

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/106619/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Kelp, Christoph and Simion, Mona 2017. Criticism and blame in action and assertion. *Journal of Philosophy* 114 (2) , pp. 76-93. 10.5840/jphil201711426

Publishers page: <http://dx.doi.org/10.5840/jphil201711426>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Criticism and Blame in Action and Assertion

CHRISTOPH KELP

Institute of Philosophy, KU Leuven

MONA SIMION

Institute of Philosophy, KU Leuven

Arché Research Group, University of St. Andrews

INTRODUCTION

Norms govern actions in all walks of life. For better or worse, we are capable of violating these norms and do so frequently. When we do break the rules, we are criticisable for breaking them. Sometimes, we are even deserving of blame. However, the latter is not always the case. Often enough we violate rules and still walk free of blame. The compulsive and small children are paradigm cases of blameless violators of norms.

While it is widely accepted that norms can be violated blamelessly, and while there is a pretty reasonable understanding of when this happens implicit in the literature, there are few if any explicit accounts.¹ This paper supplies this lack. Its first main ambition is to develop a normative framework for action in general including detailed accounts of criticisability, blamelessness and blameworthiness (section 1).

A specific type of action that has received a considerable amount of attention in recent epistemology is the speech act of assertion. Assertion is an action. And since actions in general are governed by norms, it will come as no big surprise that assertion makes no exception on this front. However, it is widely agreed that there is a norm associated with assertion in particular. More specifically, it is thought that there is an epistemic norm governing assertion. This means that there is such a thing as epistemic

permissibility and impermissibility for assertion. In particular, it is often thought that assertion is governed by a rule of the following form:

The C Rule of Assertion

One must: assert that p only if p has C.²

While many contributors to the debate on norms of assertion subscribe to *The C Rule of Assertion*, there is no consensus on the identity of the crucial property C. According to the perhaps most prominent account, C is the property of being known by the speaker. In other words,

The Knowledge Rule of Assertion (KRA)

One must: assert that p only if one knows that p.³

We accept KRA. Our second aim in this paper is to apply the normative framework to the case of assertion. We will argue that this allows its champions to defuse a prominent line of objection against KRA, which ventures to show that KRA's knowledge requirement on permissible assertion is too strong (section 2).

Finally, we will go even one step further and provide reason to believe that the argument against KRA can be turned on its head. With the normative framework from section 1 in play (and given a couple of further plausible assumptions), the cases that are supposed to show that KRA is too strong can be shown to confirm KRA and to disconfirm rival views.

I. CRITICISMS AND BLAME: A BASIC NORMATIVE FRAMEWORK

1.1. Criticisability and Norm-Specific Blamelessness. It is widely acknowledged in the literature that a permissible action is a blameless action. This applies at the level of specific norms such as the rule of Uno requiring players to call Uno when playing their penultimate card. If you do call Uno when playing your penultimate card, your action is permissible by this norm and so *blameless relative to this norm*.

Consider next a situation in which a particular norm is violated. Say, you played your penultimate card without calling Uno. You violated a rule of Uno. If so, you can, of course, be prima facie legitimately criticised for violating this norm. In the Uno case, we may do this for instance by saying: “You didn’t call Uno!”

Even so, it may be that you violate a norm and yet you are blameless for so doing. If so, you are also, of course, *blameless relative to this norm*. More specifically, there are three ways in which this may happen.

One way of blamelessly violating a particular norm is through overriding.⁴ This happens when the requirements of the norm you are violating are in conflict with the requirements of another norm that takes precedence in the situation. For instance, suppose you are playing a game of Uno and are required by the rules of Uno to call Uno when playing your next card. Suppose, at the same time, someone will kill your neighbour if you do so. What the rules of Uno require of you is in conflict with the requirements of moral norms, which take precedence here. In other words, moral norms override the rules of Uno. When, because of this, you go on to violate the rules of Uno, you do so blamelessly.

The second way of blamelessly violating a norm is by violating it because⁵ your action is not under your control. Suppose, for instance, that you have been brainwashed by your guru not to call Uno when playing your penultimate card. Here you violate the rule but are clearly blameless for doing so.

Finally, the third way manifests itself in situations in which you violate a norm because you are unaware that this is what you are doing.⁶ Suppose, for instance, that you are unaware that the rules of Uno require you to call Uno when playing your penultimate card. As a result, you do not do so. In this case you violate a rule of Uno. Again, you are blameless for doing so. Ignorance excuses also.

With regard to the second and the third way of blamelessly violating a norm, some qualifications are needed. To see why, suppose that you knew that you would undergo brainwashing were you to go back to your guru. You had also promised not to go back. However, you went anyway. The impermissible act you are made to perform may be out of your control. Even so, you are blameworthy (see below) for another act, going back to your guru, of which the impermissible act is a consequence. As a result, you are not blameless for violating the rule of Uno. Strictly speaking, then, lack of control excuses only when it is itself blameless.

Similarly, suppose that, in our toy case, you had promised to read up on the rules of Uno before playing but did not do so. In this case, you are unaware of the relevant rule of Uno. Even so, you are blameworthy for another act (in this case an omission), your failure to read up on the rules of Uno, of which your failure to call Uno is a consequence. As a result you are not blameless for not calling Uno. Strictly speaking, then, ignorance excuses only when it is itself blameless.⁷

In sum, then, we want to propose the following account of blamelessness with respect to specific norms:

Norm-Specific Criticisability

An agent is prima facie legitimately criticisable relative to a specific norm N for φ -ing iff φ -ing violates N

Norm-Specific Blamelessness

An agent is blameless relative to a specific norm N for φ -ing iff

- (1) φ -ing is permissible by N or
- (2) φ -ing violates N but the agent φ s
 - (2.a) in order to comply with the requirements of a (non-overridden) overriding norm or
 - (2.b) because this is blamelessly out of her control or
 - (2.c) because the agent is blamelessly ignorant that φ -ing violates N.

On this account, an action can be criticisable relative to a specific norm and, at the same time, blameless relative to that very same norm. While this may look odd at first sight, on reflection, it is entirely as it should be. Actions are often performed in the public sphere and, as such, are observable by others, who may pick up the forms of behaviour exhibited. When you fail to call Uno when playing your penultimate card and so violate a rule of the game, this may be observed by someone else who will pick up your behaviour and, as a result, may violate the rule in the future, too. By allowing for criticisms of

actions that violate specific norms we can work against the spread of norm-violating forms of behaviour. Since this is a good thing, it makes sense for us to allow for such criticisms. At the same time, we may also want to grant that a norm has been broken blamelessly by the agent. We do not want to hold the norm violation against her: she was blamelessly ignorant, things were blamelessly out of control and so on. If so, there is excellent reason for us to allow criticisability relative to a specific norm and blamelessness relative to the very same norm to coexist.⁸

1.2. All-Things Considered Blamelessness and Blameworthiness. Thus far we have looked at the blamelessness of an action with respect to specific norms. However, it is common to distinguish between assessments of actions with respect to specific norms and all-things-considered assessments of actions. All-things-considered assessments take into account the entire normative profile of an action and assesses whether the action was permissible, required, or forbidden in view of its entire normative profile. Unsurprisingly, then, just as an action can be blameless relative to a specific norm, it can also be all-things-considered blameless. For that reason we now want to extend the above account to all-things-considered blamelessness.

An action is all-things-considered blameless if it is all-things-considered permissible. There are, in turn, two ways in which this can happen.

First, an action is all-things-considered permissible if it is permissible by all the specific norms that apply to it (henceforth also *fully permissible* for short). Suppose you call Uno when playing your penultimate card and thus comply with the rules of Uno. Suppose, in addition, you do not thereby violate any practical and moral norms and that

no other norms apply to your act. Then your calling Uno is all-things-considered permissible.

Second, an action is all-things considered permissible if it is permissible by all (non-overridden) overriding norms that apply to it. Suppose you do not call Uno when playing your penultimate card and thus violate a rule of Uno. Suppose, in addition, you violate a practical norm in so doing: you will be punished and are less likely to win. At the same time, your neighbour will die if you call Uno and so calling Uno is prohibited by moral norms. Suppose there are no further norms applying to your action. In order to save your neighbour's life, you do not call Uno. In this case, your action is all-things-considered permissible. The moral norms override the norms of Uno and the practical norms (without being themselves overridden by further norms) and your not calling Uno is permissible by the moral norms.

To repeat, what we have seen now are two ways in which an action can be all-things-considered blameless in virtue of being all-things-considered permissible. That said, even an action that is all-things-considered impermissible can be all-things-considered blameless. What we want to suggest is that an all-things-considered impermissible action is all-things-considered blameless if the action is blameless relative to all specific norms that apply to it. Suppose that you play your penultimate card without calling Uno. However, this is because you are blamelessly unaware that there is a rule requiring to you call Uno when playing your penultimate card. Suppose that your action is permissible by moral and practical norms and that there are no further norms applying to your action. In this case, you are all-things-considered blameless for not calling Uno. Since your act is permissible by moral and practical norms, it is blameless relative to

these norms. Since you do not call Uno because you are blamelessly unaware that there is a rule requiring you to do so, you are blameless relative to this rule. Since these are all the rules that apply to your action in this case, your action is blameless relative to all specific norms that apply to it. So, it is all-thing-considered blameless.

Finally, an agent is blameworthy if and only if she is not all-things-considered blameless.

In sum, we want to propose the following:

All-Things-Considered Blamelessness

An agent is all-things-considered blameless for φ -ing iff

- (1) φ -ing is all-things-considered permissible (that is, either fully permissible or permissible by all (non-overridden) overriding norms that apply to it) or
- (2) φ -ing is all-things-considered impermissible but the agent's φ -ing is blameless relative to all specific norms that apply to it.

All-Things-Considered Criticisability/Blameworthiness

An agent is blameworthy for φ -ing iff she is not all-things considered blameless for φ -ing.

It may be worth noting that, according to this account, criticisability occurs at the level of assessments by specific norms. In this way, it is fine-grained, as it were. In contrast, blameworthiness occurs at the level of all-things-considered assessments and so is coarse-grained. Blamelessness can occur at both levels.

This completes our normative framework for criticisability, blamelessness, and blameworthiness. We would like to emphasise once more that this is a perfectly general framework, in the sense that it applies to action in general. That said, in what follows, we would like to apply the framework to a particular type of act, to wit the speech act of assertion. More specifically, we will first look at a famous problem for KRA and then show that the above framework serves to offer an appealing solution to it.

II. AN APPLICATION: ASSERTION

II.1. The Case Against KRA. There is a prominent line of argument aiming to show that KRA is mistaken. More specifically, it ventures to show that KRA's knowledge requirement on assertion is too strong. In order to achieve this, foes of KRA adduce cases in which a speaker is said to make a permissible assertion, whilst not knowing what they assert. Crucially, evidence that the assertion is permissible is supposed to reside in the fact that the relevant speakers are not deserving of criticism or alternatively that they are blameless. Here are some characteristic statements of the argumentative strategy:

I shall show that there are cases in which a speaker asserts that p in the absence of knowing that p without being subject to criticism in any relevant sense, thereby showing that knowledge cannot be what is required for proper assertion.⁹

[I]f breaching a rule makes one blameworthy, which typically it does, then, [in the relevant cases], on the knowledge account, the asserter comes out as being

blameworthy, contrary to intuition.¹⁰

The classical cases that foes of KRA have adduced against KRA are cases in which speakers assert (i) justified false beliefs and (ii) gettiered beliefs as well as (iii) cases of selfless assertion. By way of illustration, consider the following examples:

1. Assertion on justified false belief

FAKE SNOW. “[I]t is winter, and it looks exactly as it would if there were snow outside, but in fact that white stuff is not snow but foam put there by a film crew of whose existence I have no idea ... [I] assert that there is snow outside.”¹¹

2. Assertion on gettiered belief

FAKE BARNS. “[S]uppose that Wendy correctly sees the only real barn that, unbeknownst to her, is completely surrounded by barn facades and asserts to me ‘There was a barn in the field we just passed’ on this basis.”¹²

3. Selfless assertions

CREATIONIST TEACHER. “Stella is a devoutly Christian fourth-grade teacher [...] Part of this faith includes a belief in the truth of creationism and, accordingly, a belief in the falsity of evolutionary theory. Despite this, Stella fully recognizes that there is an overwhelming amount of scientific evidence against both of these beliefs [...] [S]he regards her duty as a teacher to include presenting material that is best supported by the available evidence, which clearly includes the truth of evolutionary

theory. As a result, while presenting her biology lesson today, Stella asserts to her students, ‘Modern day *Homo sapiens* evolved from *Homo erectus*,’ though she herself neither believes nor knows this proposition.”¹³

In all of these cases, the agents are said not to be subject to criticism or not to be blameworthy. Since they violate KRA, it would seem that KRA makes incorrect predictions about these cases. Hence KRA is in trouble.

Before moving on, we would like to point out that, having mentioned cases of selfless assertion, we would like to set them aside for the remainder of this paper. The reason for this is that we are primarily interested in the normative framework from section 1 and the work it can do for champions of KRA. And while, as we are about to argue, the framework does offer an attractive treatment of cases involving assertions of justified false beliefs and gettiered beliefs, it does not serve to deal with cases of selfless assertions. Now, one might wonder whether this does not just mean that KRA stands refuted and that any positive results concerning the other problem cases for KRA the framework may deliver are of negligible interest. Fortunately, the answer to both of these questions is no. The reason why the results of this section are not negligible is that cases involving assertions of justified false beliefs and gettiered beliefs are still widely considered to constitute one of the major stumbling stones for KRA. As a result, any successful defense of KRA will have to provide a workable treatment of them. This section takes on this task. To see that it does not follow that KRA stands refuted, note that even if the normative framework does not serve to handle cases of selfless assertion, all that follows is that it does not solve all of KRA’s problems. It remains entirely open to

champions of KRA to combine the account of assertions of justified false beliefs and gettiered beliefs we will offer below with a different treatment of cases of selfless assertions. And there are a number of promising candidates in the literature.¹⁴ With cases of selfless assertion thus safely set aside, let us take a closer look at cases involving assertions of justified false beliefs and gettiered beliefs.

II.2. The Blamelessness Response. Champions of KRA often argue that speakers who assert justified false beliefs and gettiered beliefs violate KRA but do so *blamelessly*. In particular, they point out that when in doing something one breaks a norm because one reasonably believes that what one does is permissible, then one is blameless. Since this is the case with agents in cases like FAKE SNOW and FAKE BARNS, these agents are blameless. These speakers assert what they assert because they reasonably believe that they know what they assert and, in consequence, that they are permitted to assert as they do. If so, it is reasonable for them to believe that they satisfy KRA. As a result, they are blameless when asserting as they do.

However, there are problems with this account. For instance, it may be that there are unsophisticated speakers who do not even have the concept of knowledge. They, too, may be in cases in which they assert justified false beliefs or gettiered beliefs. Since they do not have the concept of knowledge, they are not in a position to host reasonable beliefs about knowledge in the first place. As a result, the envisaged explanation of why agents in cases like FAKE SNOW and FAKE BARNS are blameless will not work here.¹⁵

At a more general level, foes of KRA worry that the above response is ad hoc. Here is one very clear expression of this worry:

[A] general worry with excuse maneuvers is that they form very generic ways of immunizing proposed norms. Without a principled account of when an agent is excused, every counterexample to a norm may be rebutted by upholding that the agent is excused from violating the norm. As we have seen, the proponents of [KRA] have yet to provide a viable principled account of excusability.¹⁶

The normative framework of section 1 offers exactly what Gerken claims champions of KRA are missing. If the framework predicts that the speakers in cases like FAKE SNOW and FAKE BARNS (as well as their unsophisticated counterparts) are indeed blameless, the ad hoc-ness worry can be laid to rest. The blamelessness response will clearly be available to champions of KRA. Let us ask whether we get the desired results.

First, *Norm-Specific Blamelessness* entails that an agent is blameless relative to a specific norm N for φ -ing if φ -ing violates N but the agent φ s because the agent is blamelessly ignorant that φ -ing violates N. There are a number of ways in which one can be blameless for violating a specific norm in virtue of being blamelessly ignorant that one is doing so. One such way is if one does what one does because of a reasonable belief that what one is doing is permissible. If one believes that what one is doing is permissible, then one does not believe that one is violating a norm. If so, one is ignorant of the fact that one is violating a norm. If one's belief is reasonable, then one's ignorance is blameless. (2.c) is satisfied. And, of course, this is exactly what champions of KRA have claimed is going on in cases like FAKE SNOW and FAKE BARNS. The speakers

assert out of a reasonable belief that they know and so that asserting is permissible. This means that they are blamelessly ignorant for their assertions.

While acting out of a reasonable belief that what one does is permissible is one route to blameless ignorance, it is not the only route. Others are available as well. To see this, recall our Uno case in which you are just blamelessly unaware of the fact that there is a rule requiring you to call Uno when playing your penultimate card. In this case you may have no belief either way on whether you are complying with the rules of Uno.¹⁷ So it is not as if you play in the way you do because of a reasonable belief that this is permissible. Rather, you blamelessly have no belief on the relevant rule of Uno whatsoever. Thus, according to the above framework, another way to blameless ignorance is via a blameless *lack of belief* concerning the relevant rules. And this is of course exactly what we find with cognitively unsophisticated agents, such as agents who do not even have the concept of knowledge to begin with. Any such agent is incapable of even representing KRA. If they are blameless for not having this concept, as they typically will be, they, too, will satisfy (2.c) and so come out blameless for violating KRA.

What comes to light, then, is that the envisaged response by champions of KRA is borne out by the normative framework from section 1. Moreover, the framework also delivers the right results for Gerken's cases of assertions by cognitively unsophisticated agents. As a result, not only is the blamelessness response available to them, the worry of ad hoc-ness is addressed also.¹⁸

III. THE CASE FOR KRA

III.1. An Argument for KRA. It comes to light, then, that the blamelessness response can be made to work for KRA. As a result, cases like FAKE SNOW and FAKE BARNS do not constitute evidence against KRA after all. However, on reflection, the normative framework not only allows champions of KRA to successfully defuse an objection against KRA. Rather, it serves to turn the tables on foes of KRA. Let us explain.

To begin with, recall that the framework distinguishes between norm-specific criticisability and blameworthiness. What we have seen is that, in cases like FAKE SNOW and FAKE BARNS, speakers are blameless for violating KRA in particular. Moreover, in versions of the cases that might constitute a problem for KRA, they are also all-things-considered blameless (see footnote 9). But it is of course entirely compatible with this that speakers in these cases are criticisable for violating KRA. If it can be shown that they are, these cases will confirm KRA and disconfirm rival views according to which the speakers' assertions comply with the relevant norm.

Now, it might be thought that it will be difficult if not impossible to ascertain whether the speakers in these cases are indeed criticisable for violating KRA. On second thought, however, the prospects are not so dim. Given that *Norm-Specific Criticisability* holds, we may expect that we may criticise the agent by pointing out that she violates N. One way of achieving this, in turn, is by saying "You don't ...!" where "..." specifies a requirement of the norm. This gives us

You Don't Criticisms

An agent whose φ -ing violates a specific norm N can be prima facie legitimately criticised by saying “You don’t ...!”, where “...” specifies a requirement of N.

For instance, in our toy case, we can with prima facie legitimacy criticise Uno players who do not call Uno when playing their penultimate card by saying: “You didn’t call Uno!” or “You didn’t say anything!”

Moreover, given that *Norm-Specific Blamelessness* holds, we may expect there to be a number of appropriate responses to “You don’t ...!” criticisms. For instance, the agent may disagree and maintain that she really does satisfy the relevant requirement and so her action is permissible, thus satisfying condition (1). Alternatively, she may offer an explanation to the effect that the norm was overridden by another norm (2.a), such that the action is all-things-considered permissible. We may also expect excuses pointing out (blameless) lack of control (2.b) or ignorance (2.c) to be an appropriate form of response. And given that a norm was violated, it should come as no surprise that apologies constitute an appropriate form of response also.

Specific norms are thus associated with specific kinds of appropriate response. In particular,

Appropriate Responses

Prima facie legitimate “You don’t ...!” criticisms make a certain kind of response on the part of the defendant prima facie appropriate, including (i) “I did ...!” rebuttals, (ii) explanations invoking overriding norms, excuses appealing to (blameless) (iii) lack of control or (iv) ignorance, and (v) apologies.

Let us assume that KRA is true. In conjunction with *You Don't Criticisms* and *Appropriate Responses* this assumption delivers a number of predictions. The first one is that assertions can prima facie legitimately be criticised by saying "You don't know that!", "That's not true!", "You don't believe that yourself!" and so on. Of course, champions of KRA have long pointed out that this is exactly what we find.¹⁹ Interestingly enough, it is also what we find in cases like FAKE SNOW and FAKE BARNS. In FAKE SNOW, Williamson's assertion can prima facie legitimately be criticised by saying "That's false! What you are seeing is just foam. There's a film crew shooting outside." Similarly, in FAKE BARNS, Wendy's assertion can prima facie legitimately be criticised by saying "You don't know that! This part of the country is peppered with fake barns."

Another prediction that we get is that in cases in which a prima facie legitimate "You don't ...!" criticism has been levelled against a speaker, appropriate responses include "I do ...!" rebuttals, explanations in terms of overriding norms, excuses appealing to (blameless) lack of control or ignorance, and apologies. And again this is exactly what we find. "I do know that!" is prima facie appropriate as is "He was going to shoot me if I had told the truth," "I couldn't help it," "I didn't know," and "I'm sorry." For instance, in both FAKE SNOW and FAKE BARNS, an apology and/or an excuse is entirely appropriate. It will be entirely appropriate to respond by saying "I'm sorry. I didn't know that." In this way, the putatively problematic cases for KRA, in conjunction with *You Don't Criticisms* and *Appropriate Responses* serve to confirm the assumption that KRA is true.²⁰

With these points in play, we would like to look at actions that are not constrained by certain normative requirements. In particular, we would like to focus on action types that are not governed by a rule to the effect that one must: ϕ only if one satisfies C . We would like to suggest the following here:

Illegitimate Criticisms

If we criticise an agent who ϕ s without satisfying C , perhaps by a “You don’t ...!” criticism, our criticism will be prima facie illegitimate.

For instance, there is no rule of Uno that one must call *Duo* when playing one’s antepenultimate card. Suppose we are playing a game of Uno and we ventured to criticise your antepenultimate move by saying “You didn’t call Duo!” In this case our criticism will be prima facie illegitimate.

Unsurprisingly, then, what is called for by way of response in this kind of case is not of the kind appropriate in cases of legitimate criticisms. There is no need for you to apologise, offer an excuse or explanation and so on. Rather, you may prima facie legitimately rebut the criticism as misplaced. “So what?” is prima facie a perfectly legitimate response on your part. What we want to suggest, then, is the following:

Legitimate Rebuttals

If we criticise an agent who ϕ s without satisfying C , perhaps by a “You don’t ...!” criticism, the agent can prima facie legitimately rebut our criticism as misplaced or irrelevant, perhaps by saying “So what?”.

Let us assume now that the action type assertion is not governed by KRA. In particular, let us assume that FAKE SNOW and FAKE BARNS are counterexamples to KRA. In this case, we get the following predictions, this time via *Illegitimate Criticisms* and *Legitimate Rebuttals*: criticisms of the speakers for asserting what they do not know are (i) prima facie illegitimate and (ii) can prima facie legitimately be rebutted as irrelevant or misplaced in these cases. Crucially, however, this is just *not* what we find. Suppose, for instance, in FAKE SNOW, we criticise Williamson's assertion by saying: "That's false! What you are seeing is fake snow produced by a film crew that's shooting outside." Our criticism is not prima facie illegitimate in this case. And it cannot be prima facie legitimately rebutted as misplaced or irrelevant. For instance, "So what?" is not a prima facie legitimate response to our criticism. The same is true of FAKE BARNS. Suppose we criticise Wendy by saying: "You don't know that! This part of the country is peppered with fake barns." Again, our criticism is not prima facie illegitimate and cannot be prima facie legitimately rebutted as misplaced or irrelevant. "So what?" is again not a prima facie legitimate response here. The assumption that the above cases do indeed constitute counterexamples to KRA in conjunction with *Illegitimate Criticisms* and *Legitimate Rebuttals* leads to incorrect predictions. As a result, we find this assumption disconfirmed.

III.2. Objections and Replies. Before closing, we would like to briefly consider a couple of data that might be thought to be problematic for KRA.²¹ First, when our assertions are challenged by asking "How do you know?", don't we frequently respond

by citing our evidence for what we asserted? And doesn't that support something like the justification rule for assertion rather than KRA?

By way of response to this worry, note first that the fact that we frequently respond to "How do you know?" challenges by citing our evidence is just what we may expect if KRA is true. After all, given that "How do you know?" constitutes a challenge to a speaker's assertion, the speaker would want to provide some evidence that she knows what she asserted. But of course, one excellent way of achieving this is by offering her evidence for what she asserted. Suppose, for instance, that you assert that Obama cried during his speech on gun violence in January 2016, and we challenge your assertion by asking how you know. If KRA is true, your response to our challenge had better produce some evidence that you know that Obama cried during the speech in question. One excellent way of providing such evidence is by citing your evidence for it, say that you saw him crying on a TV broadcast of the speech. But given that citing our evidence is an excellent way of producing the kind of evidence "How do you know?" challenges call for, it will come as no surprise that we should frequently do so in response to such challenges.²²

What is more, it may be worth noting that "How do you know?" challenges are standardly taken to support KRA.²³ After all, these challenges do presuppose that the speaker knows. If KRA were not in place, it is hard to see why this presupposition should be legitimate and why assertions should be challengeable in this way to begin with. In view of these considerations, we do not think that the fact that we frequently respond to "How do you know?" challenges by citing our evidence disconfirms KRA.

But perhaps the point is a slightly different one, to wit, that the fact that we frequently respond to “How do you know?” challenges by citing our evidence provides reason to believe that, at the end of the day, “How do you know?” challenges are really nothing more than “What’s your justification?” challenges. That might be a bit more troublesome for champions of KRA as, on this assumption, all that may be presupposed by these challenges is that the speaker has justification for what she asserted.

Fortunately, there is reason to believe that even this way of interpreting the objection will not refute KRA. To see why, note first that since knowledge entails justification, on KRA, we may expect that we can challenge assertions by challenging the speaker’s justification. In order to turn this into a genuine problem for KRA, it would additionally have to be shown that assertions cannot be challenged in any of the other ways KRA would predict. And this is exactly not what we find. On the contrary, we can also challenge assertions by asking: “Is that really true?” or “Do you really believe that?” Moreover, we can also challenge them by querying whether the speaker’s epistemic environment is not inhospitable. For instance, Wendy’s assertion that there is a barn in the field can be challenged by saying: “But isn’t this the part of the country that contains nearly only fake barns?” (In fact, given that assertions can be criticised by saying “That’s false!” and so on, it is unsurprising that it should also be possible to challenge them in these ways.) So, even if “How do you know?” challenges did ultimately boil down to “What’s your justification?” challenges, this fact again does not serve to disconfirm KRA.

Here is another potentially troublesome datum for KRA. Suppose, before there was widespread consensus among expert scientists on the reality of global warming, you

assert that global warming is happening and we criticise your assertion by saying: “You don’t know that! There still is a considerable number of scientists who would disagree.” Now suppose you reply along the following lines: “Well, okay... But still there is good reason to believe that it is.”²⁴ Isn’t this exactly the kind of “I do...” response that, according to *Appropriate Responses*, constitutes a legitimate rebuttal of our criticism? If so, KRA is in trouble. After all, as you also concede, you do not know that what you assert. If, nonetheless, you can legitimately rebut our criticism here, there is evidence that permissible assertion does not require knowledge, that is, that KRA does not hold.

Fortunately for champions of KRA, there is reason to think that your response does not constitute a legitimate rebuttal of our criticism. To see why, note first that assertions that go unchallenged become part of the common ground of the conversation afterwards, as do assertions such that any challenge/criticism has been successfully met. If so, and if, in the above exchange, your response to our criticism is indeed a legitimate rebuttal, we may expect that the proposition that global warming is happening to become part of the conversation’s common ground thereafter, at least assuming, as we may, that neither the original assertion nor the subsequent rebuttal are subject to any further challenges or criticisms. Crucially, however, this is not what is happening here. To see this, note that (i) subsequent assertions in the conversation cannot be defended by referring back to the proposition that global warming is happening, not even if they are obviously entailed by it. Moreover, (ii) attempted such defences can themselves be legitimately rebutted. For instance, suppose at a later stage you were to assert: “Temperatures on Earth are on the rise.” Suppose that we also criticised this assertion and that you attempted to defend your assertion by referring back to the earlier assertion that

global warming is happening, for instance by saying: “But didn’t we already say that global warming is happening?” We could rightly point out here that the answer is negative: “No, all that we said was that there is good reason to believe that global warming is happening.” There is thus reason to believe that the proposition that global warming is happening does not become part of the common ground after our exchange. But if your response to us had been a legitimate rebuttal, it would have done so. As a result, there is reason to believe that your response did not constitute a legitimate rebuttal.

Finally, there is another, more attractive interpretation of what is going on in our exchange. Rather than constituting a legitimate rebuttal of our criticism, your reply involves both a concession “Well okay...” and subsequent weakening “There is good reason to believe that global warming is happening” of the assertion. This interpretation avoids the problem of its competitor. After all, it at most predicts that the proposition that there is good reason to believe that global warming is happening becomes part of the conversation’s common ground, which seems unproblematic. At the same time, this interpretation does not generate a problem for KRA. This is because if we drop the claim that your response constitutes a legitimate rebuttal of our criticism, we no longer have any reason against KRA’s key claim that permissible assertion requires knowledge. On the more plausible interpretation of the case, then, KRA walks free once again.

CONCLUSION

In this paper, we have developed a normative framework featuring detailed accounts of both norm-specific and all-things-considered criticisability, to wit, *Norm-Specific Criticisability*, *Norm-Specific Blamelessness*, *All-Things-Considered Blamelessness*, and

All-Things-Considered Criticisability/Blameworthiness. We then turned to a specific type of action, the speech act of assertion. In particular, we focused on a popular proposal concerning the epistemic norm of assertion, to wit, KRA, and a prominent objection to this view. The thought here is that cases of assertions on justified false belief and on gettiered belief show that KRA is too strong as the speakers violate KRA whilst, intuitively, walking free of blame. We have shown that our normative framework predicts that, whilst the speakers in these cases do indeed violate KRA, they do so blamelessly. In this way, we have shown that the standard response that champions of KRA have offered when presented with these counterexamples is borne out by our normative framework. As a result a prominent worry concerning the ad hoc-ness of this response can be dispelled.

Finally, we have argued that the putative counterexamples not only fail to provide evidence against KRA. On the contrary, once we take into consideration the patterns of criticisms and responses characteristically associated with action types that are/are not governed by rules like KRA, we find that these cases confirm KRA. In this way, the argument from these cases that foes of KRA have run against KRA can be turned on its head. Rather than posing a threat for KRA, they serve to strengthen the case for KRA.

NOTES

¹ Recent literature in epistemology witnesses some interest in this issue and some steps in the direction of a worked out account are being taken as we speak. Notably, Timothy Williamson, “Justification, Excuses, and Sceptical Scenarios,” in Julien Dutant and Fabian Dorsch, eds., *The New Evil Demon: New Essays on Knowledge, Rationality, and Justification* (Oxford: Oxford University Press, forthcoming) and Clayton Littlejohn, “A Plea For Epistemic Excuses,” in Julien Dutant and Fabian Dorsch, eds., *The New Evil Demon: New Essays on Knowledge, Rationality, and Justification* (Oxford: Oxford University Press,

forthcoming) both offer conditions for blamelessness. However, their conditions are at best sufficient. Neither offers a fully-fledged normative framework of the kind we aim to give in this paper, which includes both necessary and sufficient conditions for blamelessness relative to a single norm and all-things-considered blamelessness.

² Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000).

³ Champions of KRA include Peter Unger, *Ignorance: A Case for Scepticism* (Oxford: Clarendon, 1975), Michael Slote, "Assertion and Belief," in Jonathan Dancy, ed., *Papers on Language and Logic* (Keele: Keele University Press, 1979), pp. 177–90, Keith De Rose, "Assertion, Knowledge and Context," *Philosophical Review*, CXI, 2 (2002): 167–203, John Turri, "The Express Knowledge Account of Assertion," *Australasian Journal of Philosophy*, LXXXIX, 2 (2011): 37–45, John Turri, "Knowledge and Suberogatory Assertion," *Philosophical Studies*, CLXVII, 3 (2014): 557–67, and most famously Timothy Williamson "Knowing and Asserting," *Philosophical Review*, CV, 4 (1996): 489–523, and Williamson, *Knowledge and its Limits*, *op. cit.* Both of us have also defended KRA elsewhere: Christoph Kelp, "A Practical Explication of the Knowledge Rule of Informative Speech Acts," *European Journal of Philosophy*, XXI, 3 (2013): 367–383, Christoph Kelp "Assertion: A Function First Account," *Noûs* (forthcoming), Mona Simion, "Assertion: Knowledge is Enough," *Synthese* (forthcoming). According to KRA's main rival, permissible assertion requires only knowledge-level justification (Igor Douven, "Assertion, Knowledge, and Rational Credibility," *Philosophical Review*, CXV, 4 (2006): 449–485, Igor Douven, "Assertion, Moore, and Bayes," *Philosophical Studies*, CXLIV, 3 (2009): 361–75, Jennifer Lackey, "Norms of Assertion," *Noûs*, XLI, 4 (2007): 594–626, Jennifer Lackey, *Learning From Words. Testimony as a Source of Knowledge* (Oxford: Oxford University Press, 2008)) or knowledge-level justified belief (Jonathan Kvanvig, "Assertion, Knowledge, and Lotteries," in Patrick Greenough and Duncan Pritchard, eds., *Williamson on Knowledge* (Oxford: Oxford University Press, 2009), pp. 140–60, Jonathan Kvanvig, "Norms of Assertion," in Jessica Brown and Hermann Cappelen, eds., *Assertion: New Philosophical Essays* (Oxford: Oxford University Press, 2011), pp. 233–250. Further rival accounts include the truth rule (Matthew Weiner, "Must We Know What We Say?" *Philosophical Review*, CXIV, 2 (2005): 227–251, Daniel Whiting, "Stick to the Facts: On the Norms of Assertion," *Erkenntnis*, LXXVIII, 4 (2013): 847–867) and the belief rule (Kent Bach, "Applying Pragmatics to Epistemology," *Philosophical Issues*, XVIII, 1

(2008): 68–88, Frank Hindriks, “The Status of the Knowledge Account of Assertion,” *Linguistics and Philosophy*, xxx, 3 (2007): 393–406).

⁴ Roderick Chisholm, “The Ethics of Requirement,” *American Philosophical Quarterly*, 1, 2 (1964): 147–153.

⁵ ‘Because’ here is crucial to blamelessness. See Harry Frankfurt, “Alternate Possibilities and Moral Responsibility,” *Journal of Philosophy*, LXVI, 23 (1969): 829–839.

⁶ See for example Ishtiyaque Haji *Moral Appraisability* (Oxford: Oxford University Press, 1998) and Michael Zimmerman, “Moral Responsibility and Ignorance,” *Ethics*, CVII, 3 (1997): 410–426. However, that ignorance excuses was already recognised by Aristotle in the *Nicomachean Ethics*, Aristotle, *The Nicomachean Ethics* (Indianapolis: Hackett Publishing Co, 1985), at 1110a-1111b4.

⁷ Cf. Zimmermann, “Moral Responsibility and Ignorance,” *op. cit.* on the distinction between direct and indirect blameworthiness.

⁸ This is also why we would not want to merge *Norm-Specific Criticisability* and *Norm-Specific Blamelessness* into one principle, where the latter specifies a defeat condition for the former. After all, that would mean that whenever someone is blameless for violating a specific norm, we will no longer be able to legitimately criticise her for violating that norm. Since, as we have just seen, it is useful for us to retain the right to criticise violations of a norm even when it is broken blamelessly, we will do well not to merge the two principles in the way envisaged. Thanks to an anonymous referee for pressing us on this point.

⁹ Lackey, “Norms of Assertion,” *op. cit.*, at p. 595.

¹⁰ Douven, “Assertion, Knowledge, and Rational Credibility,” *op. cit.*, at p. 477.

¹¹ Williamson, “Knowledge and Its Limits,” *op. cit.*, at p. 257.

¹² Lackey, “Learning from Words,” *op. cit.*, at p. 544.

¹³ Lackey, “Norms of Assertion,” *op. cit.*, at p. 599.

¹⁴ For instance Ernest Sosa, “Value Matters in Epistemology,” *Journal of Philosophy*, CVII, 4 (2010): 167–90, E.J. Coffman, “Lenient Accounts of Warranted Assertibility,” in Clayton Littlejohn and John Turri, eds., *Epistemic Norms. New Essays on Action, Belief, and Assertion* (Oxford: Oxford University Press, 2014), pp. 33–58, John Turri, “You Gotta Believe,” in Clayton Littlejohn and John Turri, eds., *Epistemic Norms. New Essays on Action, Belief, and Assertion* (Oxford: Oxford University Press, 2014), pp. 193–200

and John Turri, “Selfless Assertions: Some Empirical Evidence,” *Synthese*, CXCII, 4 (2015): 1221–1233 (2015). In fact, one of us has offered a principled solution to the problem posed by these cases elsewhere (Kelp, “Assertion: A Function First Account,” *op. cit.*), one that drops out of an independently motivated account of the normativity of assertion. For those interested, here is the briefest sketch of the general argumentative strategy pursued in that paper. Its key aim is to offer an account of the normativity of assertion that derives from its function. In particular, it is argued that assertion has the etiological function of generating knowledge in hearers. Etiological functions have normative import. In particular, they give rise an evaluative norm according to which a token of type T with the etiological function of producing effect E is a good T if and only if it has the disposition to produce E by functioning normally when in normal conditions. Since the etiological function of assertion is to generate knowledge in hearers, we get the result that a good assertion is one that has the disposition to generate knowledge in hearers by functioning normally when in normal conditions. It is easy to see that, even though selfless assertions violate KRA, they are still good assertions on this account, which accommodates the problematic intuition. It may also be worth noting that the same is not true of assertions of justified false beliefs and of gettiered beliefs. That is why the present paper is still essential towards a successful defense of KRA.

¹⁵ Mikkel Gerken, “Warrant and Action,” *Synthese*, CLXXVIII, 3 (2011): 529–47, in sections 4–5.

¹⁶ *Ibid.*, at p. 544.

¹⁷ If this is not immediately obvious, note that you may know that you have not been filled in on all the rules of the game yet and are taught various rules as you go along.

¹⁸ Finally, note that we have so far argued that agents in cases like FAKE SNOW and FAKE BARNS are blameless for violating KRA in particular. One might wonder whether they are also all-things-considered blameless. As a first observation, note that even though the question does indeed arise, the key point for champions of KRA is surely that speakers in the problem cases violate KRA blamelessly. With this point in play, there is every reason to believe that, on any version of the relevant cases that stands a chance of working as counterexample to KRA to begin with, the speakers come out as all-things-considered blameless as well. After all, it is important that speakers in these cases are intuitively not deserving of criticism or are blameless for their assertions. However, if speakers are indeed not deserving of criticism/are indeed blameless, the cases had better be described such that speakers do not violate any other

norms governing the assertion. If they did violate other norms of assertion, the key intuition that speakers are not deserving of criticism/are blameless will be compromised. Given that this is so, on any version of the cases such that they are in the ballpark for constituting counterexamples for KRA, speakers will turn out not only blameless relative to KRA, but also all-things-considered blameless.

¹⁹ Williamson, "Knowledge and Its Limits," *op. cit.*

²⁰ Again, it may be worth pointing out that it makes perfect sense for us to allow for criticisms of assertions that fall short of knowledge. Assertions have an etiological function, to wit, the function of generating knowledge in others (Kelp, "Assertion: A Function First Account," *op. cit.*, recall also fn.7). It is of some importance that assertions fulfil this function reliably. If they are too unreliable at generating knowledge in hearers, they will stop generating knowledge in hearers altogether. Allowing for criticisms of unknown assertions contributes to ensuring the reliability of assertion's fulfilling its etiological function. To see this, notice first that, in the vast majority of cases, an assertion will generate knowledge in hearers only if the speaker also knows what she asserts. In other words, cases in which assertion generates knowledge are the exception and cases in which it transmits it are the rule. Moreover, hearers who are otherwise prepared to form a testimonial belief on the basis of a speaker's say-so will refrain from doing so when the speaker has been criticised for asserting what she does not know. When Williamson asserts that it is snowing outside and you criticise his assertion in the way just outlined, I will not form a belief that it is snowing outside even though I otherwise might have. Your criticism has prevented me from forming a belief that falls short of knowledge, thus contributing to ensuring that assertion reliably fulfils its etiological function of generating knowledge in hearers. See Kelp, "Assertion: A Function First Account," *op. cit.* for a detailed explanation of why assertion should be governed by KRA in terms of this etiological function.

²¹ Thanks to an anonymous referee for bringing these objections to our attention.

²² Incidentally, we are not so sure just how frequently we actually do respond to "How do you know?" challenges by citing our evidence. To see this, note that one perfectly legitimate response you may give in the Obama case is that you saw the speech on TV. It is hard to deny, however, that the fact that you saw Obama give a certain speech on TV is not evidence for the proposition that Obama cried during that speech. Rather, your evidence for this is that you saw him crying during the speech. But that is not what you offer in response to the challenge. What is going on here is that you cite the source of your belief that

Obama cried—that is, seeing a TV broadcast of the speech—implying that it is a source of knowledge for the proposition in question. These considerations suggest that we do not need to cite our evidence in order to successfully rise to a “How do you know?” challenge. How frequently we actually do cite our evidence seems to us to be an open empirical question.

²³ For instance, Williamson, “Knowledge and Its Limits,” *op. cit.*, at p. 252.

²⁴ Cf. Douven, “Assertion, Knowledge, and Rational Credibility,” *op. cit.*, at p. 468, fn. 26.