, to make application of the principle of utility is to take one account of the feelings of the two opposite kinds—of the pleasures of all sorts on the one side, of the pains of all sorts on the other side—which, in all breasts that seem likely to be in any way affected by it, seem liable and likely, in the two opposite cases of the act's being done and of its being left undone, to take place.   (Bentham, 1983, p. 168)

From (B24), it is clear that applying the principle of utility to assess an act involves comparing what would happen were one to perform that act with what would happen were one to forbear from performing it. This raises the puzzle of how to understand the algorithm in (B7), which at first sight looks like it only takes into account what would happen if the act in question were performed.

That Bentham simply forgot about preventions when he wrote the algorithm in (B7) does not seem credible. First, note that the short chapter presenting the felicific calculus in (B7) starts as follows:

(B25)   Pleasures then, and the avoidance of pains, are the *ends* which the legislator has in view: it behoves him therefore to understand their *value*.   (Bentham, 1970, p. 38)

Given this initial emphasis on avoiding pains, it would be strange if the calculus presented in the same short chapter did not account for the avoidance of pains. Second, note that in the penultimate paragraph of that chapter, commenting on the usefulness of the just presented calculus, Bentham states that

(B26)   An article of property, an estate in land, for instance, is
         valuable, on what account? On account of the pleasures of
         all kinds which it enables a man to produce, and what
         comes to the same thing the pains of all kinds which it
         enables him to avert.   (Bentham, 1970, p. 40)

Hence Bentham emphasizes averting pains just before and just after the
felicific calculus. In light of this, the forgetfulness solution is implausible.

But (B26) provides a further clue to the puzzle: In (B26) and through-
out *An Introduction*, Bentham stresses that preventions of pains 'comes
to the same thing' as pleasures in his theory and, similarly, preventions of
pleasures 'comes to the same thing' as pains:

(B27)   By utility is meant that property in any object, whereby it
         tends to produce benefit, advantage, pleasure, good, or
         happiness, (all this in the present case comes to the same
         thing) or (what comes again to the same thing) to prevent
         the happening of mischief, pain, evil, or unhappiness to the
         party whose interest is considered   (Bentham, 1970, p. 12)

(B28)   Having taken a general view of these two grand objects (viz.
         pleasure, and what comes to the same thing, immunity
         from pain) in the character of *final* causes; it will be
         necessary to take a view of pleasure and pain itself, in the
         character of *efficient* causes or means.   (Bentham, 1970,
         p. 34)

(B29)   For danger is nothing but the chance of pain, or, what
         comes to the same thing, of loss of pleasure.   (Bentham,
         1970, p. 144)

This suggests that preventions of pains should be counted as pleasures in
the algorithm for the felicific calculus and, analogously, that preventions
of pleasures should be counted as pains. If we read (B7) with this in mind,
the felicific calculus takes preventions of pain and pleasure into account
and makes a lot more sense. Hence we have a solution to our puzzle.

Thus, to summarize the textual evidence, we have: (i), from Section 3,
that Bentham thinks that one should avoid pain; (ii), from (B24), that he
holds that applying the principle of utility takes in to account both the
effects of an act's being done and those of its not being done; (iii), from
(B23), that he holds that the tendency of things in general should take pre-
ventions into account; (iv), from (B27)–(B29), that he regards throughout

An Introduction that producing pleasure (pain) comes to the same thing as averting pain (pleasure); and (v) that he emphasizes the avoidance of pains both in the introduction to felicific calculus—(B25)—and in the discussion following it—(B26).

In the light of (i)–(v), there is, I think, strong textual evidence that any pains that appear to be averted by the act should be included with the pleasures that appear to be produced by it in steps 1 and 3 of the algorithm in (B7) and, similarly, that any pleasures that appear to be averted by the act should be included with the pains that appear to be produced by it in steps 2 and 4.[16] This provides, combined with (B1), the following reading of the calculus, where like in the rest of my interpretations 'pain' and 'pleasure' are used in their standard senses that do not include averted pains and pleasures:[17]

(30) An act conforms to the principle of utility if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *greater than* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.[18]

[16] If 'pains' and 'pleasures' are used in Bentham's inclusive manner in (10*) and (11*), then (10*) and (11*) become a version of the binary view and hence, by the argument in Section 8, a form of maximizing rather than satisficing.

[17] Regarding *Rationale of Judicial Evidence*, this reading of 'pain' and 'pleasure' is plausible in (B3) given his (1827, vol. 1, p. 192n; 1838–1843, vol. 6, p. 259n2) claim that

> Whatever act affords any the minutest particle of satisfaction, of pleasure,
> or removes or prevents any the least particle of pain, is, in so far, good.

See also Bentham's (1827, vol. 1, p. 185n; 1838–1843, vol. 6, p. 257n1) definitions of 'good' and 'evil', which take preventions of pains or pleasure into account. Hence it seems that (B3) should be read so that preventions are taken into account, and, given my reading of 'pain' and 'pleasure', it would be.

[18] This way of understanding Bentham's calculus suggests that he was a total rather than average utilitarian. In *Institute of Political Economy*, Bentham (1952–1954, vol. 3, p. 318) makes clear, contrary to what Peter Singer (1976, pp. 81–82) claims, that he was aware of the distinction between total and average views:

> Opulence, though so nearly of kin to wealth, or rather for that very reason,
> requires to be distinguished from it: opulence is *relative* wealth, relation
> being had to population: it is the ratio of wealth to population. Quantity
> of wealth being given, the degree of opulence is therefore not directly, but
> inversely, as the population, i.e. as the degree of populousness—as the num-
> ber of those who are to share in it: the fewer the shares, the larger is each
> one's share.

One passage in *Pannomial Fragments* supports a total-utilitarian interpretation. Ben-

(31)   An act deviates from the principle of utility if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *less than* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

Combining (30) and (31) with (4) and (5) respectively, we get the following sufficient criteria for right and wrong:

(32)   An act is right if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *greater than* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

(33)   An act is wrong if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *lesser than* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

---

tham (1838–1843, vol. 3, p. 228) argues that it is better to distribute wealth so that 10,000 already existing rich people get richer than to use the money to give existence to some very poor people:

> a greater addition to the aggregate quantity of happiness would be made by dividing among the first 10,000 the whole additional quantity of wealth, than by making any addition to the number of persons brought into existence.

Here, Bentham uses money as a more easily measurable substitute for happiness, as he often did—see, for example, Bentham (1998, p. 252) and Schofield (2006, p. 43). So far, the passage is consistent with average utilitarianism, but Bentham (1838–1843, vol. 3, p. 228) continues:

> For, supposing the whole 10,000 having each of them the minimum of the matter of subsistence on any given day,—the next day, in consequence of some accident, they might cease to have it, and in consequence cease to have existence: whereas, if of this same 10,000, some had, in addition to his minimum of the matter of subsistence, particles one or more of the matter of abundance, here would be a correspondent mass of the matter of wealth, capable of being by the legislator so disposed of as to be made to constitute the matter of subsistence to those who, otherwise being without subsistence, would soon be without existence.

Hence it is because he holds that it is likely that the poor people who would live at the minimum of subsistence would soon suffer some accident and pass away, not because it would lower the average level of happiness. This rationalization makes sense on the total view but not on the average view, where the average would be lowered by the addition of the poor even if they managed to stay alive.

5. Bentham's account of obligation

Having found these sufficient criteria for right and wrong, let us return to Bentham's views on obligation. Commenting on (B2), Quinton questions whether Bentham has any positive account of obligation.[19] Bentham does, however, present a stronger account of obligation in the penultimate chapter of *An Introduction*:

(B34)   Every act which promises to be beneficial upon the whole
             to the community (himself included) each individual
             ought to perform of himself   (Bentham, 1970, p. 285)

Hence it seems that Bentham holds that

(35)   An act ought to be done if it conforms to the principle of utility.[20]

But, if (35) is what Bentham had in mind, why did he weaken his claim in (B2) by adding the second disjunct in 'one may always say either that it is one that ought to be done, or at least that it is not one that ought not to be done'? One might think that Bentham meant that some acts that conform to the principle of utility are obligatory but other acts that conform are not obligatory and merely not obligatory to avoid. But, if so, (B2) would not provide the more or less complete meaning of *ought*, since we would also need to know which acts falls into the first category and which falls into the second. And that would conflict with the last sentence of (B2), which suggests that this paragraph provides the meaning of *ought*. Instead, I think we should read the quote as saying that, if an act conforms to the principle of utility, it would always be correct to say each of the following: (i) the act ought to be done and (ii) it is not the case that the act ought not to be done.[21] The addition of (ii) makes clear, in combination with (i), that the act conforming to the principle of utility will be obligatory in a non-dilemmatic manner.

[19] Similarly, Bright (1991, p. 215) takes all these qualifications to show that 'Bentham did not suppose there to be any very simple or direct connection between utility and obligation.'

[20] Here, I follow Kelly (1990, p. 258) in interpreting Bentham as not requiring any sanctions for there being a moral obligation.

[21] von Wright (1968, p. 21) points out that we ordinarily understand statements of the form 'you may work or relax' as expressing its being permitted that you work and also its being permitted that you relax.

So we can safely conclude from (B34) that Bentham holds (35). Then, from (30) and (35), we get that Bentham holds the following sufficient criterion for obligation:

(36)   An act ought to be done if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *greater than* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

Moreover, an alternative path to get to the same interpretation is that (36) follows from the combination of (18) and (33).[22]

6. The case of ties

So far, we have not considered what Bentham would say about cases where the balance of pain and pleasure would be the same whether or not one does a certain act. Consider

Case 3

Act	Pleasures	Pains
a	2	1
not-*a*	1	0

Read strictly, (B6) seems to say that an act's being right requires that it conforms to the principle of utility, and, likewise read strictly, (B1) seems to say that an act's conforming to the principle of utility requires that it augments happiness. We get that act *a* in Case 3 is not right, since it does not affect the balance between pain and pleasure. Similarly, in *Rationale of Judicial Evidence*, Bentham claims that

(B37)   An immoral sort of act, is that sort of act, the tendency of which is, in some way or other, to lessen the quantity of happiness in society.   (Bentham, 1827, vol. 5, p. 303; 1838–1843, vol. 7, p. 474)

---

[22] Some commentators, such as Bright (1991, p. 229n12) and Shaver (2013, p. 295), think that Bentham must have held a more complicated view because of his (1970, pp. 28–9nd) remarks in a dialogue with an intuitionist about preventing a mischievous act, that it is one's duty, 'if it is what lies in your power, and can be done without too great a sacrifice, to endeavour to prevent it.' Yet this seems consistent with the simpler sum-total account: the qualification is necessary because if the sacrifice is large enough it will outweigh the mischief that would have been produced by the prevented act. And then the prevented act would no longer be mischievous, and preventing it would deviate from the principle of utility.

This suggests that he holds that acts are not wrong unless they diminish happiness. So acts, such as *a* and not-*a* in Case 3, that neither augment nor diminish happiness are, according to Bentham, neither right nor wrong, which is also strongly suggested by (B3).[23] This raises the puzzle of what Bentham takes to be the deontic status of such acts. Earlier in *Rationale of Judicial Evidence*, Bentham mentions, discussing moral intuitionism, a further deontic category beyond right and wrong:

(B38)   And if the tendency of an action to produce most pleasure or most pain be the criterion and measure of its claim to be pronounced right, indifferent, or wrong, in any one case; in what other can it fail of being so?   (Bentham, 1827, vol. 1, p. 123; 1838–1843, vol. 6, p. 239)

p. 102

Hence it seems that acts that neither augment nor diminish happiness are indifferent rather than right or wrong.[24] Combining this with (32), (33), and (36), my binary interpretation of Bentham's utilitarianism can be stated as follows:

The Binary Interpretation

(39)   An act *ought to be done* if and only if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *greater than* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

(40)   An act is *right* if and only if it ought to be done.

(41)   An act is *indifferent* if and only if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is *equal to* the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

(42)   An act is *wrong* if and only if the sum total of the values of the pleasures it appears to produce and the pains it appears to avert is

[23] As we noted earlier, this additional deontic category also seems to follow from the combination of (B13) and (B14). Moreover, the category seems required, in (B15), to allow for the possibility of an act that is not commanded (i.e. neither right nor obligatory) and not prohibited (i.e. neither wrong nor negatively obligatory).

[24] This view that some acts are neither right nor wrong has been defended more recently by Carlson (1995, p. 103). von Wright (1951, pp. 3–4) also treats 'indifferent' as an deontic category roughly corresponding to Bentham's notion.

less than the sum total of the values of the pains it appears to produce and the pleasures it appears to avert.

These criteria are binary, because they just compare what would happen if the act were performed with what would happen if the act were not performed.

It might seem weird that acts ought to be done if and only if they are right, but this fits with (B2), which explains both *ought* and *right* in terms of conformance to the principle of utility.[25] Moreover, note that, in (B20), Bentham regards it as logically true that 'what is *right ought* to be done'. This conflicts with contemporary usage of these terms, but it accords with (40).

7. Maximizing ends

If we accept the binary interpretation of Bentham contained in (39)–(42), there is still a major puzzle left relating to maximization. Slote (1984, pp. 153–4) claims that a footnote which was added for the 1823 second edition of *An Introduction*, defends 'a typical optimizing form of (utilitarian) act-consequentialism' in conflict with the mostly unaltered main text. This footnote states, regarding 'the principle of utility', that

(B43)   To this denomination has of late been added, or substituted, the *greatest happiness* or *greatest felicity* principle: this for shortness, instead of saying at length *that principle* which states the greatest happiness of all those whose interest is in question, as being the right and proper, and only right and proper and universally desirable, end of human action: of human action in every situation, and in particular in that of a functionary or set of functionaries exercising the powers of Government.[26]   (Bentham, 1970, p. 11na)

While this is not a clear endorsement of Moorean maximization, it seems at least clear that Bentham holds that

[25] See footnote 11. In *Pannomial Fragments*, Bentham (1838–1843, vol. 3, p. 218) claims that 'You have a *right* to perform whatever you are not under obligation to abstain from the performance of.' This might seem to conflict with (40) and (41). Yet note that he is not claiming that an act is right if it is not obligatory to abstain from it. Bentham tries to clarify here what is meant by someone's *having* a right.

[26] Much the same footnote was also added to the second edition of *A Fragment on Government*; see Bentham (1977, p. 446).

(44) The right and proper end for human action is the greatest
      happiness for all whose interest is in question.[27]

This, however, is not a doctrinal departure from the account presented in the first edition. The final chapter of the first edition states that

(B45)   Ethics at large may be defined, the art of directing men's
            actions to the production of the greatest possible quantity
            of happiness, on the part of those whose interest is in view.
            (Bentham, 1789, p. 308; 1970, p. 282)

The change announced in (B43) is not a revision of the principle of utility but merely a new name for it. Bentham makes this clear in a letter to his translator, Étienne Dumont, written some months after (B43), stating:

(B46)   *Principle of Utility*—is dead and gone. Greatest happiness
            principle has succeeded to it: that is to say in English.
            (Bentham, 2000, p. 149)

Thus there is little support, I think, for Slote's suggestion that (B43) shows that Bentham had changed his mind about maximization.[28] Nonetheless, there is still an apparent tension between the binary interpretation in (39) –(42) and the claim in (44) that the right and proper end is the greatest happiness. While Bentham does not state his explicit criteria for the rightness and wrongness of acts in terms of 'the greatest happiness', he often states that the greatest happiness is the right and proper *end*.[29] The talk about 'the greatest happiness' in (B43), (B45) and other places raises the puzzle of how this can be squared with the binary interpretation.[30]

    In addition, there are several passages in which Bentham claims that his theory requires one to produce the greatest happiness that could be produced in a situation. In *Constitutional Code Rationale*, Bentham claims that

(B47)   The greatest happiness of the greatest number of the
            members of the community in question being the right and

----

[27] See also Bentham (1952–1954, vol. 3, p. 318; 1993, p. 5).

[28] Furthermore, Bentham (1977, p. 393) used the maximizing slogan 'the greatest happiness of the greatest number' as early as 1776 in the preface to *A Fragment on Government*.

[29] Postema (2002, p. xv) and Shaver (2013, p. 294).

[30] Bentham's maximizing passages in this section are, of course, especially problematic for the satisficing interpretation.

> proper, and sole right and proper, end of government,
> misrule has place in so far as by the line of conduct
> pursued by the rulers positive unhappiness is produced, or
> any quantity of happiness that by a different line of conduct
> might have been produced by them fails of being produced.
> (Bentham, 1989, p. 270)

Here, it seems clear that, when one assesses a feasible line of conduct, every alternative line of conduct that might be pursued instead is relevant.

Another example comes from Bentham's essay 'On Retrenchment', where he compares the principle of utility—then known as the greatest-happiness principle—with a misunderstanding based on the principle's earlier locution 'the greatest happiness of the greatest number':[31]

(B48)   Bring to view in supposition two communities. Number of the individuals—in the one, 1,000: in the other, 1,001: of both together, 2,001. By the greatest happiness, the arrangement prescribed would be that by which the greatest happiness of all together would be produced. But wide indeed from this effect might be the effect of the application of the principle, if the arrangement productive of the greatest happiness of the greatest number—no regard being shewn to the happiness of the smaller number—were understood to be the arrangement prescribed by that same principle.

So long as the greatest number—the 1,001—were in the enjoyment of the greatest degree of comfort, the greatest possible degree of torment might be the lot of the smallest of the two numbers—the 1,000: and still the principle stating as the proper object of endeavour the greatest happiness of the greatest number be actually conformed to—not contravened.[32]   (Bentham, 1993, p. 352)

[31] Bentham (1970, p. 11) announced 'the greatest happiness' label in (B43).

[32] See also Bentham (1983, pp. 309–10). It should be clear from (B48) that Bentham does not defend the following view, which Bykvist (2010, p. 20) ascribes to him:

> Total happiness is not the only important thing; it also matters how many people are benefited. In a choice between two outcomes that contain the same amount of total happiness, we should realize the one in which more people are benefited.

There is, I think, a more plausible reading of the 'greatest number' part of 'the greatest happiness of the greatest number' slogan. Unlike on other moral views that were popular

Here, Bentham seems to consider a case where we have a choice between several distinct arrangements for the two communities, and, from all of these arrangements, the principle of utility prescribes the arrangement 'by which the greatest happiness of all together would be produced'. Hence, to find out whether an act is right, it seems that one has to consider what would happen for each of the available options, not just for the act and for the corresponding negative act.

A final example comes from *Codification Proposal*, where Bentham provides the following rough version of his theory:

(B49)   *Good* is pleasure or exemption from pain: or a cause or instrument of either, considered in so far as it is a cause or instrument of either.

        Evil is pain or loss of pleasure; or a cause or instrument of either; considered in so far as it is a cause or instrument of either.

        Happiness is the sum of pleasures, deduction made or not made of the sum of pains.   (Bentham, 1998, p. 256)

(B50)   To warrant the employment of *evil*, whether in the character of punishment or in any other character, two points require to be made out: 1. that, by means of it, *good* to a preponderant amount will be produced: 2. that, at any less expense of *evil*, *good* in so great proportion can not be produced.   (Bentham, 1998, p. 256)

when Leibniz (1700, p. 378) introduced the slogan, utilitarianism takes the happiness of *everyone* into account, not just the happiness of a smaller set of people such as just oneself or the ruling class. The 'greatest number' part stresses that as many people as possible—that is, everybody—should be included when one calculates the sum total of happiness. This reading is especially plausible, I think, for Leibniz's (1700, p. 378) original formulation of the slogan, translated in Hruschka (1991, p. 166):

> to act in accordance with supreme reason is to act in such a manner that the greatest quantity of good available is obtained for the greatest multitude possible and that as much felicity is diffused as the reason of things can bear.

In other words, one should, taking the greatest multitude of people into account, maximize their collective sum total of happiness. Given this reading, the slogan is not open to the often repeated charge, put forward by Edgeworth (1881, p. 117–8) and von Neumann and Morgenstern (1944, p. 11), that it self-contradictorily recommends maximizing two independent functions. But, if—as I have suggested—the maximization of number is within the scope of the maximization of happiness, the principle is no more self-contradicting than, for example, maximin. Crimmins (2011, p. 61) asks 'Why did Bentham not twig to this defect in the utilitarian formula earlier?' If I am right, however, the formula's content was not defective; the problem was just its presentation.

The second point at the end of (B50) suggests that Bentham holds that whether an act that employs some evil should be done depends on more than just the course of action one would take were one to employ some evil and the course one would take were one to not employ it.

Thus, as we have seen in this section, there are a lot of passages where Bentham seems to accept a maximizing form of utilitarianism which requires one to produce the greatest happiness among all feasible options. In the next section, we shall explore how the maximizing described in these passages can be squared with the binary interpretation.

8. The role of negative and complex acts in Bentham's maximizing

The key to understanding the relation between Bentham's criteria for right and wrong in (39)–(42) and the end of realizing the greatest happiness in (44) is his theory of action, which includes negative and complex acts. We begin with negative acts, which as we saw in (B17) consists in forbearing to do a certain act—that is, if a is an act, then forbearing to do a is doing the negative act not-a.

Consider the following case where one has a choice between three acts, which are mutually exclusive and jointly exhaustive:

Case 4

Act	Pleasures	Pains
a	3	0
b	2	0
c	1	0

If one performs one of the acts b and c, one will perform the negative act not-a. And, regardless of how one does not-a (that is, regardless of which one of b and c is performed), doing a instead would have had a better balance of pleasure. Hence doing not-a is wrong according to (42). The only way to avoid doing something wrong according to (42) is to pursue the line of conduct that produces the greatest happiness.

Case 4 also illustrates a peculiar feature of binary utilitarianism. The outcome of a negative act such as not-b depends, if there are more than two courses of action, on how not-b is done.[33] Suppose we wonder

---

[33] One might object that, if Bentham thought of negative acts in this way, one would expect him to talk about different instances of a single negative act; but he does not. Yet

whether *b* is right or wrong. If *a* were performed, then not-*b* would produce the same amount of pain and pleasure as *a* so *b* would be wrong, because *b* would produce less happiness than not-*b* would. On the other hand, if *c* were performed then not-*b* would produce the same amount of pain and pleasure as *c* so *b* would be right, because *b* would produce more happiness than not-*b* would. This phenomenon—that the deontic status of an act in a situation depends on what one does in the situation—is called *normative variance*. The most common objection to normative variance is that it leads to problems with action-guidance.[34] Bentham is not open to this objection because one only needs to know that the act that produces the greatest possible happiness in the situation (in Case 4, for example, act *a*) is guaranteed to not be wrong.

This also illustrates how binary maximization differs from Moorean maximization. These two forms of maximization are similar in that the only way to avoid doing anything wrong is to perform the act with the greatest possible happiness in the situation. The two forms differ, however, in how they handle non-optimal courses of action. On Moorean maximization, *b* is wrong in Case 4 regardless of what acts are performed in the situation. But, on binary maximization, as we have just seen, *b* would be right if *c* were performed.

p. 107

If there are two or more optimal acts, things get slightly more complicated. Consider the following case:

Case 5

Act	Pleasures	Pains
a	3	0
b	3	0
c	2	0
d	1	0

If one would perform *b* if one were to perform not-*a* and one would perform *a* if one were to perform not-*a*, then neither *a* nor *b* would be obligatory; they would both be indifferent since *a*, not-*a*, *b*, and not-*b* would all produce the same amount of pain and pleasure.

To handle this case, we need to take into account not just negative acts but also complex acts. In *An Introduction*, Bentham describes complex acts as follows:

this absence can be explained by the general scarcity of discussions of situations with more than two alternatives in Bentham's writings.

[34] See, for example, Carlson (1995, pp. 100–1).

(B51)   acts may be distinguished into *simple* and *complex*: simple, such as the act of striking, the act of leaning, or the act of drinking […]: complex, consisting each of a multitude of simple acts, which, though numerous and heterogeneous, derive a sort of unity from the relation they bear to some common design or end; such as the act of giving a dinner, the act of maintaining a child, the act of exhibiting a triumph, the act of bearing arms, the act of holding a court, and so fourth.   (Bentham, 1970, pp. 78–9)

Consider the complex act that consists in the negative acts not-c and not-d, that is, (not-c)-and-(not-d). The simple parts of this complex act derive a unity by their common end, that is, the only right and proper end, the greatest happiness. If one were to perform neither of the optimal acts a and b, one would perform one of c and d, that is, one would perform the disjunctive act c-or-d or, put in more verbose terms of negative and complex acts, one would perform not-((not-c)-and-(not-d)), that is, the negative act that consists in not performing the complex act (not-c)-and-(not-d). Yet performing not-((not-c)-and-(not-d)) would be wrong according to (42).[35] So, again, we have that in order to not do something wrong according to (42), one has to do an act that will realize the right and proper end of the *greatest* happiness for the concerned individuals.

In this manner, we can show that, in any situation, the only way to avoid doing a wrong act according to Bentham's binary utilitarianism is to do the act that produces the greatest happiness out of all feasible acts in the situation. Hence, unlike rival interpretations, the binary inter-p. 108pretation renders Bentham consistent. On the binary interpretation, we can make sense of Bentham's felicific calculus, his criterion of right and wrong, and his talk of maximization in his slogans and elsewhere. Hence there is, I think, a convincing case that Bentham was a binary act utilitarian.

[35] One might object that performing this complex act would perhaps not be intentional in the sense that one actively contemplated its performance. But Bentham (1970, pp. 75–6nd) explains that to be punishable an act need only be intentional in the sense that it could have been avoided if the agent '*had* so willed'.

References

Bentham, Jeremy (1789) *An Introduction to the Principles of Morals and Legislation*, London: T. Payne.

— (1821) *The Elements of the Art of Packing, as Applied to Special Juries, Particularly in Cases of Libel Law*, London: Effingham Wilson.

— (1827) *Rationale of Judicial Evidence, Specially Applied to English Practice*, ed. John S. Mill, London: Hunt and Clarke, 5 vols.

— (1831) *Parliamentary Candidate's Proposed Declaration of Principles: Or Say, a Test Proposed for Parliamentary Candidates*, London: Heward, published anonymously.

— (1838–1843) *The Works of Jeremy Bentham, Published under the Superintendence of His Executor John Bowring*, ed. John Bowring, Edinburgh: William Tait, 11 vols.

— (1952–1954) *Jeremy Bentham's Economic Writings: Critical Edition Based on His Printed Works and Unpublished Manuscripts*, ed. W. Stark, London: George Allen & Unwin, 3 vols.

— (1970) *An Introduction to the Principles of Morals and Legislation*, eds. J. H. Burns and H. L. A. Hart, The Collected Works of Jeremy Bentham, London: Athlone.

— (1977) *A Comment on the Commentaries and A Fragment on Government*, eds. J. H. Burns and H. L. A. Hart, The Collected Works of Jeremy Bentham, London: Athlone.

— (1983) *Deontology Together with A Table of the Springs of Action and the Article on Utilitarianism*, ed. Amnon Goldworth, The Collected Works of Jeremy Bentham, Oxford: Clarendon Press.

— (1989) *First Principles Preparatory to Constitutional Code*, ed. Philip Schofield, The Collected Works of Jeremy Bentham, Oxford: Clarendon Press.

— (1993) *Official Aptitude Maximized; Expense Minimized*, ed. Philip Schofield, The Collected Works of Jeremy Bentham, Oxford: Clarendon Press.

— (1998) *'Legislator of the World': Writings on Codification, Law, and Education*, eds. Philip Schofield and Jonathan Harris, The Collected Works of Jeremy Bentham, Oxford: Clarendon Press.

— (2000) *The Correspondence of Jeremy Bentham Volume 11: January 1822 to June 1824*, ed. Catherine Fuller, The Collected Works of Jeremy Bentham, Oxford: Clarendon Press.

— (2010) *Of the Limits of the Penal Branch of Jurisprudence*, ed. Philip

Schofield, The Collected Works of Jeremy Bentham, Oxford: Clarendon Press.

Bright, Robert W. (1991) *Foundations of Utilitarianism*, Ph.D. thesis, Dalhousie University.

Bykvist, Krister (2010) *Utilitarianism: A Guide for the Perplexed*, London: Continuum.

Carlson, Erik (1995) *Consequentialism Reconsidered*, Dordrecht: Kluwer.

Clapham, Sir John (1944) *The Bank of England: A History*, Cambridge: Cambridge University Press, 2 vols.

Crimmins, James E. (2011) *Utilitarian Philosophy and Politics: Bentham's Later Years*, London: Continuum.

Edgeworth, F. Y. (1881) *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*, London: Kegan Paul.

Hruschka, Joachim (1991) 'The Greatest Happiness Principle and Other Early German Anticipations of Utilitarian Theory', *Utilitas* 3 (2): 165–177.

Kelly, P. J. (1990) 'Utilitarian Strategies in Bentham and John Stuart Mill', *Utilitas* 2 (2): 245–266.

Leibniz, G. W. (1700) 'Observationes de Principio Juris', *Monatlicher Auszug/Aus allerhand neu-herausgegebenen/nützlichen und artigen Büchern* (July): 371–382, published anonymously.

Moore, G. E. (1912) *Ethics*, London: Williams & Norgate.

Postema, Gerald J. (1998) 'Bentham's Equality-Sensitive Utilitarianism', *Utilitas* 10 (2): 144–158.

— (2002) 'Introduction', in Gerald J. Postema, ed., *Bentham: Moral Political and Legal Philosophy*, pp. xi–xxxii, Aldershot: Ashgate.

Quinton, Anthony (1973) *Utilitarian Ethics*, London: Macmillan.

Rosen, Frederick (2003) *Classical Utilitarianism from Hume to Mill*, London: Routledge.

Schofield, Philip (2006) *Utility and Democracy: The Political Thought of Jeremy Bentham*, Oxford: Oxford University Press.

Shaver, Robert (2013) 'Utilitarianism: Bentham and Rashdall', in Roger Crisp, ed., *The Oxford Handbook of the History of Ethics*, pp. 292–311, Oxford: Oxford University Press.

Singer, Marcus G. (1977) 'Actual Consequence Utilitarianism', *Mind* 86 (341): 67–77.

Singer, Peter (1976) 'A Utilitarian Population Principle', in Michael D. Bayles, ed., *Ethics and Population*, pp. 81–99, Cambridge, MA: Schenkman.

Slote, Michael (1984) 'Satisficing Consequentialism', *Proceedings of the Aristotelian Society, Supplementary Volumes* 58:139–163.

von Neumann, John and Oskar Morgenstern (1944) *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

von Wright, Georg Henrik (1951) 'Deontic Logic', *Mind* 60 (237): 1–15.

— (1968) *An Essay in Deontic Logic and the General Theory of Action*, Amsterdam: North-Holland.

— (1981) 'On the Logic of Norms and Actions', in Risto Hilpinen, ed., *New Studies in Deontic Logic: Norms, Actions, and the Foundations of Ethics*, pp. 3–35, Dordrecht: Reidel.