Philos Stud (2008) 138:211–223
DOI 10.1007/s11098-006-9032-5

ORIGINAL PAPER

# Conversation and conditionals

**J. Robert G. Williams**

Received: 7 April 2006 / Accepted: 26 June 2006 / Published online: 30 August 2006


**Abstract** I outline and motivate a way of implementing a closest world theory of indicatives, appealing to Stalnaker's framework of open conversational possibilities. Stalnakerian conversational dynamics helps us resolve two outstanding puzzles for a such a theory of indicative conditionals. The first puzzle—concerning so-called 'reverse Sobel sequences'—can be resolved by conversation dynamics in a theory-neutral way: the explanation works as much for Lewisian counterfactuals as for the account of indicatives developed here. Resolving the second puzzle, by contrast, relies on the interplay between the particular theory of indicative conditionals developed here and Stalnakerian dynamics. The upshot is an attractive resolution of the so-called ''Gibbard phenomenon'' for indicative conditionals.

**Keywords** Conditionals · Counterfactual · Indicative · Subjunctive · Conversation · Pragmatics · Gibbard · Sobel · Stalnaker

Stalnakerian conversational dynamics can help us resolve two outstanding puzzles for a ''closest-world'' modal theory of indicative conditionals. I begin the paper by outlining and motivating a new way of implementing a closest world theory of indicatives, appealing to Stalnaker's framework of open conversational possibilities. Stalnaker's framework itself shows its utility in application to conditionals by allowing us to explain a puzzling feature of conditionals—concerning so-called 'reverse Sobel sequences'—in a theory-neutral way. The explanation has application to any ''closest worlds'' account of indicative or counterfactual conditionals, as well as to other truth-conditional accounts of conditionals. My favoured closest world theory of indicative conditionals, when combined with Stalnakerian dynamics, gives an attractive resolution of the so-called ''Gibbard phenomenon'' for indicative conditionals.

J. R. G. Williams (✉)
School of Philosophy, University of Leeds, Woodhouse Lane, Leeds, West Yorks LS2 9JT, UK
e-mail: j.r.g.williams@leeds.ac.uk

## Conditionals

Consider the following two conditionals:

(1) If Oswald didn't shoot Kennedy, someone else did

(2) If Oswald hadn't shot Kennedy, someone else would have

The first is an *indicative* conditional; the second is a *counterfactual* conditional. I take it that the former is true, and the latter is false. Indicative and counterfactual conditionals, then, must differ semantically.[1]

A popular semantics for the counterfactual conditionals gives it the following truth-conditions:[2]

'$A \mathrel{\Box\!\!\rightarrow} C$' is true at $w$ iff $C$ holds at all the closest $A$-worlds to $w$.[3]

In such analyses, 'closest' is here a technical term, which might then be further analysed: Lewis (1973a, 1979) holds it should be unpacked in terms of the similarity of possible worlds to one another.

We know from (1) and (2) above that we cannot offer exactly the same semantics for indicative conditionals. Surely, though, the presumption has to be that the semantics of conditionals share a common form: syntactically they involve the same connective 'if'; they exhibit the same logical behaviour.

Let us therefore consider what is going on in the Oswald–Kennedy cases. It seems right to say that the counterfactual ''if Oswald hadn't shot Kennedy, someone else would have'' is false because at the closest worlds where Oswald doesn't shoot Kennedy (e.g. his gun doesn't go off; or he has a last minute change of heart; or he fires and misses), nobody else shoots Kennedy. Intuitively, why don't these close worlds likewise undermine the truth of the *indicative* conditional? I suggest it is because such worlds are inconsistent with what is taken as *common ground* about the case: that Kennedy got shot. Worlds where Kennedy doesn't end up being shot, I suggest, are ignored when evaluating an indicative conditional. The analysis of indicative conditionals suggested by these thoughts is the following:

'$A \rightarrow C$' is true at $w$ iff $C$ holds at all the closest *open* $A$-worlds to $w$.[4]

[1] The example is due to Adams (1975).

[2] I use '$\Box\!\!\rightarrow$' for the counterfactual conditional, '$\rightarrow$' for the indicative conditional; and '$>$' for an arbitrary conditional.

[3] $A$-worlds are worlds at which $A$ is true. The account may need to be amended if the so-called 'limit assumption' is denied—see Lewis (1973b)—but not in ways that effect the present discussion.

[4] Compare Nolan (2003), who also tries to build a 'closest world' account of indicatives on top of the Lewisian treatment of counterfactuals, by appealing to knowledge. Nolan's approach is to define a new notion of 'closeness', whereby worlds incompatible with what we know are ipso facto 'far away' from actuality. Stalnaker (1975, p. 71), giving one of the earliest defences of a 'closest world' account of indicatives, is even closer to the present proposal. My account, in contrast, eliminates such worlds altogether from the space of worlds over which the conditional is defined (that is, my framework is not *universal* in the sense of Lewis, 1973, p. 120). This difference will have substantial knock-on effects when we come to look at Gibbard cases. Note also that we do not have to presuppose some substantive account of counterfactual closeness (say, along the lines of Lewis, 1979 to buy into my proposal: even if the notion of counterfactual closeness is primitive, or analysed *in terms of counterfactuals* rather than the other way round, we can obviously still appeal to it when giving the truth-conditions of indicatives, in the way described in the main text.

The account of indicative conditionals just mentioned explains why the Oswald–Kennedy conditionals give different results in the indicative and counterfactual formulations. In formulating this account, however, we have had to appeal to a new notion: *openness* of worlds. I suggest we spell this out using notions drawn from Stalnaker (1978, 1984). Stalnaker defines the *context set*, at a given stage of a conversation, to consist of all those worlds which are *not* collectively presupposed *not* to obtain.[5] This fits quite nicely with the gloss given above: that of open worlds as those which are compatible with all facts that are taken as common ground.[6]

My suggestion, therefore, is that Stalnaker's context set has more than a pragmatic role in the analysis of indicative conditionals: it *defines* which worlds are 'left out' of the space of possibilities over which indicative conditionals are defined, and thus enters into the semantics of such conditionals.[7]

The standard way in which Stalnaker appeals to the context set is quite different. For the next two sections, we shall concentrate on describing how the Stalnakerian framework can be applied to the special case of conditionals, independently of the particular thesis about their truth-conditions just advocated. We shall see, in Sect. 3, that the Stalnakerian framework can do substantial explanatory work in a theory of conditionals, by explaining in a theory-neutral way puzzles over so-called 'Sobel' and 'reverse Sobel' cases. In Sect. 4, I turn to the interaction between conversational dynamics and the particular account of semantics for the indicative conditional given above. This combination turns out to give an attractive explanation of one outstanding problem for a truth-conditional account of indicatives: the Gibbard phenomenon.

## Conversation

Presuppositions change as conversations progress. As Stalnaker defines it, this will mean that the *context set* (those worlds which are not ruled out by one of the presuppositions in force) will be being updated constantly. Stalnaker, indeed, holds that we can analyse the dynamics of conversations by looking at the effects that

[5] Presumably one can presuppose something which is false: in which case the *actual world* may not be an open possibility. (In Stalnaker's framework, in presupposing $p$ where $p$ is in fact false, we are in effect presupposing that the actual world does not obtain.) For this observation, and discussion of its potential impact on the logic of indicative conditionals, see Nolan (2004): the threat is that instances of modus ponens fail. The threat of unacceptable logical revisionism does not arise for those (such as Nolan, 2003; Weatherson, 2001) who use factive attitudinal states (such as knowledge) instead of presupposition, within their account of indicative closeness. There are a number of options at this point: (Nolan, 2004) canvasses some options. Since I am not inclined to regard modus ponens as negotiable, the relevant alternatives for me are (a) to defend some broadly epistemic (and factive) substitute for Stalnakerian presuppositions; or (b) to tweak the setup laid out in the main text so that the actual world will always count as an "open possibility" even if it presupposed not to obtain. I discuss these issues in other work.

[6] One can make judgements about indicative conditionals in non-conversational situations. But there is a natural understanding about what 'open worlds' are in that setting: the worlds compatible with what we believe (or know) to be the case.

[7] Formally, take the set of worlds $W$ and closeness-ordering $\sigma$ over which the counterfactual conditional is defined. Given the context set $C$, the set of worlds over which the indicative conditional is defined as follows: $W' := C \cap W$. The ordering is just that induced by $\sigma$: if $x$ is closer than $z$ to $y$ on the original space of worlds, and $x,y,z \in W'$, then $x$ is closer than $z$ to $y$ on the derived space of worlds.

assertions, denials and the like have on the context set. This very general idea is used to provide illuminating explanations of a number of otherwise puzzling linguistic phenomena.[8]

Stalnaker's explanations are based on a very simple idea about how assertions interact with the context set. Recall that worlds within the context set are 'live conversational possibilities': for all that we can presuppose in making an assertion, any one of those possibilities might be actual. By uttering *S* assertorically, one is committing oneself to *S*'s being true *no matter which world is actual*. If the assertion is taken on board by one's conversational partners, then the presupposition that *S* is true *no matter which world is actual* is thereby established. This will typically change the context set: worlds which, if actual, would make *S* false, will be eliminated. The new context set will contain only worlds compatible with the new common ground. The situation is illustrated in Fig. 1.

An example: in the pre-existing context, we shall take it that no-one has said anything to establish any presuppositions about what minerals bananas contain. Worlds where bananas contain potassium, and worlds where they don't, are both compatible with all facts that are part of the common ground. Someone now asserts *bananas contain potassium*. If this is not challenged, then it becomes part of the new common ground: the context set is updated by eliminating worlds where bananas do not contain potassium.

Another example: in the pre-existing context, we take it that no propositions about the constitution of the watery-stuff in our environment are part of the common ground: worlds where this role is played by $H_2O$ and worlds where it is played by XYZ are both included in the context set. If I utter ''water is H2O'', I eliminate from the context set every world where uttering ''water is H2O'' would express a falsehood. For example, worlds where the watery stuff in my environment (and featuring in the causal history of my usage of 'water') is XYZ, for example, are eliminated.[9]

What does this general story say about the effects of uttering a conditional, of whatever sort? In accordance to the general line described above, when we utter 'if *A*, *C*' this should eliminate all worlds which that utterance expresses something false. There is one immediate effect of this. On *any* truth-conditional account of the conditionals (indicative or counterfactual), the conditional with antecedent *A* and consequent *C* will be *false* if uttered at an $A \wedge \neg C$ world. So no matter what truth-conditions we think conditionals have, we should say that uttering such a conditional eliminates from the context set all $A \wedge \neg C$ worlds. The greyed-out area in Fig. 2 corresponds to worlds eliminated by asserting ''if *A*, *C*'':

Depending on the details of the truth-conditions assigned, some of the $\neg A$-worlds may also be eliminated. On a Lewisian account of counterfactuals, any $\neg A$ world which is closer to an $A \wedge C$-world than it is to an $A \wedge \neg C$-world will be eliminated from the context set. On the material conditional account of conditionals, by contrast, the *only* worlds where uttering the conditional is uttering a falsehood will be $A \wedge \neg C$-worlds; so (absent further argument) these are the only worlds that get

[8] A very nice example is the treatment of negative existentials in Stalnaker (1978).

[9] Compare Stalnaker (1984, ch.4.). Given the actual world is an H2O world, Kripkean orthodoxy has it that ''water is H2O'' is true relative to *every world*. Nevertheless, those same sounds, uttered on twin-earth, would express a falsehood (they would not express the same proposition).

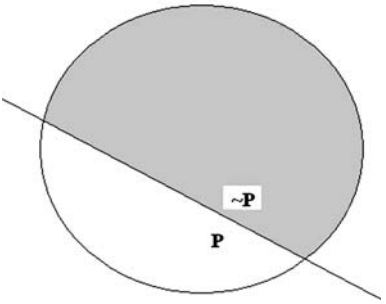


**Fig. 1** Effect of uttering "*S*" on context set: greyed-out worlds are eliminated. (*P*-worlds are worlds where *S* expresses a truth)

eliminated. Nothing in what follows will depend on what happens to the $\neg A$ worlds, so we can ignore these differences for present purposes.[10]

Below, I look at one way in which applying general Stalnakerian conversational dynamics to the case of conditionals can explain puzzling features of conditionals: the infelicity of reverse Sobel sequences. This explanation will be applicable to many truth-conditional conditionals: counterfactual as well as indicative. However, since on my account the semantics of indicative conditionals itself makes appeal to the context set, we can expect some interesting interactions between Stalnakerian dynamics and conditionals *specific* to my account. The account of the Gibbard phenomenon in Sect. 4 illustrates the power of this combination.

## Reverse Sobel sequences

In the previous section I outlined a very general feature of the effect of asserting a conditional on the context set of a conversation: it eliminates all $A \wedge \neg C$ worlds, without exception, from that set.

You might think this is a peculiar result for the 'closest-worlds' conditional theorist. For famously, natural language conditionals give rise to *Sobel sequences*, e.g:[11]

1. If Oswald didn't shoot Kennedy, someone else shot Kennedy
2. If Oswald didn't shoot Kennedy and the *National Inquirer* was right about the existence and character of space aliens[12], no-one shot Kennedy

The point here is that natural language conditionals 'if A, C', seem to *leave it open* that $A \wedge \neg C$ is true at 'far off' worlds. But the conversational story we've just given

[10] I suggest that the conversational effects of asserting an indicative conditional with the truth-conditions suggested earlier are the same as those described for the material conditional: once the $A \wedge \neg C$ worlds are eliminated from the context set, there *are no* open $A \wedge \neg C$ possibilities, and hence vacuously, every world in the context set is such that $C$ holds at every closest open $A$-world. Of course, this doesn't mean that there are no differences between the material conditional and closest-world conditional under consideration, but just that their typical effects on the context set are identical.

[11] The case below is formulated with indicative conditionals. For Sobel sequences of counterfactuals, see Lewis (1973a, ch.1.).

[12] I am taking it that the *National Inquirer* holds that space aliens are generally inclined to fake the death of celebrities in order to cover up the fact that they have kidnapped them.

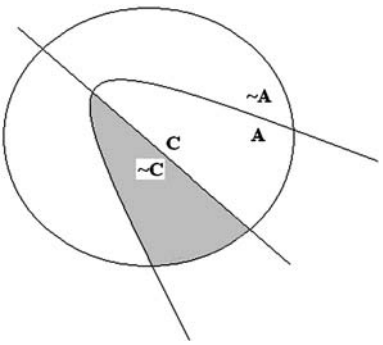


**Fig. 2** Effect on context set of "if $A$, $C$" (greyed out area=worlds eliminated)

says that *all* $A \wedge \neg C$ worlds *without exception*—even 'far-off' ones—are eliminated by the assertion of an ordinary indicative conditional.[13]

I think that we do indeed rule out all such worlds when asserting an conditional. Nevertheless, I think the acceptability of Sobel sequences can be accounted for, and further, we can use the setting to explain some *prima facie* puzzling results for the variable strict theory of indicative conditionals.

The story about the acceptability of asserting Sobel sequences is the following. In asserting 'If Oswald didn't shoot Kennedy, someone else shot Kennedy', we mean to eliminate *all* conversational possibilities where neither Oswald nor anyone else shot Kennedy. What of possibilities where the *National Inquirer* is right about space aliens faking the death of celebrities in order to kidnap them? Clearly, there is normally a presupposition that the actual world is not this way: such a world is not standardly one of the conversational possibilities. So we need not take it into consideration when considering whether the assertion was appropriate.

On the other hand, there are ways of putting such possibilities into the context set: one such way, it seems, is just by asserting 'it might be that the *National Inquirer* is right...'.[14] Another way, I would contend, is by uttering a conditional whose antecedent is true only at such a conversational possibility.[15] Thus, when I state the second element of the Sobel sequence, the context set first *expands* to incorporate worlds where the *National Inquirer* is right; and then we eliminate all the '*National Inquirer*, Oswald not shooting, someone else shooting' worlds. The net effect, then, of asserting (1) and then (2), is that we end up with an expanded context set, but one without either of the following classes of worlds:

- Worlds where *National Inquirer* speaks nonsense, Oswald didn't shoot Kennedy, and no-one shot Kennedy (eliminated by (1)).
- Worlds where the *National Inquirer* speaks the truth, Oswald didn't shoot Kennedy, and someone else shot Kennedy (eliminated by (2)).

[13] Analogous remarks apply to the assertion of a counterfactual conditional.

[14] This role for 'might' utterances is advocated in Swanson (2005).

[15] That is, I contend that in uttering 'if $A$ then $C$', there must either be within, or one must introduce into, the context set worlds where $A$ is true. Compare Stalnaker (1975, p. 81), who advocates the rule 'it is appropriate to make an indicative conditional statement or supposition only in a context which is compatible with the antecedent'. If an indicative conditional is put forward where the pre-existing context set does not fit this description, then we *change* the context set to accommodate the assertion, by bringing in extra worlds.

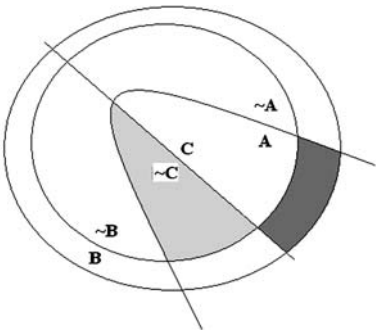


**Fig. 3** Result on context set of uttering Sobel sequence: "if $A$, $C$"; "if $A$ and $B$, $\neg C$" (worlds eliminated by the former in light grey; worlds eliminated by the latter in dark grey)

This is illustrated in Fig. 3.

This seems a perfectly sensible description of the result of making such assertions, and there is nothing to suggest anything bad going on. Nothing in the story so far rests on any particular truth-conditional account of conditionals: so it may be appealed to by material-conditionals theorists as well as closest worlds accounts, whether of counterfactuals or indicatives.

On the other hand, consider the following 'reverse Sobel sequence':[16]

2. If Oswald didn't shoot Kennedy and the National Inquirer was right about the existence and character of space aliens, no-one shot Kennedy.

    1. If Oswald didn't shoot Kennedy, someone else shot Kennedy.

The idea here is to consider a situation where one utters (2) *and then* (1). In point of view of *consistency*, of course, nothing depends on the order in which things are asserted, so the 'closest-world' conditional theorist is standardly taken to be committed to saying exactly the same thing about this sequence as the standard Sobel assertions. Yet, the reverse Sobel sequence are far less comfortable—even infelicitous—to utter.

The point I want to make here is that we can *explain* the difference between the respective felicity of Sobel and reverse-Sobel patterns of assertion via the conversational dynamics just sketched. In the reversed case, as before, our initial context set will standardly not contain any worlds at which the *National Inquirer* reports truly. But in uttering (2), we introduce such worlds into consideration. The distinctive effect of asserting that conditional, as before, is to eliminate all "National Inquirer" worlds where Oswald didn't shoot Kennedy, but someone else did.

What would happen if we were now to utter (1)? By our initial discussion of the conversational dynamics of asserting conditionals, we know it will eliminate *all* Oswald-not-shooting and no-one-shooting worlds: in particular, it will eliminate those conversational possibilities where Oswald didn't shoot Kennedy, where no-one else did, *and* where the *National Inquirer* is right. Hence, uttering (1) after (2) would lead to the context set ruling out all three of the following classes of worlds:

- Worlds where the *National Inquirer* speaks the truth, Oswald didn't shoot Kennedy, and someone else shot Kennedy (eliminated by (2)).

[16] Reverse Sobel sequences seem to first appear in print in a paper by von Fintel (2001). von Fintel credits Irene Heim for communicating the phenomenon to him.

- Worlds where *National Inquirer* speaks nonsense, Oswald didn't shoot Kennedy, and no-one shot Kennedy (eliminated by (1)).
- Worlds where the *National Inquirer* speaks the truth, Oswald didn't shoot Kennedy, and no-one shot Kennedy (also eliminated by (1)).

This is illustrated in Fig. 4.

So, the context set resulting from asserting (1) and (2) in the 'normal' order leaves open more worlds than the context set resulting from asserting (2) and (1) in the reverse order. We will be able to explain why the Sobel sequence is fine, while the reverse Sobel sequence is bad, if we can make the case that it is inappropriate to assert something that eliminates situations of the final class (*National Inquirer* speaks truth, neither Oswald nor anyone else shot Kennedy). The disputed class is highlighted in Fig. 5.

It seems that such a case is available, for in the envisaged situation we have no grounds for thinking that such worlds are *not* actual. If the *National Inquirer is* right, then all our pro-shooting-of-Kennedy evidence is the result of a massive conspiracy, so for all we know, Kennedy wasn't shot at all. The difference in the conversational effects of Sobel vs. reverse Sobel sequences thus explains the difference in the acceptability of those patterns of assertion.[17]

The general utility of appeal to conversational dynamics to understand the behaviour of conditionals is thus illustrated. To emphasize once again: only conversational effects common to (almost) all truth-conditional accounts of conditionals have been appealed to in explaining the potential infelicity of reverse Sobel cases (all that is required is that a conditional never be true at a world where it's antecedent is true and its consequent false). So again, this explanation is available to strict and material conditional theorists, Lewisian theorists of counterfactuals, as well as analyses of indicative conditionals of the style advocated in this paper.[18]

## The Gibbard phenomenon

Gibbard (1980) argues against truth-conditional theories of conditionals. It is well known that material conditional theorists avoid these worries; but they remain a standing challenge to stronger accounts of the truth-conditions of indicative conditionals. I shall show how my account of the truth-conditions of indicatives, far from falling to Gibbard's challenge, gives an attractive account of the phenomenon to which Gibbard points.

I shall follow the presentation of Gibbard phenomenon given by Bennett (2003). The situation to be considered features a system of water-gates and levers, which are constructed according to the following rules:

[17] You might eliminate such worlds as a byproduct of ruling out worlds where the *National Inquirer* is right about such matters; but if this is ones reason, Gricean reasoning would suggest that one should simply assert this straight out, rather than uttering the indicative. Andrew McGonigal pointed out another predication of the present account that seems to be borne out: in cases where the final class of worlds is *empty*, then we shouldn't expect Sobel and reverse Sobel to differ. There are trivial cases of this (where the antecedent is repeated, or adjoined to its contrary). But potentially there are more interesting test cases: where the Sobel sequences are $A > C$, $(A \wedge B) > \neg C$, and there is a standing presumption against $A \wedge B \wedge \neg C$ worlds.

[18] Gillies (2005) (independently of the present work) gives a detailed explanation of the Sobel/reverse Sobel difference for the particular case of a strict conditional account of counterfactuals, in ways that are similar in spirit to those given above.

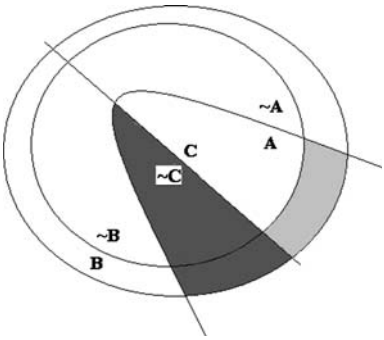


**Fig. 4** Result on context set of uttering reverse Sobel sequence: ''if $A$ and $B$, $\neg C$''; ''if $A$, $C$'' (worlds eliminated by the former in light grey; worlds eliminated by the latter in dark grey)

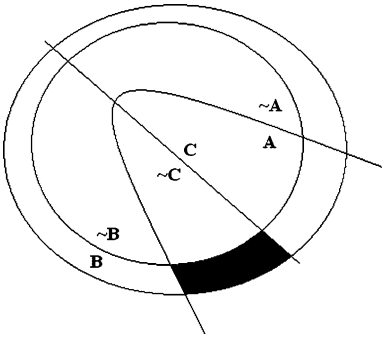


**Fig. 5** Blacked out region contains worlds eliminated by reverse Sobel sequence but not by Sobel sequence

- Lever 1 down and top gate raised: water flows left
- Lever 2 down and top gate raised: water flows right
- Except! that if both levers down, then top gate can't be raised.

We have two informants, Abel and Cain. They are each standing next to a lever, out of sight and earshot of each other. Their job is to tell me what will happen to the water. Abel can see that the first lever is down. Cain can see that the second lever is down. They say:

- Abel: ''If the top gate is raised, water will flow left, not right''
- Cain: ''If the bottom gate is raised, water will flow right, not left''

*Both utterances are perfectly appropriate*. However, it is clear that what they say will happen in the event of the gate being raised *can't both take place*. The intuitive, and correct, conclusion for me to draw from this information is that *the top gate will not be raised*.

The following is then an argument for the *no truth value* approach to the indicative conditional. The argument is schematic: using the Abel/Cain case to rule out each potential truth-conditional approach to the indicative conditional $\theta$:

1. Abel and Cain's utterances are both true
2. Truth-conditional account $\theta$ of indicative conditionals supports **conditional non-contradiction** : $\neg((A \rightarrow B) \wedge (A \rightarrow \neg B))$

3. Conditional non-contradiction means that it is not the case that Abel and Cain's utterances are both true
4. *Therefore*: $\theta$ is not the right account of indicative conditionals.

The first premiss is non-negotiable for present purposes. The argument is plainly of a valid form, so if one is to resist it to defend a particular truth-conditional account of indicative conditionals, one must resist either premise (2) or (3).

A material conditional account of indicative conditionals allows one to resist at premise (2): for conditional non-contradiction does not hold for material conditionals. Famously, material conditionals suffer from the *paradoxes of material implication*. Whenever, $A$ is false, $A \supset C$ is true. Thus, whenever $A$ is false, $(A \supset C) \wedge (A \supset \neg C)$ is true. Note, too, that if $(A \supset C) \wedge (A \supset \neg C)$ is to be true, it follows that $A$ is false. So, in the Abel/Cain case, if we take on board what each of our informants say, we would expect to be able to derive 'the top gate will not open''. This seems exactly right: it is what we *should* conclude from that testimony.

Material conditional accounts of indicative conditionals are, however, independently objectionable. Our interest is in seeing whether ''closest world'' accounts of indicatives fall to Gibbard-style considerations. I contend that the approach advocated in this paper allows one to rebut the argument directly.[19]

We begin by noting a loophole in the argument available to *all* modal accounts of the conditional: strict conditional theorists as well as 'closest world' theorists. (2), as stated, fails in some special instances. The failure of conditional non-contradiction in the material case is founded upon the paradoxes of material implication. The failure in the modal conditional case is founded on a similar result. The so-called *paradoxes of strict implication* state that if $A$ is false at all worlds, then $(A \dashv C)$ is true no matter what $C$ is. Under the same conditions, the 'closest worlds' conditional $A > C$ is true:[20] in this case, the result is known as the *vacuous truth of conditionals with impossible antecedents*.[21]

However, these are the *only* exceptions to conditional non-contradiction modal accounts allow. So, from the truth of Cain and Abel's utterances, we can apparently derive that opening the gate is *impossible*. If we patch the argument scheme by adding the premise that the relevant antecedent is possible, then we appear to have sealed the loophole just sketched.

But this response crucially depends by what is meant by 'impossible'. If we are giving a standard Lewisian account of the counterfactual conditional, then coun-

[19] The solution is 'incompatibilist' in the sense of (Lycan, 2001, ch. 8.): the two informants are vindicated, but only by denying the relevant instance of conditional non-contradiction. Unlike the proposals of Nolan (2003) and Lycan (2001), (and many of the proposals outlined in Lycan, 2001, ch. 8 fn. 4) I deny premiss (2) rather than premise (3) of the relevant instance of the argument scheme above. Lycan (2001) says:

...one further option open to us, which no one has yet explored, is to deny Conditional Non-Contradiction in the first place. (Lycan, 2001, p. 176)

What I note below is that standard modal accounts of the conditional *already* have a loophole whereby they can deny instances of Conditional Non-contradiction. Plausibly to exploit this loophole and resolve the Gibbard phenomenon (and setting aside the material conditional account of indicatives), one needs something like the modal account of the indicative conditional that I have given.

[20] Recall, we use $>$ as a variable over conditionals.

[21] In the gloss of the truth-conditions given above, the conditional is to be true if the consequent is true at *all closest A-worlds*. If there are *no A* worlds, then vacuously, the consequent is true at *all A worlds*, and a fortiori *all closest A worlds*. Closest worlds accounts of conditionals are often formulated with an additional clause making the vacuous truth case explicit.

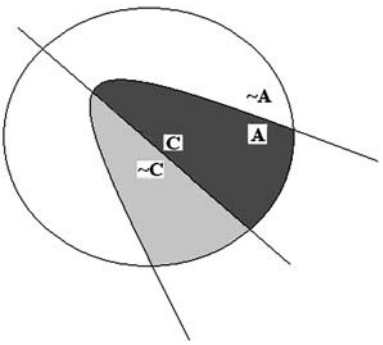


**Fig. 6** Assertion of "if $A$, $C$" eliminates light grey worlds; assertion of "if $A$, $\neg C$" eliminates dark grey worlds

terfactuals $AC$ will be vacuously true only when $A$ is impossible in the sense of *being false at all metaphysically possible worlds*. But on the account of indicative conditionals I advocate, the relevant 'possible worlds' are all and only those within the *context set*—those which are 'open possibilities' in our conversational context. Thus, in the relevant sense, for $A$ to be impossible is for *no A world to be within the context set*—that is, for $\neg A$ to be implicit in the common ground of that stage of the conversation. My account of the truth-conditions of indicative conditionals, therefore, looks like it is well placed to *exploit* the loophole in the Gibbardian argument, and allow for both Abel's and Cain's utterances to both be true in a single context. Unlike the analysis of Sobel and reverse-Sobel discussed earlier, this way of responding to the Gibbard phenomenon *crucially depends* on the particular analysis of indicative conditionals I laid out in Sect. 1.[22]

I end by noting that the account of conversational dynamics with which we have been working *predicts* that the result of taking on board the reports of both Abel and Cain in the Gibbardian situation is exactly one where the common antecedent is false throughout the open conversational possibilities. That is, applying general principles of conversational dynamics, we predict the emergence of a single context where both conditionals are true.

Recall that the basic effect of asserting the indicative conditional 'if $A$, $C$' is to remove any $A \wedge \neg C$ worlds from the context set. So if conversational participants assert *both* 'if $A$, $C$' and 'if $A$, $\neg C$', then all $A \wedge \neg C$ and all $A \wedge C$ worlds will be eliminated from the context set. That is, all $A$ worlds *whatsoever* are eliminated from the context set.[23] In the Gibbardian case, the net result of the two assertions is a situation where only $\neg A$ worlds remain in the context set. Thus, as required, conversational dynamics predicts the emergence of a context where (on my account of the indicative conditional) Abel and Cain's utterances are both true. The situation is illustrated in Fig. 6.

We have exploited the loophole in the Gibbardian argument to defend an account of the Gibbard phenomenon whereby both Abel and Cain's utterances are true

[22] Thus, it will have nothing to say about putative examples of the Gibbard phenomenon involving counterfactual rather than indicative conditionals. Orthodoxy suggests there are no such cases; for debate, see Edgington (1995, 1997) and Morton (1997).

[23] One might think, for Gricean reasons, that one should simply assert $\neg A$; but of course, the present case is one in which two *different* individuals are each asserting one piece of information.

within a single context: we do this because, on our favoured account, Stalnaker's context set provides the range of open possibilities which feed into the semantics of the indicative conditional. On this interpretation, both conditionals are true just when the context set contains no *A*-possibilities: which is exactly what we find emerging once both assertions have been made.[24]

**Acknowledgements** Thanks to Ross Cameron, George Darby, Andy McGonigal and Daniel Nolan for discussion of this material. This paper was stimulated by thinking through the Gibbard phenomenon with students from my class on conditionals at the University of Leeds in 2006. Thanks to all involved.

## References

Adams, E. W. (1975) *The logic of conditionals*. Dordrecht: Reidel.
Bennett, J. (2003) *A philosophical guide to conditionals*. Oxford: Oxford University Press.
Edgington, D. (1995) On conditionals. *Mind, 104*, 235–329.
Edgington, D. (1997) Truth, objectivity, counterfactuals and Gibbard. *Mind, 106*, 107–116.
Gibbard, A. (1980) Two recent theories of conditionals. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs: Conditionals, belief, decision, chance and time*. The Netherlands: Kluwer Academic Publishers.
Gillies, T. (ms. 2005). Counterfactual scorekeeping.
Lewis, D. K. (1973a) *Counterfactuals*. Oxford: Blackwell.
Lewis, D. K. (1973b) Counterfactuals and comparitive possibility. *Journal of Philosophical Logic, 2*, 418–446. Reprinted with postscripts in Lewis, *Philosophical Papers vol. II* (Oxford University Press, 1986), pp. 3–31.
Lewis, D. K. (1979) Counterfactual dependence and time's arrow. *Noûs, 13*, 455–476. Reprinted with postscript in Lewis, *Philosophical Papers II* (Oxford University Press, 1986), pp. 32–51. Also reprinted in Jackson (Ed.), *Conditionals* (Oxford University Press, 1991), pp. 46–76.
Lycan, W. G. (2001) *Real conditionals*. Oxford: Oxford University Press.
Morton, A. (1997) Can Edgington Gibbard counterfactuals? *Mind, 106*, 101–106.

[24] My account need have no truck, therefore, with the suggestion that the way to deal with Gibbard cases is to think that Abel and Cain's utterances are only true *in their respective contexts*. Such contextualist responses to the Gibbard phenomenon appear to me highly unattractive. There are also worries that if conditionals are sensitive to highly local context in this way, then apparent disagreements about conditionals will be depicted as people 'talking past' one another. It is a virtue of the present account that all such worries are quieted.

Nolan (2003) offers an account of the truth-conditions of indicative conditionals in some ways very close to mine. As noted above, he suggests that worlds incompatible with what we know to be the case are *ipso facto* further away. However, this means that Nolan *cannot* avail himself of the loophole to Gibbard's argument I identify: since Abel and Cain's utterances will not be vacuously true. To give a sense in which they are both true, he must think that the *metric of worlds* shifts according to whether it is Abel or Cain that is speaking. Relative to metric on worlds generated shared conversational knowledge, Abel and Cain's conditionals will probably turn out false, on Nolan's account. Another advocate of worlds-semantics for indicative conditionals is Weatherson (2001) (I read Weatherson's excellent article only after completing this paper). Weatherson's framework differs from the one I advocate at a number of points (e.g. unlike the approaches favoured by Nolan and myself, counterfactual closeness of worlds does not play a significant role in Weatherson's characterization of indicative closeness of worlds). However, his account and mine are similar in allowing both informants in the Gibbard scenario to be speaking truly in 'violation' of the relevant instance of conditional non-contradiction—Weatherson achieves this by introducing an ''impossible world'' where everything is true, and allowing its closeness to actuality to vary according to the epistemic states of the agents concerned. In recent work, Weatherson has abandoned this account of the Gibbard phenomenon in favour one based on relativism about the truth-values of conditionals.

Nolan, D. (2003) Defending a possible-worlds account of indicative conditionals. *Philosophical Studies, 116*, 215–269.
Nolan, D. (ms. 2004). Is Stalnaker inconsistent about the indicative conditional?.
Stalnaker, R. (1975). Indicative conditionals. *Philosophia, 5*. Reprinted in Stalnaker, *Context and Content* (Oxford: Oxford University Press, 1999). Also reprinted in Jackson (Ed.) *Conditionals* (Oxford University Press, 1991), pp. 136–154.
Stalnaker, R. (1978). Assertion. *Syntax and Semantics, 9*, 315–332. Reprinted in Stalnaker, *Context and Content* (Oxford University Press, Oxford, 1999), pp.78–95.
Stalnaker, R. (1984) *Inquiry*. Cambridge, MA: MIT Press.
Swanson, E. (ms. December 2005). Something 'might' might mean.
von Fintel, K. (2001). Counterfactuals in a dynamic context. In M. Kenstowicz (Ed.), *Ken Hale: A life in language*. MIT Press.
Weatherson, B. (2001). Indicative and subjunctive conditionals. *Philosophical Studies, 51*, 200.