

# SOSA'S SAFETY NEEDS SUPPLEMENTING, NOT SAVING: A REPLY TO COMESAÑA AND MCBRIDE

John N. WILLIAMS

ABSTRACT: Juan Comesaña argues that *Halloween Party* shows that Sosa's (2002) disjunctive safety condition on knowledge is too strong. Mark McBride agrees, and proposes a modification to that condition in order to evade *Halloween Party*. I show that that *Halloween Party* is not a counterexample to Sosa's disjunctive safety condition. However the condition, as well as McBride's modification to it, is insufficient for true belief (or acceptance) to be knowledge. Sosa's condition needs supplementing in some way that would yield a full analysis of knowledge.

KEYWORDS: Juan Comesaña, *Halloween Party*, Mark McBride, safe belief, Ernest Sosa

For Robert Nozick, your true belief counts as knowledge just in case it is sensitive to falsehood and adherent to truth, meaning roughly that you would not have the belief were it false and that it would still be true were you to hold it under slightly changed circumstances.<sup>1</sup> Dissatisfied with this, Ernest Sosa proposed replacing Nozick's adherence to truth condition with its logically independent contrapositive, roughly that were you to hold the belief under slightly different circumstances then it would still be true. He called this 'safely true' belief. After proposing two non-disjunctive formulations of the safety condition,<sup>2</sup> he proposed a

---

<sup>1</sup> Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981), 179.

<sup>2</sup> Ernest Sosa, "How to Defeat Opposition to Moore," *Noûs* 333, Supplement (1999): 141-153, "How Must Knowledge Be Modally Related to What Is Known?" *Philosophical Topics* 26, 1/2 (1999): 373-384, "Tracking, Competence, and Knowledge," in *The Oxford Handbook to Epistemology*, ed. Paul Moser (New York: Oxford University Press, 2002), 264-86.

John N. Williams

disjunctive condition,<sup>3</sup> before moving away from an analysis of knowledge as true safe belief (or ‘acceptance’) to one in terms of “apt and adroit” belief.<sup>4</sup>

In “Unsafe Knowledge,” Juan Comesaña presents *Halloween Party*, and argues that this example shows that Sosa’s (2002) disjunctive safety condition on knowledge is too strong, predicting ignorance where there is knowledge.<sup>5</sup> In “Saving Sosa’s Safety,” Mark McBride agrees, and proposes a modification to that condition in order to evade *Halloween Party*.<sup>6</sup> I show that that contrary to Comesaña and McBride, *Halloween Party* is not a counterexample to Sosa’s disjunctive safety condition. There is no need to save Sosa’s condition from *Halloween Party*. However the condition, as well as McBride’s modification to it, is insufficient for true belief (or acceptance) to be knowledge. Sosa’s condition needs supplementing in some way that if it is to yield a full analysis of knowledge.

## 1. Sosa’s Disjunctive Safety Condition

Sosa’s (2002) *disjunctive* safety condition on knowledge is as follows.

### *Disjunctive Safety*

S knows that  $p$  on the basis of an indication  $I(p)$  only if either (a)  $I(p)$  indicates the truth outright and  $S$  accepts that indication as such outright, or (b) for some condition  $C$ ,  $I(p)$  indicates the truth dependently on  $C$ , and  $S$  accepts that indication as such not outright but *guided* by  $C$  (so that  $S$  accepts the indication as such *on the basis of*  $C$ ).<sup>7</sup>

Some terminology needs explication, as follows.

An indication  $I$  that  $p$  indicates the truth that  $p$  outright just in case  $I$  entails that  $p$ .

An indication  $I$  that  $p$  indicates the truth that  $p$  dependently on  $C$  just in case  $I$  does not entail that  $p$  but  $C$  obtains and the conjunction of  $C$  with  $I$  entails that  $p$ .

$S$  accepts an indication  $I$  that  $p$  as such outright just in case  $S$  accepts that  $p$  solely

---

<sup>3</sup> Sosa, “Tracking.”

<sup>4</sup> Ernest Sosa, *A Virtue Epistemology: Apt Belief and Reflective Knowledge Volume I*. (Oxford: Oxford University Press, 2007), *Reflective Knowledge: Apt Belief and Reflective Knowledge Volume II*. (Oxford: Oxford University Press, 2009).

<sup>5</sup> Juan Comesaña, “Unsafe Knowledge,” *Synthese* 146, 3 (2005): 395-404.

<sup>6</sup> Mark McBride, “Saving Sosa’s Safety,” *Logos & Episteme* III, 4 (2012): 637-652.

<sup>7</sup> Sosa, “Tracking,” 275-276.

Sosa's Safety Needs Supplementing, Not Saving: A Reply to Comesaña and McBride

on the basis of *I*.

*S* accepts an indication *I* that *p* as such not outright but guided by *C* just in case *S* accepts that *p* on the basis of the conjunction of *C* with *I*.

At this point we should note the mention of “*some condition*” in disjunct (b). Next, consider Comesaña’s *Halloween Party*, as follows.

There is a Halloween party at Andy’s house, and I am invited. Andy’s house is very difficult to find, so he hires Judy to stand at a crossroads and direct people towards the house (Judy’s job is to tell people that the party is at the house down the left road). Unbeknownst to me Andy doesn’t want Michael to go to the party, so he also tells Judy that if she sees Michael she should tell him the same thing she tells everybody else (that the party is at the house down the left road), but she should immediately phone Andy so that the party can be moved to Adam’s house, which is down the right road. I seriously consider disguising myself as Michael, but at the last moment I don’t. When I get to the crossroads, I ask Judy where the party is, and she tells me that it is down the left road.<sup>8</sup>

Comesaña argues that this shows Sosa’s disjunctive safety condition to be too strong, predicting ignorance where there is knowledge. The indicator *I* is Judy’s testimony to me that the party is down the left road. Is disjunct (a) satisfied? No. Judy’s testimony to me does not indicate outright the truth that the party is down the left road, because her testimony to me that it is down the left road does not *entail* that it is down the left road. Had I disguised myself as Michael, then her testimony to me would have remained the same, but the party would not be down the left road, but down the right road. Is disjunct (b) satisfied? To decide that, we have to decide what counts as condition *C*. Comesaña takes this as the condition *that I do not disguise myself as Michael*. He points out the following.

I am unaware of the relevance of the respective condition to the truth of Judy’s testimony: I would have believed that *p* whether or not I looked like Michael to Judy.<sup>9</sup>

Echoing this correct observation, McBride makes the following remark.

As the case is set up, I’ll accept Judy’s testimony whether or not I appear to her Michael’ly. So I don’t accept the indication ‘guided by,’ or ‘on the basis of,’ *C*.<sup>10</sup>

Comesaña immediately concludes as follows.

---

<sup>8</sup> Comesaña, “Unsafe,” 396.

<sup>9</sup> Comesaña, “Unsafe,” 398.

<sup>10</sup> McBride, “Saving,” 640.

John N. Williams

Therefore, HALLOWEEN PARTY is a counterexample to the safety condition even taking into account dependent indication.<sup>11</sup>

In other words (b) is not satisfied, and so my acceptance that the party is down the left road is unsafe.

This does not follow. The question still remains of whether there is “some condition,” even if not *C*, that satisfies disjunct (b). Since I do not know that Andy does not want Michael to go to the party, presumably I also know nothing about his instructions to Judy about Michael. What then is the basis on which I accept her testimony to me? It appears to be nothing more than condition *C'* —*that she is telling me the truth*. After all, her testimony to me could hardly *guide* me into accepting that the party down the road on the left—and with it, to the party itself—if she were not telling me the truth. That condition obtains. She is indeed telling me the truth. I accept Judy’s testimony to me not as outright but guided by the condition that she is telling me the truth, and her testimony to me, plus the fact that she is telling me the truth, entails that the party is down the left road. On this very plausible reading of the example, disjunct (b) is satisfied, and we have knowledge incorporating safe acceptance. Sosa’s disjunctive safety condition survives *Halloween Party!*

Comesaña considers my conclusion that “my belief that the party is at the house down the left road is safe after all,”<sup>12</sup> replying as follows.

... my belief does not satisfy Sosa’s definition of safety: it *could* easily have happened that I had the *same belief* on the *same basis* and yet the belief was false.<sup>13</sup>

But here it seems that Comesaña has a different safety condition in mind, one that Sosa also proposes, that fixes both the content and the basis of the belief from actuality across close possible worlds, as follows.

*Content-and-basis-fixed Safety*

If *S* knows that *p* on basis *B*, then *S* could not easily form the false belief that *p* on basis *B*.<sup>14</sup>

---

<sup>11</sup> Comesaña, “Unsafe,” 398. McBride agrees, writing that “... so Sosa’s updated (2002) safety principle – as Comesaña notes – cuts no ice against HALLOWEEN PARTY. By its lights we still have unsafe knowledge” (“Saving,” 640).

<sup>12</sup> Comesaña, “Unsafe,” 399.

<sup>13</sup> Comesaña, “Unsafe,” 399, my italics.

<sup>14</sup> As Comesaña observes (“Unsafe,” 403, note 4) this is incorporated by Sosa’s (2002) requirement

In other words, if  $S$  knows that  $p$  on basis  $B$ , then  $S$ 's belief that  $p$ , formed on basis  $B$  in close possible worlds, is true. *Halloween Party* indeed shows that this condition is too strong. In close possible worlds in which I disguise myself as Michael, and in which I form the belief that the party is down the left road on the basis of Judy's testimony to me that it is down the left road, the party is not down the left road, but down the right road. My belief that the party is down the left road is unsafe. But surely I know that the party is down the left road.

I have looked at what would be the case *after* I receive Judy's testimony, for only then can I be guided by condition  $C'$  —*that she is telling me the truth*. In contrast, Comesaña insists that we must look at what would be the case *shortly before* I receive Judy's testimony and just after I decide to not disguise myself as Michael. He writes as follows.

... it seems to me simply false that, in HALLOWEEN PARTY, after I decide not to dress up as Michael it is no longer a close possibility that I have a false belief. When considering whether the proposition that  $p$  obtains safely at  $t$  in the actual world, we consider whether it obtains in possible worlds that differ from the actual world just slightly *right before*  $t$ . And, in HALLOWEEN PARTY, I seriously consider dressing myself up as Michael just before driving to the intersection where Judy is standing.<sup>15</sup>

But what needs to be considered is not whether the “proposition that  $p$ ” obtains safely in the actual world, but whether the *belief* that  $p$ —in this case my belief that the party down the road on the left—is actually safe. This might seem like a minor quibble, but it helps to explain why Comesaña continues as follows.

Moreover, we can change the case so that the time when I decide not to dress up as Michael is even closer to the time when I believe that the party is at the house down the left road. We can suppose, if we want, that I *was* dressed up as Michael, and that I decided to take the disguise off at the last minute, just before arriving at the intersection where Judy is. We can also make more radical changes to the case, by imagining that I am dressed up as Michael, but that I'm going to the party with Alex, and that we decide at the last moment that he will ask Judy for directions, not me. In any of those cases, there are possible worlds that differ from the actual world just in what happens *right before* I believe that the party is at the

---

that  $S$ 's belief must be based on a reliable indication, one that would not have been present without it being so that  $p$ .

<sup>15</sup> Comesaña, “Unsafe,” 399. He does not tell us which time ‘ $t$ ’ is supposed to denote. It appears to be the time at which I decide not to disguise myself as Michael.

John N. Williams

house down the left road, and that are such that *my belief is false*.<sup>16</sup>

But until I actually receive Judy's testimony, I form *no actual belief* about the location of the party. Until then, although the *proposition* that the party is down the road on the left might turn out to be false (if I receive Judy's testimony while disguised as Michael) I have no *belief* that the party is down the road on the left that could easily or not easily be false. In deciding whether my actual belief that the party is down the road on the left could easily be false, we must look at close possible worlds in which I *have* that belief, and these are close possible worlds in which I have *already received* Judy's testimony. Given that I have received it while not disguised as Michael, Judy's testimony is truthful, with the result that my actual belief that the party is down the road on the left is one that pre-theoretically, "could not easily be false." Given that I have acquired the evidence of her testimony while not disguised as Michael, it is not a stroke of luck that the party is down the road on the left. It seems then, that my belief that the party is down the road on the left is safe according to Sosa's disjunctive safety condition.

But Sosa's disjunctive safety condition incorrectly allows you to know that it is 4:30 pm in *Stopped Clock*, as follows.

You habitually nap between 4:00 pm and 5:00 pm. Your method of ascertaining the time you wake is to observe, between 4:00 pm and 5:00 pm, the position of the hands of your clock, one you know has always worked perfectly reliably. Awaking at 4:30 pm, you see that its hands point to 4:30 pm. Accordingly, you form the belief that it is 4:30 pm. And it is indeed 4:30 pm because exactly twenty-four hours ago a stray fleck of dust chanced to enter the clock's mechanism, stopping it.

Disjunct (a) is not satisfied. The indication *I*—that the hands of your clock point to 4:30 pm—does not indicate the truth outright that it is 4:30 pm, because the fact that the hands of your clock point to 4:30 pm does not entail that it is 4:30 pm. At all other times during the hour that you nap the hands still point to 4:30 pm without it being 4:30 pm. Is disjunct (b) satisfied? There is a condition *C* that obtains—that *the hands of your clock point to the correct time*. This does not indicate the truth outright that it is 4:30 pm, because the fact that the hands of your clock point to the correct time does not entail that it is 4:30 pm. But the conjunction of *C* with *I* (that the hands of your clock point to 4:30 pm) does indeed

---

<sup>16</sup> Comesaña, "Unsafe," 399, my italics.

entail that it is 4:30 pm. And surely you are guided by  $C$ . You accept the hands of your clock pointing to 4:30 pm as a veridical indication of the time on the basis of  $C$ . Thus is there is some condition, namely  $C$ , such that the hands of your clock pointing to 4:30 pm indicates the truth dependently on this condition, and you accept that indication as such not outright but *guided* by it (so that you accept the indication as such *on the basis* of it). So disjunct (b) is satisfied. Your acceptance that it is 4:30 pm is safe according to Sosa's formulation. Yet although it is indeed 4:30 pm, surely you do not know that it is 4:30 pm. Your acceptance that it is 4:30 pm is luckily true. You were lucky to look at your clock exactly twenty-four hours after it stopped working, at the only instant during the hour when you nap at which its hands could have pointed to the correct time. Thus Sosa's disjunctive safety condition must be supplemented in some way to produce a full analysis of knowledge.

## 2. McBride's Modification to Sosa's Disjunctive Safety Condition

Following Comesaña in thinking mistakenly that *Halloween Party* shows Sosa's disjunctive safety condition to be too strong, McBride proposes a modification of it, as follows.

### *McBride's Modification of Sosa's Disjunctive Safety*

S knows that  $p$  on the basis of an indication  $I(p)$  only if EITHER (a)  $I(p)$  indicates the truth outright and S accepts that indication as such outright, OR (b) either (i) for some condition  $C$ ,  $I(p)$  indicates the truth dependently on  $C$ , and S accepts that indication as such not outright but guided by  $C$  (so that S accepts the indication as such on the basis of  $C$ ), or (ii) for some non-trivial condition  $C_{\text{SAFE}}$ ,  $I(p)$  indicates the truth dependently on  $C_{\text{SAFE}}$ , and S accepts that indication not-as-such outright.<sup>17</sup>

This is more complicated. Here is an explication of the extra terminology.

A condition  $C$  is  $C_{\text{SAFE}}$  just in case  $C$  obtains, and if  $C$  were the case in the way described in the thought-experiment under consideration, then  $C$  would hold in all close possible worlds.<sup>18</sup>

A condition  $C$  is trivial with respect to  $p$  just in case  $C$  is  $p$  or  $C$  entails that  $p$ .<sup>19</sup>

---

<sup>17</sup> McBride, "Saving," 642-643.

<sup>18</sup> McBride, "Saving," 642.

<sup>19</sup> McBride, "Saving," 643.

John N. Williams

*S* accepts an indication *I* that *p* not-as-such outright just in case *S* accepts that *p* on the basis of *I* but not solely on the basis of *I*.<sup>20</sup>

McBride's modification is supposed to make my acceptance that the party is down the left road safe. Disjunct (a) is not satisfied. Judy's testimony to me that the party is down the road on the left does not indicate the truth that the party is down the road on the left outright. Her bare testimony to me *itself* does not entail its truth. Is (b)(i) satisfied? Yes, for the same reason that disjunct (b) is satisfied in Sosa's disjunctive safety condition. There is a condition *C*—*that Judy is telling me the truth*. That condition obtains. She is indeed telling the truth. I accept Judy's testimony to me not as outright but guided by the condition that she is telling the truth, and her testimony to me, plus the fact that she is telling the truth, entails that the party is down the left road. Thus we need not inspect (b)(ii) since disjunct (b) is already satisfied. My acceptance that the party is down the road on the left is safe. So far so good for McBride's modification.

But McBride's modification is also shown to be insufficient by *Stopped Clock*. Disjunct (a) is not satisfied. The indication *I* that the hands of your clock point to 4:30 pm does not indicate the truth outright that it is 4:30 pm, because the fact that the hands of your clock point to 4:30 pm does not entail that it is 4:30 pm. Is (b)(i) satisfied? Yes. You accept the indication *I* that the hands of your clock point to 4:30 pm as indicating the truth that it is 4:30 pm, dependently on *C*—*that the hands of your clock point to the correct time*. You are guided by this condition. That condition indeed obtains. And the conjunction of *C* with *I*—that the hands of your clock point to 4:30 pm—indeed entails that it is 4:30 pm. McBride's modification renders your acceptance that it is 4:30 pm safe. Yet surely you do not know on the basis that the hands of your clock point to 4:30 pm that it is 4:30 pm. McBride's modification of Sosa's disjunctive safety condition must also be supplemented in some way to produce a full analysis of knowledge.

### 3. Concluding Remarks

In sum, McBride does not need to supplement Sosa's disjunctive safety condition in order to evade Comesaña's *Halloween Party*, because *Halloween Party* does not show that this condition is too strong. However it, as well as McBride's modification, is too weak. Thus McBride's modification appears to give Sosa no advantage. But it is better for a condition to be too weak than too strong. There is

---

<sup>20</sup> McBride, "Saving," 643, note 24.

Sosa's Safety Needs Supplementing, Not Saving: A Reply to Comesaña and McBride  
hope yet that Sosa's disjunctive (2002) safety condition may be supplemented with some further condition in some way to yield a full analysis of knowledge. What that condition might be is a separate question.