Trolling as speech act (or, the art of trolling, with a description of all the utensils, instruments, tackling, and materials requisite thereto: With rules and directions how to use them)

P. J. Connolly

Department of Philosophy, University of Sheffield, Sheffield, UK

Correspondence: P. J. Connolly, Department of Philosophy, University of Sheffield, Sheffield, UK. Email: pjconnolly1@sheffield.ac.uk

1 INTRODUCTION

The acceptance of trolling as a feature of internet communication and the clustering together of behaviors that range from playfulness to abusive bullying under the term “trolling” has interesting moral implications. For it seems that sometimes such behavior is dismissed as “mere trolling” and consequently the trolls who perform these acts are granted a partial exemption from responsibility for the content or consequences of their speech. A result of this is that some acts which might ordinarily be considered bullying, abuse or threats when uttered outside of the internet can be dismissed as mere trolling when carried out on the internet. One popular response to this is to categorize acts of trolling based on perceived intentions or to argue that, for more virulent acts particularly, “trolling” is not an appropriate label to be applied. Although I agree that, for example, there should be no special exemption granted for racist speech on the basis that it is used in an act of trolling, I argue here that we also should not ignore the structural similarities of the range of speech acts often described as “trolling.” Application of tools from the philosophy of language and consideration of an act of trolling as a complex speech act can make clear some of these similarities. Once these similarities are made clearer we can then start to understand a little more about what trolling is. I suggest that central to acts of trolling is a notion of seriousness and its interplay between the troll, their target, and any onlookers to the act of trolling. By viewing trolling in this way we can understand better the rhetorical nature of an act of trolling and why we can often end up facing what I call the trolling dilemma—that is, in many acts of trolling the directions of reply are either to respond seriously to an utterance and so be trolled, or to acknowledge it as an unserious act of trolling and so dismiss it as mere trolling.
Though far from a complete account of the speech act of trolling, what I present here are the grounds for understanding trolling as a systematically divisive act. In an act of trolling the audience targeted by the troll is a mere prop to be manipulated for the entertainment of the troll and their onlookers. This in-built division between different audiences makes trolling a particularly powerful political device, and the trolling dilemma makes responding to trolling difficult. So when we consider that the act itself is essentially divisive and difficult to respond to effectively, we can start to understand why regimes and movements which thrive on division have successfully embraced trolling as a propaganda tool.

2 | TROLLING PRELIMINARIES

Some terminological notes to begin. *Trolling* is carried out by *trolls* on *targets*. A particular target need not necessarily be pre-determined by the troll but is simply anyone who responds in a way amenable to the act of trolling. The target is an audience that trolling is performed to, but an act of trolling has an audience (which includes the troll) it is performed for. I call this audience *onlookers*.

2.1 | Trolling

Trolling is a complex linguistic behavior and the usage of “trolling” is still in flux. As such, defining it both pithily and satisfactorily is difficult. This is an occasion, though, when etymology is perhaps helpful. “Trolling” originally referred to a method of fishing whereby bait is placed on a hook, the hook is dangled from a boat, and the angler gently moves the boat downstream waiting for a fish to bite. With a few slight changes, we can see how internet trolling is similar. The bait is generally a provocative comment, the hook the social media it is written on, and the troll leaves it dangling waiting for a target to bite.

Trolling is not an entirely new type of behavior either. Expressions such as “baiting,” “codding,” “winding-up,” “teasing,” and “goading” probably capture similar acts (the first two or three are even plausibly fishing-derived terms too). Just like “trolling,” those terms can also sometimes refer to what appears to be amusing, innocuous, or playful behavior, but also cases of abuse and harassment. As I argue later, trolling bears relation to certain types of humor, and it is in these terms we might best characterize it. What we might think makes trolling worthy of close attention, however, is the scale of its proliferation and successes as a communicative tool. It is a conspicuous feature of internet discourse, an implement of corporate marketing and, most importantly, it has become a tool used for disseminating political ideas.

2.2 | Examples

The following three examples have all been described as trolling. They have been chosen for their differences, and to reflect some of the range of acts that have been termed “trolling.” I present them here only with some relevant contextual information and analysis follows in subsequent sections.

2.2.1 | KenM

Kenneth McCarthy posts intentionally confused messages on the internet under the username KenM and is generally regarded as a troll. McCarthy has spoken publicly about his trolling and describes it
as “playing a well-meaning moron on the internet.” He claims that he does it primarily as a reaction to the harshness of internet discourse. For example, (T1) was posted under a 2015 article titled “Clinton, Bush fundraising steady amid GOP summer Trump slump.”

(T1):

Ben Franklin said politicians are like pampers, they both stink and they act like babies.

The comment received the following response from user Karl: “ken hate to tell you pampers have only been around 50 years. True moron.”

2.2.2 | Donald Trump

On July 14, 2019, the then president Donald Trump posted the following comment on social media platform Twitter.

(T2):

So interesting to see “Progressive” Democrat Congresswomen, who originally came from countries whose governments are a complete and total catastrophe … Why don't they go back and help fix the totally broken and crime infested places from which they came.

This references four non-white US congresswomen, Alexandria Ocasio-Cortez, Rashida Tlaib, Ayanna Pressley and Ilhan Omar. There are numerous ways in which one might view (T2) as racist, for current purposes though most striking is that it provides a rendition of the “send them back to their own countries” rhetoric eternally popular with racists. Although some of the media responses to Trump’s comment condemned it as racist, interestingly some of these same critics noted he was also trolling. In one report the Guardian said “The presidential trolling may have been meant as a distraction from immigration raids that were due in major cities on Sunday.” An LA Times editorial on the day of the tweets also said; “He is just trolling, as usual. He is just trying to get a rise out of us. He is baiting us.”

2.2.3 | The RIP troll

The following example is a case of so-called “RIP trolling,” and was posted under a YouTube video about the 2011 Christchurch earthquake.

(T3)

I and the rest of the world are pleased your piece of shit family … are dead squashed filthy shit rotting in the ground. Especially those two filthy babies that were squashed

REST IN PISS YOU FUCKING RODENT PIECES OF SHIT.
RIP trolls target articles, videos, or memorial pages relating to tragedies and deaths. Generally, it is supposed that RIP trolls post such comments for the amusement of other RIP trolls. The motivations for RIP trolling and the “humor” it evokes will be touched upon in Section 4.3.12

3 | TROLLING, SINCERITY, AND SERIOUSNESS

3.1 | Academic work

There is some call in academic discussion for the need to distinguish between different types of trolling, for example classifying some acts as “kudos trolling” and others as “flame trolling.” So whereas the “kudos troll” provokes for the entertainment of others, the “flame troll” is said to be simply carrying out the abuse that is not intended to be humorous. Without knowing the context from which the examples above come from, we might naively think that (T1) is a case of kudos trolling, and (T3) a case of flame trolling, for example. It is plausible that (T3) could, on the face of it, appear to be a case of simple abuse, but as previously mentioned (Section 2.2.3), and as will be discussed further in Section 4.3, even (T3) has a component of entertaining an audience. There are clear merits to drawing distinctions between particular acts of trolling in the sense that some acts of trolling are certainly more offensive or damaging than others. However, there is a fundamental problem with trying to draw such a distinction; we might think that any troll is at least, in some way, entertaining themself, and beyond a warped version of utilitarianism I do not see any good moral distinction between abusing someone by saying P to entertain only oneself, on the one hand, and saying P to entertain oneself and some others, on the other. As such I think it very plausible to say that with every act of trolling there is at least one person the troll intends to entertain—themself. When analyzing trolling the speech act, therefore, I would not distinguish between acts of trolling in such a way.

There are other arguments we find that suggest that labeling abusive behavior as mere trolling is dangerous as doing so suggests that online abusers are mere trolls, rather than, in many cases, anti-social criminals, and so a term such as “e-bile” or “cyberabuse” might be a more apt name. I agree with the aims of such a project; there should be no moral redemption for an act of abuse simply because it was deemed an act of trolling. However, I will argue that paying attention to similarities between the seemingly disparate acts that get called “trolling” is also an important part of the task of avoiding dismissing abuse as mere trolling. By viewing trolling as a type of speech act with some unifying properties, we can understand how it operates as a feature of conversation and how it has spread. We can also understand that there is nothing “mere” about an act of trolling. It is an intrinsically divisive and manipulative act, regardless of the specific goals of the troll. This does not entail an act of trolling is necessarily morally bad, manipulation is sometimes justifiable after all. The interesting consequence is that abuse should not be considered somehow more acceptable as a component of an act of trolling (“mere trolling”); it is in fact even more questionable when used in such a way. Once we see that an act of trolling is by nature divisive and manipulative, an act of abusive trolling is, therefore, better thought of not as mere trolling, but as abusive and divisive and manipulative. Just as if I do something wrong and lie about it I would be held liable for two wrongs should the lie be discovered (the lie and the act I lie about), so too using racist speech, for example, does not somehow become more acceptable when used as a means of trolling. The moral implications, therefore, are cumulative not diminishing.
3.2 | Defining “troll”

The next task, then, is to find commonality between these seemingly disparate trolling behaviors. Claire Hardaker\textsuperscript{15} offers a well-cited definition of what it is to be a troll;

\textit{(HT)}

A troll is a [computer mediated communication] user who constructs the identity of sincerely wishing to be part of the group in question, including professing, or conveying pseudo-sincere intentions, but whose real intention(s) is/are to cause disruption and/or to trigger or exacerbate conflict for the purposes of their own amusement. (237)\textsuperscript{16}

Although useful as a starting point, this is too restrictive for current purposes. Consider the examples (T2) and (T3) above, Trump nor the RIP Troll appear to be making any attempt to construct an identity of wanting to be a part of a group they wish to disrupt, yet both have been described as trolling. Ordinary language usage of “trolling” has changed since Hardaker proposed (HT), and it no longer seems correct to suggest that a troll attempts to construct an “identity of sincerely wishing to be part of the group,” so I’ll remove this from consideration. Hardaker also defines the troll rather than the act of trolling, my focus is on the speech act of trolling and so I adjust accordingly. The main point of departure, however, will be to move from talk of “sincerity” to talk of “seriousness,” I discuss this next.

3.3 | Sincerity and seriousness

As Hardaker's definition above seems to suggest, and as we might expect considering the role sincerity often plays when analyzing duplicitous behavior, it may be thought that the role of sincerity is essential for an understanding of trolling. As mentioned though, I do not want to discuss trolling in terms of sincerity. The main reason for this being that when thinking about sincerity it seems right that our focus is on the relationship of the speaker to utterance.\textsuperscript{17} However, what is most interesting about trolling is not the relationship the speaker has to her utterance (not directly, at least), but rather it is the relationship the speaker has to her intended audiences.\textsuperscript{18} More specifically, the main point of interest is how a troll intends her multiple audiences to take her utterance.

A trolling utterance is generally placed on widely accessible internet-based social media platforms; it is most often a type of public activity. As with most internet comments, it is therefore potentially viewable by a wide audience. As is suggested in Section 2, in an act of trolling the audience can be crudely divided into target and onlooker. Dividing the audience in such a way means the targets are the audience the troll intends to troll and the onlookers (including the troll) are the audience the troll performs the act of trolling for. If we think it seems plausible to presume that broadly speaking the aim of a troll in an act of trolling is to provoke a certain type of response from the target and a particular type of response from the onlooker, then it will turn out that it is not of direct importance whether a trolling utterance is made sincerely or not. Though this is not to say that on inspection of individual acts of trolling (and their case-specific moral ramifications) that this utterer-utterance relationship should not be a consideration. The point is simply that a troll can sincerely utter \textit{u} intending to troll a target, but a troll can just as plausibly insincerely utter \textit{y} intending to troll a target. In an act of trolling the troll can say something she believes to be true, something she believes to be false, or even something she has no particularly strong beliefs about—it does not matter so long as she gets the
response(s) she is aiming for. This type of relationship to the truth of one's own utterance is perhaps closely related to Harry Frankfurt's description of the bullshitter, of which he says.

[T]he motive guiding and controlling [his speech] is unconcerned with how the things about which he speaks truly are … He does not care whether the things he says describe reality correctly. He just picks them out, or makes them up, to suit his purpose.19

Whereas in the case of Frankfurt's bullshitter the purpose is to convince an audience of something, we might think that in the case of trolling the primary purpose is to provoke a particular type of action from a target. So we might say that for the purposes of an act of trolling a troll is indifferent to their utterance beyond its utility within the act of trolling. Let us next consider the first two examples in this light.

(T1):

Ben Franklin said politicians are like pampers, they both stink and they act like babies.

Knowing as we do that KenM is a contrived character, we might expect that the author of (T1) did not utter it sincerely. He does not genuinely believe Franklin said politicians are like pampers.20 On many accounts of sincerity, then, McCarthy uttered (T1) insincerely, and yet this is generally considered a case of trolling, it seems, however, that he is still aiming for a particular type of response. Consider too the second part of the Trump example (T2).

(T2a):

Why don't they go back and help fix the totally broken and crime infested places from which they came.

It seems plausible that Trump is making a sincere statement in (T2a). We might have good reasons to think he may genuinely wish the congresswomen would “go back” to the places “from which they came.” Whether he sincerely thinks this or not, though, is not salient to understanding it as an act of trolling. As an utterance used in an act of trolling it is the intended response(s) that is of most importance. Whether or not Trump is sincere in uttering (T2a) is secondary to the fact that he did utter it, and the nature of what it is he has said (and the use of racist tropes) makes a particular type of response likely. With this in mind, what I suggest then is that the particular type of response both KenM and Trump are aiming at is related to the notion of seriousness. I use this notion of seriousness to help characterize how a speech act of trolling works and so next I'll offer a few thoughts as to what I take this to be. A good starting point is to consider the seriousness and its relationship to joking.21

Say S makes a joke to an audience, conventionally we might expect that in making the joke, S intends to make her audience laugh.22 It could plausibly be the case that S sincerely believes the joke expresses some truth, and as the clichés suggest, many a true word is said in jest, and sometimes something is funny because it's true. However, there is certainly no requirement that a joke is of this nature. Satire, parody, irony, sarcasm, and absurdity are all well-worn jocular devices, and by their nature, none of these require that an utterance is made sincerely.23 Take irony, for example. Traditionally philosophers have treated irony as saying P to convey something like not-P. A speaker, therefore, can make a joke by making some utterance sincerely or insincerely.24 So although sincerity might be
salient when discussing specific jokes, for a general account it is of little interest. What I suggest is of interest when thinking about jokes, in general, is the way in which the audience takes the utterance rather than the sincerity or insincerity of the joker. And this is the case with trolling. Of importance when considering trolling, in general, is not whether the troll is sincere, but rather the seriousness with which she intends her audiences to take her utterance.

I will return to the discussion of seriousness in Sections 4.2 and 4.2, next though I speak more specifically about trolling and the relationship of troll to onlooker and troll to target. What I want to say is that in uttering a troll intends that the target take it to likely be a serious utterance, and that the onlooker takes u to likely be an unserious utterance.25

4 PERLOCUTIONARY INTENTIONS

4.1 Perlocutionary acts

As my interest is in trolling as a speech act, and as Section 3.3 discusses the role of intended audience responses (under the guise of seriousness) it seems apt to next consider some of the foundational work on speech act theory, specifically what J.L. Austin says of perlocutions. Austin says; “[s]aying something … produce[s] certain consequential effects upon the feelings, thoughts, or actions of the audience … and it may be done with the design, intention or purpose of producing them.”26 Perlocutionary acts, then, aim at certain effects, and as Austin says later, “‘perlocutionary act[s] always include some consequences’; and are ‘what we bring about or achieve by saying something, such as convincing, persuading, deterring” (1962, 108). So if I were to persuade you of something by uttering u it requires that you are persuaded by u. The perlocutionary effect is that you were persuaded, and my act in persuading you was the perlocutionary act. François Recanati27 refers to the intention to produce these types of responses as a perlocutionary intention, and so a perlocutionary intention is an intention that a perlocutionary effect is produced by some utterance. Considering (HT) and what has been said so far about seriousness, this seems to correlate with an act of trolling. Trivially speaking, a troll will naturally have perlocutionary intentions when making a trolling utterance, and these can be numerous and diverse even within a single act. We cannot hope to account for all such intentions so what I aim to do next is to suggest some primary intentions common to acts of trolling.

Typically perlocutionary effects are considered to be such things as convincing, deterring, surprising and misleading. In the case of a specific act of trolling, a troll's perlocutionary intentions might, for example, be to cause disruption, anger, confusion, or to fluster or unnerve a target in some way. However, this will be case-specific, and indeed even specifying one of these perlocutionary intentions for a single act of trolling might prove difficult, for a troll could very plausibly be aiming for any or none of these responses. As such, we need to take a step away from these more traditional types of perlocutionary effect and think of something more general. That it is more general, though, is not to say it is less complex, indeed what I propose is a multifaceted notion. It seems to be that there are two distinct sets of perlocutionary intentions attached to an act of trolling which relate to the trolls relationship to her multiple intended audiences. Recall Section 3.3, there trolling is discussed in terms of two distinct audiences, the first is the target of the act of trolling (the fish the troll intends to catch) the second (which includes the troll) is the onlooker (the group that the act of trolling is performed for the benefit of). Let’s look now, then, at how the relationship between the troll and these two audiences differ by elaborating on the corresponding sets of perlocutionary intentions.28
4.2 Perlocutionary intentions: Troll-target

The first set of perlocutionary intentions relates to troll and target. We might expect this first set includes any (or more, or less) of the above mentioned traditional types of perlocutionary intentions (depending on the specific act), plus the effect that the target takes the troll’s utterance to be taken seriously (another way we might put this is to say it be taken to be worthy of serious consideration or to be taken as a serious statement). Regardless of the case-specific contents of this set, it will also include the perlocutionary intention that the target act in a specific way. To put this another way, although some specific trolls will likely have a number of specific perlocutionary intentions, each of these intentions is secondary to the primary perlocutionary intention to provoke a target into a certain type of response. This certain type of response, I suggest, is that a target takes the trolling utterance to be an utterance requiring serious attention and act accordingly. That is, the intention is for a target to respond to the trolling utterance as if it were a serious statement worthy of a serious response. This does not, however, preclude that a target might be uncertain as to the level of seriousness with which they should take the trolling utterance. A target may strongly suspect that the troll is indeed a troll and be suspicious that the troll’s intention is to make some target respond in such a way, yet even suspecting this to be the case does not entail they would not respond as if the troll has made a serious statement. Take the Trump example (T2), many of those he targets might assume he’s trolling— that he has some intention to make them respond in a way that entertains him and his onlooker audience—yet they might still respond in a way amenable to his trolling.

4.3 Perlocutionary intentions: Troll-onlooker

The second set of perlocutionary intentions relates to troll and onlooker. The intentions in this set might include entertaining, amusing, or provoking laughter, for example. It's not immediately obvious that the troll has any specific action-based perlocutionary intentions as is the case with the troll-target set and an onlooker need not play an active role in a trolling interaction. What we can say, though, is that included in this set of perlocutionary intentions is that the onlooker does not take the trolling utterance to be a serious statement in the way the target does. The intention of the troll is that an onlooker view their utterance as something like a joke, or an unserious utterance. This in itself adds a layer of complication that can also help us distinguish seriousness from sincerity. For even if an onlooker takes the utterance to be unserious, as they might do a joke, this is not to say that they do not endorse it or agree with it. Let's consider (T2a) again.

(T2a):

Why don't they go back and help fix the totally broken and crime infested places from which they came.

It's certainly plausible that some members of the onlooker audience will believe Trump uttered (T2a) sincerely, and they themselves might even endorse the view expressed in the tweet. However, even if this is the case, it is not incompatible with them taking the utterance to be unserious. Consider again the correlation between joking and unseriousness. A joke can be taken to be amusing by an audience for a range of reasons. A joke might strike us as humorous because it is a shrewd observation, or surprising, or cleverly constructed, for example, but there is also an element of humor that develops from the confounding
of social expectation or convention—jokes often play a taboo-breaking social role. It is in this socially confounding way that the onlooker takes some utterance to be unserious. Recall, then, that in Section 3.3 it was suggested that seriousness, unlike sincerity, does not take as its locus the relationship between speaker and utterance, but rather the focus is on the relationship of utterance and audience. So although Trump may well sincerely utter (T2), and some onlooker may endorse his view, there is also no requirement for an act of trolling that either do endorse it, all that is required is the recognition of the way the utterance functions socially. In this case, what may be happening is a challenge to the idea that it is socially unacceptable to say racist things.\(^{29}\) We can see this type of social taboo-breaking in an even more extreme form by looking at (T3) again.

\[(T3)\]

I and the rest of the world are pleased your piece of shit family … are dead squashed filthy shit rotting in the ground. Especially those two filthy babies that were squashed REST IN PISS YOU FUCKING RODENT PIECES OF SHIT.

There is clearly a social norm by which it is rarely acceptable to *speak ill of the dead*. There also seems to be something even more socially distasteful about insulting the mourning families of the deceased, and yet another more shocking layer is added when making light of the death of babies. We do not know much about the disposition of the author of (T3) or their beliefs regarding the comment, so it could plausibly be that (T3) is a sincere utterance and the author could simply be intending to do nothing more than cause distress, for example. However, the fact that this is not an isolated example, and indeed as RIP Trolling is a well-documented phenomenon, we have good cause to think that the author is not merely abusing the families of the victims, though they are indeed doing this. If it is an act of trolling, then they are engaged in abuse and also using this abusive speech as a tool in an act of trolling. Phillips (2015, Ch. 7) argues that some of the rationale behind RIP trolling is a reaction to a perceived mawkishness in internet responses to tragedies, and utterances such as (T3) are used to respond to this wider social trend. If this is the case, then utterances such as (T3) can be viewed from certain angles as a form of satire. Which again moves us back toward the dual-intentional nature of an act of trolling.

Suppose, then, that the author of (T3) wishes to satirize responses to tragedy found on the internet by making an utterance of polar-extremity to those customarily left on memorial pages. If this is the case, then in this sense (T3) is to be viewed by onlookers as an unserious statement—it’s satire.\(^{30}\) However, clearly, the posting of such a comment in a social space wherein most other comments are sympathetic in nature will have a particular effect, and addressed as it is specifically at the families of the deceased, it seems that part of the “joke” is that it is such an offensive thing to say that a target will likely have a strong emotional reaction to it. Unless they accept it as satire, then they might seem to have little choice but to respond seriously to the utterance even if they suspect it is an act of trolling. The problem for the target, then, is that how to respond is not clear. This, I suggest, is the trolling dilemma.

5  \textbf{THE TROLLING DILEMMA}

Central to the spread and success of the phenomenon of trolling and its complex socio-linguistic role is something we might call the *trolling dilemma*.\(^{31}\) The dilemma faced is in how to respond to an act of trolling. As such, it does not really arise when an act of trolling is successful in the sense that the target is unaware that they are being trolled and responds seriously to a trolling utterance. We do see
the dilemma, however, in almost all other cases wherein a potential target is suspicious they might have encountered a troll. Consider again Trump and the RIP troll. Many readers encountering (T2) or (T3) might be suspicious that these are trolling utterances. In being suspicious in such a way, the reader might also suspect that the intentions of Trump and the RIP troll are for the target to take these utterances seriously, in some sense, and respond accordingly. Responding to the trolling utterance in such a way, though, allows the troll to successfully troll—the respondent thus becomes the target and does precisely what the troll wants. However, to not respond as if trolling utterances are serious statements, and simply dismiss them as acts of trolling, can therefore allow a president to use racist tropes, or an RIP troll to direct abuse at mourners without challenge. In effect what this means is that a president using racist tropes can be dismissed as engaging in mere trolling.

If the division I suggest that exists between target and onlooker reflects the structure of an act of trolling, there might be another unpalatable conclusion to be drawn too. In not responding we might, by default, be an onlooker too. If we assume Trump is just joking or the RIP troll is merely satirising in order to troll, then we take their utterances to be unserious. We need not endorse (T2) or (T3), and we might find them repugnant, but we take them to be unserious in the sense that we do not take them to be utterances worthy of serious response. This could even be the case in the example (T1), though with much less worrying moral implications.

(T1):

Ben Franklin said politicians are like pampers, they both stink and they act like babies.

It's plausible that a reader of (T1) might assume it to be a joke; it's verifiably wrong and makes little sense. (T1) did receive what appears to be serious responses though; recall it was pointed out to KenM that pampers did not exist during Franklin's time. Again we have no knowledge of what the respondent thought of (T1), but even if they were suspicious that (T1) was not intended to be a serious remark, not knowing for certain might still be reason enough to respond in a way amenable to the troll's troll-target intentions. Suppose the respondent was someone with an acute concern about the well-documented spread of false information on the internet. This being the case, then, even an example as seemingly innocuous as (T1) might push someone toward the dilemma. From that particular point of view, the dilemma is: should one respond seriously to a suspected act of trolling, and so become a target in a successful act of trolling, or should one let the demonstrably false claim go unchallenged?

That such a dilemma arises out of an act of trolling, I suggest, is part of the reason why trolling has been embraced by political actors in recent years, particularly those wishing to propagate divisive violent ideologies. The ambiguous dual-identity of an act of trolling—as speech that can be responded to either seriously or unseriously—allows those with racist views, for example, to make public their racist ideas, but with enough ambiguity that an effective response is difficult to formulate. That in such a case the trolling dilemma can give at least the appearance that there is no good response—the troll either succeeds in trolling when taken seriously, or racist views remain unchallenged or dismissed as mere trolling when taken unseriously—provides the racist statement with a status not befitting its actual content. That is, it can appear like a view or argument that has no adequate response. However, in the case of trolling, it is not the argument itself that carries this dilemma, but rather it is the ambiguity built-in to the rhetorical device (an act of trolling) used to deliver it. There is another closely related side to the appeal of trolling to these groups too, and this is the topic I discuss next; the cloak of humor.
So far the characterization of the speech act of trolling has been based around the notion of seriousness. This notion, I claim, is closely related to joking. As such it is worth considering the role of humor in trolling in a little more detail. That trolling is often malignant, bullying and abusive, and is increasingly used as a tool of propaganda, might make it seem odd to suggest that a form of humor or unseriousness underpins it. However, when we remind ourselves that humor can be used to dehumanize, poke fun, and bully, and can also be used to mark out “us” and “them,” then it need not be thought to somehow dampen the force of an act of trolling. A racist “joke” is still a racist remark, just as a sexist “joke” is still sexist. That it is shrouded in the cloak of humor makes it no less potent. In fact, there is an argument to be made that it becomes even more forceful in this guise. Indeed consider what Hobbes describes in his famous description of what has come to be known as the superiority view of humor.

I may therefore conclude, that the passion of laughter is nothing else but sudden glory arising from some sudden conception of some eminency in ourselves, by comparison with the infirmity of others … It is no wonder therefore that men take heinously to be laughed at or derided, that is, triumphed over.

Although Hobbes's account does not seem to fit with many forms of humor, it captures something of the nature of trolling. If the target is merely a prop in an act of trolling, then in Hobbes's terms, the troll's aim is to triumph over the target, and they do this for the entertainment of the onlooker. If this is correct, then what seems to be built-in to the act of trolling is a natural division between onlooker and target. The target becomes a mere figure of fun, and as I will discuss in Section 6.2, this has some serious consequences. Before considering this though, I want to discuss how trolling acts can provide a level of deniability to trolling utterances.

6.1 Figleaves

There are good reasons why we might want to avoid dismissing Trump using racist tropes as mere trolling, however, this type of deniability is invoked in Trump's defence (even implicitly by his critics). That this type of defence might seem available gives us reason to consider that in some respects trolling operates in a similar way to a racial figleaf, such as discussed by Jenny Saul.

A successful racial figleaf operates by offering a block on the inference from an utterer making a racist comment to the conclusion that the utterer is racist (“I'm not racist but …,” being a familiar example). If it is accepted that Trump is trolling when uttering racist tropes, then it leaves open the possibility to read his utterance unseriously. So if it is viewed by an onlooker as merely a means to “wind-up” those who do take it seriously, or to force them to confront the trolling dilemma rather than taking it as a serious utterance, then this changes what it becomes acceptable to say. As Saul puts it; “[i]f the audience accepts that the figleaf blocks the concern about racism arising from the utterance of racist sentence R, then R becomes seen as something one can say without being racist.” Similarly, if someone were to accept that by uttering (T2) Trump is merely trolling, this too might provide a block to the concern of racism that Trump's utterance of the racist (T2) might ordinarily entail. A conclusion from that being that if (T2) is uttered in an act of trolling, then (T2) is not necessarily racist. So whereas with the type of figleaf characterized by Saul a speaker will offer some qualifying remark, in an act of trolling the act itself operates in this way. Again, this should give us cause for concern.
and points again to the problem posed by the trolling dilemma. For if we dismiss \( \text{T2} \) as an act of mere trolling, we give credence to the notion that \( \text{T2} \) can be uttered (without caveat) in a non-racist way (or perhaps less strongly, in a “less-racist” way), we are in danger of accepting it as an effective figleaf. The dilemma is such though that if we do respond to Trump’s racism, we then become props in Trump’s act of trolling.39

6.2 | Humor and propaganda

The use of humor in propaganda is not new. Humorous media such as cartoons and caricatures have been used effectively in recent history as a means of disseminating political messages. What seems to set trolling apart though is that not only can it be used to disseminate ideas, but the people those ideas reach are then equipped to use trolling as a means to disseminate them further—even if they are “just joking.” So even if a troll is initially drawn solely to the seeming irreverent and taboo-breaking “humor” of using, for example, racial slurs in an act of trolling, and even if they do not take themselves to sincerely hold racist beliefs, when they use them in an act of trolling they are still spreading racist ideas. In doing this the troll is still contributing to a divisive climate both in the content of what they say and also with the type of speech act they are using. For it seems that making unpalatable ideas appear in some way less serious can detract from their insidiousness, yet can still be used to broadcast them further. Not only can it be used to broadcast them, but it can also make them more appealing; humorousness and unseriousness are attractive qualities after all.

Following on from this it should also be noted that trolling has become an effective method of spreading authoritarian and far-right ideas in recent times, and based on the previous discussion, we might have some good reasons to expect trolling to be such a natural vehicle for this type of idea. One of Jason Stanley’s overall points in his analysis of the tactics used by fascists to gain power is that they seek to separate populations.40 Trolling as an act does this regardless of the content of the utterance. It is built-in to an act of trolling that there be a target and an onlooker, and that the target is objectified—the target is a prop in an act of trolling and their worth is as a mere source of amusement. This in itself need not be socially divisive, such as in the case of \( \text{T1} \), which seems so absurd as to not be genuinely divisive. However, even innocuous-seeming \( \text{T1} \) includes targets to be laughed at, and even \( \text{T1} \) can make a reader confront the trolling dilemma. So when already divisive political ideas are loaded into the act of trolling, the divisiveness of the act itself, in combination with the divisiveness of the content, makes for a powerfully disruptive act. If I’m right about the humor component of trolling too, there is further cause for division. To the onlooker, a target is lacking humor, they do not get the “joke.” So there is a division between those who “get the joke” and those who do not. The power of that as an enforcer of disharmony should not be underestimated. Humor is not to be taken lightly, and if trolling is a form of humor, then this makes it no less of a serious matter.

6.3 | Responding to an act of trolling

One important question this all leads to is on the matter of what the best response to an act of trolling might be.41 One frequent exhortation we see in online communicative environments is “do not feed the trolls.” Such an approach of ignoring those suspected of trolling behavior can of course be useful in some cases in which the troll is not using slurs or hate speech as a means of trolling, though when applied to cases such as the Trump tweet \( \text{T2} \), for example, such an approach is not a realistic response. And it might seem contrary to the moral commitments many of us might have to challenging
hate speech when we encounter it (and when we are in a position to challenge it) to simply ignore such utterances. This approach would seem to allow the unchecked propagation of racist ideas such as those found in (T2). Though of course if my analysis of the dynamics of trolling (in which an act of trolling forces upon the target the trolling dilemma) is correct, then part of the problem with a serious response is that it allows the troll a victory of sorts in having a claim to have trolled a target.

Further to this problem the structures of social media and the way information spreads on the internet entails that any well-meaning response to an act such as (T2) also necessarily draws attention to (T2) itself. Part of the power of the trolling dilemma is in the potential for paralysis of response it creates. I think the best solution to this problem will vary depending on the nature of a specific act and will require cross-disciplinary work. For example, although structurally similar to the other examples we might think how we respond to cases such as (T1) is not as important as how we respond to cases such as (T2) or (T3) as the stakes are much lower. Whether one is trolled by KenM and how one responds is of little overall societal importance. Yet how we respond to racist, misogynistic, or xenophobic trolling utterances is of great importance. And so this is the type of trolling that we should be most interested in developing responses to. What exactly the most effective response is, however, is not something I can develop alone. What is needed for this would be a collaboration between psychologists, linguists, sociologists, and philosophers. For although there are clear moral issues relating to acts of trolling that philosophers are well placed to consider, there is a need too for us to collaborate to understand some of the ramifications. So my inclination on it is that the most morally acceptable response is going to coincide with the most practically effective response. And for this, we would benefit from the insights and data of linguists, psychologists, sociolinguists, and sociologists.

The author is very grateful to Jenny Saul, Komarine Romdenh-Romluc, and Will Hornett for early discussions on this topic. And to Jenny, Komarine, Neri Marsili, Eliot Michaelson, and two anonymous reviewers for helpful feedback on various drafts.

ORCID

P. J. Connolly © https://orcid.org/0000-0003-1076-6205

ENDNOTES

1 Such use of “trolling” dates back to at least Nobbes (1682)—which gives me (part of) my title.
2 For example, Gallucci (2019).
3 This stretches from the use of large-scale so-called “troll factories” by regimes (see, for example, Gallagher 2019). And Giles (2019), to the use of a form of trolling known as “shitposting” by political parties in election campaigns (for example, see Read 2019). Stokel-Walker (2019), and by far-right groups, such as documented in Evans (2019).
4 See Edwards (2019) and Grossman (2019).
5 Edwards (2019).
6 The original article on AP News has since been removed, see https://horsey surprise.tumblr.com/image/13047 7486506 for image, and https://www.reddit.com/r/KenM/ for more examples.
7 One aspect of internet discourse that we might think important when considering an act of trolling is the role anonymity plays in creating an environment in which such behaviors flourish. For example, Goldberg (2015) provides an outline and explanation of the link between anonymity and the epistemic quality of assertions. In this explanation Goldberg does a great deal to show how actors we might describe as “trolls” work to exploit the diminishment of the warranted expectations that hearers and audiences have toward one another. The Trump case is interesting because clearly he is not making anonymous utterances. I do not think this poses a problem for Goldberg’s account, however. In a fuller account of a speech act of trolling one promising direction we might go in would be to draw parallels between the types of lowered expectations we find in cases of anonymous assertion and how Trump’s very public carefree relationship to the truth creates space for a similar diminishment of expectations.
8 Pengelly (2019).
9 LA Time Editorial (2019).
10 See Phillips (2011) and Phillips (2015).
11 Cited in McCosker (2014).
12 See McCosker (2014) and Phillips (2015, Ch. 7) for more detailed discussions of RIP trolling.
13 Bishop (2014) and March and Marrington (2019), both use such terminology, Phillips (2015) also differentiates types of trolling along similar lines.
14 See Jane (2014); Jane (2015); March and Marrington (2019).
15 Hardaker (2010).
16 It is worth noting Hardaker in later work deems this definition insufficient due in part to its narrowness and also its focus on trolls rather than acts of trolling. She suggests a revised definition: “Trolling is the deliberate use of impoliteness/aggression. Deception and/or manipulation in [computer mediated communication] to create a context conducive to triggering or antagonizing conflict, typically for amusement’s sake.” See Hardaker (2013).
17 A standard way of thinking about the role of sincerity in communicative practice, for example, is along the lines of that found in Searle (1969). Loosely put, Searlean sincerity is determined by the relationship between an utterance $u$ made by $S$ and $S$’s beliefs about $u$. So if $S$ believes $u$ (in the case of assertion), then $S$ is sincere in uttering $u$. Conversely, if $S$ utters $u$ yet does not believe $u$ (assertion), then $S$ utters $u$ insincerely. So for Searle a speech act is sincere when the speaker has the mental state expressed by their utterance, and insincere when they do not. The Searlean view meets with counterexamples Chan and Kahane (2011); Peacocke (2000); Ridge (2006); Stokke (2018), for example, however, the direction of these challenges generally aims to question what it is to have some mental state in relation to some utterance. I do not here want to challenge any view of sincerity, however, my point is that for the present task, sincerity is not quite the notion we should be considering.
18 As an anonymous reviewer points out though, in Hardaker’s definition it could be taken that what is important is not necessarily how the troll actually stands in relation to the content of the utterance, but rather that the troll at least gives the impression of standing in a sincere relation to the content.
19 Frankfurt (2005).
20 It is possibly a nod to a quote misattributed to Mark Twain about politicians and diapers needing to be changed frequently.
21 The relationship between trolling to humor is enduring and inescapable. To the type of subcultural troll discussed in Phillips (2015, Ch 2); a form of laughter, “lulz,” is very often the overall aim of any act of trolling. She says; “A corruption (or as the trolls might argue, perfection) of “Laugh Out Loud,” lulz celebrates the anguish of the laughed-at victim.”
22 Or to amuse them in some way, this is very simplistic and not necessarily the case, of course, it may be to humiliate or belittle.
23 At least in terms of the literal content of the utterance.
24 I also see no reason why they might not have no sincere/insincere relationship to their utterance. For example, a comedian improvising during a set could utter anything that occurs to them with the aim of provoking amusement. There seems no reason to expect that the comedian need have any particular beliefs about the things they say at the point of utterance.
25 Though I stop short of a precise definition of the speech act of trolling, what I have in mind is something along these lines;

Linguistic trolling is a perlocutionary act in which $A$ utters $u$ and:

1. $A$’s beliefs about $u$ are irrelevant to the act of trolling;
2. $A$’s set of target-directed perlocutionary intentions include that $u$ is taken seriously by some target(s) $B$;
3. A’s set of target-directed perlocutionary intentions include that B acts in response to \( u \) based on (2), and:
4. A’s set of onlooker-directed perlocutionary intentions include that \( u \) is taken unseriously by some onlooker(s) C.

26 Austin (1962).
27 Recanati (1986, 216, fn5).
28 As one anonymous reviewer helpfully points out, we might question why this account relies on perlocutionary acts rather than illocutionary acts as its basis. My thoughts on this are that what is most striking about acts of trolling is the intended multi-faceted perlocutionary consequences. It is for this reason that I offer an analysis in terms of perlocutionary acts.
29 That it is the actual president making the utterance makes this an even more profound challenge to the norm.
30 This does not exclude that satire can be serious, however, if there is such a point to be conveyed by an utterance such as (T3) it would take a lot of presupposition for such an interpretation. That is, faced with simply what is said in (T3) alone, a lot of background knowledge is required to arrive at the understanding of it as satire.
31 Thanks to Komarine Romdenh-Romluc for the discussion that led me to think about this.
32 Interestingly enough, the availability of an utterer’s intentions in an act of trolling seems to differ from many other types of manipulative speech, such as misleading, where it is essential for the success of the act that the intention to manipulate remains hidden. For example consider the covert collateral acts such as discussed in Bach and Harnish (1979). These acts “are performed with intentions that are intended not to be recognized.” A similar notion, too, is used in Saul (2018) which suggests that dogwhistles are a strain of covert perlocutionary act which are defined as an act that does not succeed if the perlocutionary intention is recognized. Even recognizing the troll’s trolling intentions, however, does not provide escape from the dilemma.
33 See Bowker (unpublished manuscript) for a similar dilemma.
34 The issues this brings to mind are not modern we can see a correlation in ancient discussions of rhetoric. Think of Plato’s observation in the Protagoras of Protagoras’s use of argument as a spectator sport (335a4-8). Also, Aristotle’s rejection in Rhetoric II.24 of the sophists claims to make the “weaker argument the stronger” (1402a23–5), in which Aristotle explains how the methods of Protagoras used in supposedly doing this were rejected once it was realized they relied on rhetorical maneuvering. On this note, it is worth directing you to the excellent Barney (2016).
35 See Phillips (2019). For more on the role of humor in acts of trolling.
36 Hobbes (1845, Ch. IX.13).
37 Saul (2017).
38 Trump did also offer what Saul terms a diachronic figleaf following the original utterance, tweeting “I don’t have a racist bone in my body.”
39 Indeed, claims of “joking” can often be used to block other types of inference, for example, a threat. Consider the phone call between Rudy Giuliani and a journalist on November 14, 2019 in which Giuliani responded to a question as to whether he was worried if Trump might “throw him under the bus” to avoid impeachment, Giuliani responded “I’m not, but I do have very, very good insurance”, his lawyer, also on the phone call, then interjected to say “He’s joking” (Kirchgaessner 2019).
40 Stanley (2018).
41 My thanks to the two anonymous reviewers for pressing me on how we should respond to acts of trolling. For those interested in some data on how social media users respond to suspected trolling see Hardaker (2015).

REFERENCES
Austin, J. L. 1962. How to Do Things With Words. Oxford: Clarendon Press.
Bach, K., and R. M. Harnish. 1979. Linguistic Communication and Speech Acts. Cambridge, MA: MIT Press.
Barney, R. 2016. “[Aristotle], On Trolling.” Journal of the American Philosophical Association 2: 193–5.
Bishop, J. 2014. “Representations of “Trolls” in Mass Media Communication: A Review of Media-Texts and Moral Panics Relating to “Internet Trolling.” International Journal of Web Based Communities 10 (1): 7. https://doi.org/10.1504/IJWBC.2014.058384.

Bowker, M. Unpublished manuscript. “Edgelording: Provoking Criticism Through Offence.” (manuscript).

Chan, T., and G. Kahane. 2011. “The Trouble With Being Sincere.” Canadian Journal of Philosophy 41 (2): 215–34. https://doi.org/10.1353/cjp.2011.0013.

Edwards, P. 2019. “The World’s Greatest Internet Troll Explains his Craft.” Vox Website. Accessed August 1. https://www.vox.com/2016/5/6/11597960/ken-m-internet-troll.

Evans, R. 2019. “Shitposting, Inspirational Terrorism, and the Christchurch Mosque Massacre.” Bellingcat. Accessed October 9. https://www.bellingcat.com/news/rest-of-world/2019/03/15/shitposting-inspirational-terrorism-and-the-christchurch-mosque-massacre/.

Frankfurt, H. G. 2005. On Bullshit. Princeton, NJ: Princeton University Press.

Gallagher, P.. 2019. “Putin’s Army of Pro-Kremlin Bloggers.” Independent, November 15.

Gallucci, N.. 2019. “13 of the Absolute Best and Worst Brand Trolls in Recent History.” MashableUK. Accessed July 1. https://mashable.com/article/best-brand-trolls-social-media/?eu= true.

Giles, K. 2019. “Putin’s Troll Factories.” The World Today 71: 48.

Goldberg, S. C. 2015. “Anonymous Assertion”, Assertion: On the Philosophical Significance of Assertoric Speech, 222. Oxford: Oxford University Press.

Grossman, S.. 2019. “The 30 Most Influential People on the Internet.” Time. Accessed July 10. https://time.com/4258291/30-most-influential-people-on-the-internet-2016/.

Hardaker, C. 2010. “Trolling in Asynchronous Computer-Mediated Communication: From User Discussions to Academic Definitions.” Journal of Politeness Research 6: 242. https://doi.org/10.1515/jplr.2010.011.

Hardaker, C. 2013. “‘Uh. . . . Not to Be Nitpicky,,,,,But…the Past Tense of Drag Is Dragged, Not Drug.’: An Overview of Trolling Strategies.” Journal of Language Aggression and Conflict 1 (1): 86.

Hardaker, C. 2015. “‘I Refuse to Respond to this Obvious Troll’: An Overview of Responses to (Perceived) Trolling.” Corpora 10 (2): 201–29. https://doi.org/10.3366/cor.2015.0074.

Hobbes, T. 1845. The English Works of Thomas Hobbes of Malmesbury. London: Longman, Brown, Green, and Longman.

Jane, E. A. 2014. “‘Your a Ugly, Whorish, Slut’: Understanding E-bile.” Feminist Media Studies 14 (4): 531–46.

Jane, E. A. 2015. “Flaming? What Flaming? The Pitfalls and Potentials of Researching Online Hostility.” Ethics and Information Technology 17: 65–87. https://doi.org/10.1007/s10676-015-9362-0.

Kirchgassner, S. 2019. “Rudy Giuliani Says Trump Will Stay Loyal to Him But Jokes That He Has “Insurance”.” The Guardian. Accessed November 14. https://www.theguardian.com/us-news/2019/nov/14/rudy-giuliani-donald-trump-insurance-loyal.

LA Time Editorial. 2019. “Trump is Truly America’s Bigot-in-Chief.” Los Angeles times. Accessed July 27. https://www.latimes.com/opinion/editorials/la-ed-trump-aoc-squad-ilhan-bigoted-tweets-20190714-story.html.

March, E., and J. Marrington. 2019. “A Qualitative Analysis of Internet Trolling.” Cyberpsychology, Behavior, and Social Networking 22 (3): 192–7. https://doi.org/10.1089/cyber.2018.0210.

McCosker, A. 2014. “Trolling as Provocation: YouTube's Agonistic Publics.” Convergence 20: 201–17. https://doi.org/10.1177/1354856513501413.

Nobbes, R.. 1682. The Compleat Troluer : or, The Art of Trolling, With a Description of All the Utensils, Instruments, Tackling, and Materials Requisite Thereto : With Rules and Directions How to Use Them: As Also a Brief Account of Most of the Principal Rivers in England. London: T. James for Tho. Helder.

Peacocke, C. 2000. Conscious Attitudes, Attention, and Self- Knowledge. In Knowing Our Own Minds, edited by C. Wright, B. C. Smith, & C. Macdonald, 63–98. Oxford: OUP.

Pengelly, M.. 2019. “‘Go Back Home’: Trump Aims Racist Attack at Ocasio-Cortez and Other Congresswomen.” The Guardian. Accessed July 24. https://www.theguardian.com/us-news/2019/jul/14/trump-squad-tlaib-omar-press-ley-ocasio-cortez.

Phillips, W. 2011. “LOLing At Tragedy: Facebook Trolls, Memorial Pages and Resistance to Grief Online.” First Monday, 16.

Phillips, W. 2015. This is Why We Can't Have Nice Things: Mapping the Relationship Between Online Trolling and Mainstream Culture. Cambridge, MA: MIT Press.
Phillips, W. 2019. “It Wasn’t Just the Trolls: Early Internet Culture, “Fun”, and the Fires of Exclusionary Laughter.” *Social Media Society* 5 (3): 205630511984949. https://doi.org/10.1177/2056305119849493.

Read, J. 2019. “Tories Accused of Trolling Remainers by Using Comic Sans in Pro-Brexit Post.” *The New European*. Accessed October 25. https://www.theneweuropean.co.uk/top-stories/tories-accused-of-trolling-remainers-by-using-comic-sans-in-brexit-post-1-6335865.

Recanati, F. 1986. “On Defining Communicative Intentions.” *Mind & Language* 1: 213–41. https://doi.org/10.1111/j.1468-0017.1986.tb00102.x.

Ridge, M. 2006. “Sincerity and Expressivism.” *Philosophical Studies* 131 (2): 487–510. https://doi.org/10.1007/s11098-005-2218-4.

Saul, J. M. 2017. “Racial Figleaves, the Shifting Boundaries of the Permissible, and the Rise of Donald Trump.” *Philosophical Topics* 45: 97–116. https://doi.org/10.5840/philtopics201745215.

Saul, J. M. 2018. “Dogwhistles, Political Manipulation, and Philosophy of Language.” In *New Work on Speech Acts*, edited by D. Fogal, D. W. Harris, and M. Moss, 383. Oxford: Oxford University Press.

Searle, J. R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Stanley, J. 2018. *How Fascism Works: The Politics of Us and Them*. New York: Random House.

Stokel-Walker, C. 2019. “‘They're Doing This Badly on Purpose’: Why the Tories' Latest Online Ads Look So Ugly.” *The New Statesman*. Accessed October 26. https://www.newstatesman.com/politics/media/2019/10/they-re-doing-badly-purpose-why-tories-latest-online-ads-look-so-ugly.

Stokke, A. 2018. *Lying and Insincerity*. Oxford: Oxford University Press.

**AUTHOR BIOGRAPHY**

**P. J. Connolly** is a research affiliate at the University of Sheffield. Although interested in language more generally, his current research is particularly focused on the application of theories from the philosophy of language to internet speech. As the topic of internet speech is so broad Patrick also draws upon work from social epistemology, political philosophy, philosophy of mind, linguistics, sociology and psycholinguistics. He is also interested in deceptive linguistic behavior and matters relating to the semantic-pragmatic distinction.

**How to cite this article:** Connolly, P. J. (2021). Trolling as speech act (or, the art of trolling, with a description of all the utensils, instruments, tackling, and materials requisite thereto: With rules and directions how to use them). *Journal of Social Philosophy*, 00, 1–17. https://doi.org/10.1111/josp.12427