

Copyright page

This article is the uncorrected proof. To view the final, published version please follow the doi links. To cite this paper, please use the following:

Coles, B. A., & West, M., (2016). Trolling the trolls: Online Forum Users Constructions of the Nature and Properties of Trolling, *Computers in Human Behavior*, 60, 233-244. doi: <http://dx.doi.org/10.1016/j.chb.2016.02.070>

©2016 This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>



Uncorrected Proof

Trolling the trolls: Online Forum Users Constructions of the Nature and Properties of Trolling

Dr. Bryn Alexander Coles*, Melanie West*

Department of Psychology and Counselling, Newman University, Bartley Green, Birmingham, B32 3NT, United Kingdom

Uncorrected Proof

*Corresponding author

Email address: a.coles@newman.ac.uk (Dr. Bryn Alexander Coles)

Trolling the trolls: Online Forum Users Constructions of the Nature and Properties of Trolling

Abstract

‘Trolling’ refers to a specific type of malicious online behaviour, intended to disrupt interactions, aggravate interactional partners and lure them into fruitless argumentation. However, as with other categories, both ‘troll’ and ‘trolling’ may have multiple, inconsistent and incompatible meanings, depending upon the context in which the term is used and the aims of the person using the term.

Drawing data from 15 online fora and newspaper comment threads, this paper explores how online users mobilise and make use of the term ‘troll’. Data was analysed from a discursive psychological perspective.

Four repertoires describing trolls were identified in posters online messages: 1) that trolls are easily identifiable, 2) nostalgia, 3) vigilantism and 4) that trolls are nasty. A final theme follows these repertoires – that of identifying trolls. Analysis also revealed that despite repertoire 01, identifying trolls is not a simple and straight-forward task.

Similarly to any other rhetorical category, there are tensions inherent in posters accounts of nature and acceptability of trolling. Neither the category ‘troll’ nor the action of ‘trolling’ has a single, fixed meaning. Either action may be presented as desirable or undesirable, depending upon the aims of the poster at the time of posting.

Keywords: Trolling, online behaviour, internet forum, discursive psychology, social construction

1. Introduction & background

‘Trolling’ refers to a specific type on malicious online behaviour, intended to aggravate, annoy or otherwise disrupt online interactions and communication (Binns, 2012; Bishop, 2012a). Trolling is also a topical, important issue attracting an increasing

5 amount of attention in the public eye (Bishop, 2014) Cases of individuals identified as trolls, targeting other high-profile individuals are not only being widely reported in the press, but are also forming the basis of court actions.

Trolling appears to be pervasive throughout online media, having been observed in such diverse locations as online magazines (Binns, 2012), social networking sites
10 (Cole, 2015; Bishop, 2012a), online computer games (Thacker & Griffiths, 2012), online encyclopedia (Shachaf & Hara, 2010), online newspapers (Ruiz et al., 2011) and even on Government e-petition pages (Virkar, 2014).

Trolling can have serious consequences for both the perpetrators and the victims of such behaviours, not only in their online spaces, but also in their daily life (Binns,
15 2012). The presence of trolls in online spaces may serve to create a hostile online space, unwelcoming to new posters, inhibiting the development of online communities. Alternatively, the consequence of uncivil online behaviour may include the polarisation of opinions and beliefs within that online group, as people move to reject the subject of uncivil discourse, or more damagingly, to accept it as normative. This
20 may result in harsher opinions or judgements being formed, acting as a mechanism for the maintenance of prejudicial attitudes (Anderson et al., 2014). Amongst the more serious consequence of trolling is an increase in the risk of suicidal ideation and self harm amongst the victims of such behaviours (Hinduja & Patchin, 2010; Bauman et al., 2013). As such, trolling may be understood not only as an unpleasant behaviour, but
25 as an unethical one which holds the potential to do great harm.

1.0.1. Characterising the troll

A number of academics have offered definitions of 'trolling' alongside attempts to account for trolls behaviours (Shachaf & Hara, 2010; Hardaker, 2010). Trolling may be understood as the posting of subtly or unsubtly offensive messages in order to
30 create offence, start an argument or lure the unwary into pointless debate (Binns, 2012). Trolling may also be understood as unconstructive messages designed to provoke a reaction, to draw targets (and others) into fruitless argument, and to disrupt the avowed purpose of the group gathering (Bishop, 2012a). Trolling may also be understood as repetitive, harmful actions which violate a websites terms of use (Shachaf & Hara,

35 2010). Trolls following this latter definition should be demonstrably active throughout all sections of the website community which they are attacking, rather than just limiting their activity to a limited subsection of the online space.

As with the off-line world, one of the facilitating features of malicious actions online is anonymity (Shin, 2008; Suler, 2004). Anonymity is thought to provide a facilitating condition for disinhibition, leading in turn to greater self-disclosure, deindividuation and the emergence of counter-normative behaviours (Bishop, 2013b). Similarly with off-line behaviours, this may not necessarily be connected with notions of deindividuation, but rather may be associated with a shift in ones sense of identity and self-salience, and a loss of self monitoring (Suler, 2004).

45 Suler (2004) argues that factors which may encourage counter-normative online behaviour include dissociative anonymity; invisibility; asynchronicity; solipsistic introjection; dissociative imagination and the minimisation of authority. When such conditions are met, people do and say online what they would not do or say in an offline environment. Taken to an extreme, this 'online disinhibition effect' may be termed
50 'toxic disinhibition'.

Hardaker (2010) explored users definitions of the phenomenon of trolling, drawing upon an extensive archive of data collected over a nine year period from a single forum. Hardaker (2010) collected instances of posters using the word 'troll' or mobilising various euphemisms, such as making reference to other members of the online group 'living under a bridge'. This investigation revealed that posters definitions of a
55 troll typically contain four characteristics. Those of: deception (hiding ones motivations); aggression (attempting to rile other posters); disruption (disturbing the flow of interaction); success (if the trolls failed to provoke anyone, they were not considered successful).

60 However, not all trolling may necessarily be considered malicious. Interestingly, some authors have also noted that more positive definitions of trolling may exist. Bishop (2012a) describes 'kudos trolling', where users may post irrelevant information seemingly in good faith. The disruption of online interactions here may be considered an unfortunate and unintended consequence.

65 Herring et al. (2002) explore in depth the requirements of successful trolling, de-

scribing the stages which a troll in a specific case study moves through. These stages include: outward manifestations of sincerity, laying the ‘flame bait’, and attempting to provoke others into engaging in futile arguments.

Personality variables have also been noted to play a role in trolling, with some authors noting that trolls may exhibit a ‘dark tetrad’ of sadism, psychopathy, machiavelianism and narcissism or may otherwise display symptoms of personality disorders (Suler, 2004; Buckels et al., 2014; Bishop, 2013b).

Aside from personality characteristics or situational factors contributing to disinhibition, trolls have also been shown to be motivated by circumstantial factors such as boredom, attention seeking, revenge and the perception of their targets or online spaces as sources of entertainment (Shachaf & Hara, 2010).

1.0.2. Troll management

Strategies suggested for the management of trolls vary, according to the aims and sophistication of the online space in which the trolls are operating. At its simplest, users of online spaces are simply admonished ‘do not feed the trolls’ (Binns, 2012; Shachaf & Hara, 2010). Deprived of oxygen, these flammers are expected to quickly die down.

An alternative, more involved approach to troll management has been termed ‘gamification’. This refers to the use of video game elements in non-gaming contexts (Detterding et al., 2011). Specifically, online spaces may have elements designed into them which are intended to discourage trolling behaviours. This may include a requirement to ‘sign in’ to the online space in order to be able to interact – an action which should also reduce the anonymity and deindividuation effects of the individual, as it would render them traceable. Alternatively, ‘tokens’ may be awarded for good online behaviour. These tokens may in turn allow members of that online space to engage in additional activities, privileges, or to simply ‘collect a high score’ (Binns, 2012; Bishop, 2012a).

However, requiring individuals to link all their online activities may have an unanticipated downside in inhibiting peoples freedom of expression, where there is a fear that what is said could be taken out of context, be misunderstood or otherwise misrepresented.

resented.

Yet another solution to the scourge of trolling may be to depend upon the use of moderators, to control what may be posted online. Moderators may clearly delineate the boundaries of civil behaviour (Binns, 2012; Bishop, 2012a). This may be achieved
100 through the use of a network of moderators, who may work within the online community in order to establish norms of civil discourse which all members of the online space can conform to (Lampe et al., 2014).

This however may rapidly become a labour intensive activity, if a large number of individuals begin posting and to the online space, and each comment requires consider-
105 ation before it can be published. Potential resolutions to this issue include distributing moderation duties amongst a broad selection of the online forum (Lampe et al., 2014), or by automating the detection of malicious posts. Software algorithms may be deployed which are able to monitor information posted online and automatically filter out unwanted online activity (Galán-García et al., 2014).

Another attempt to resolve the problem of trolling is legislative in nature. Trolling
110 then becomes a criminal offence, and the rule of law may be exercised in order to prevent it (Butler et al., 2009; The Select Committee on Communications, 2014; Bishop, 2013a).

1.1. A discursive psychological approach to trolling

115 The potential solutions to the problem of trolling cited above all rely upon the notion that trolls are easily identifiable, and universally undesirable. However, such an assumption may be problematised. A number of authors have suggested that both trolls and the act of trolling may be divided into a number of sub-categories, such as ‘flame trolls’ Bishop (2013a), ‘hater’ and ‘snert’ Bishop (2012a). Bishop (2012a) Further
120 identifies a category of troll whose actions do not represent an attempt to disrupt the flow of interaction online. Rather, contributions from ‘kudos trolls’ seem to represent a genuine attempt to contribute to the interaction. That disruption occurs is simply an unfortunate side-effect.

The case of the kudos troll may be taken as an example of the tensions inherent in
125 the nature and understanding of ‘trolling’, as well as the action-orientation of language

(Potter & Edwards, 1999; Potter, 2003). The discourse produced by the kudos troll does not produce the intended result (dialogue and information exchange). Rather, the action achieved is to be understood as a disruptive influence in the online space. Information provided in good faith results in the appellation of 'troll'.

130 It is not simply the case that what is said (or perhaps, typed) may be misunderstood or misconstrued between speakers. Discursive psychologists have demonstrated that contradiction and inconsistency form an inherent aspect of communication and interaction. Contradiction and inconsistency are present both between speakers and within a single speakers utterances (Potter & Wetherell, 1987; Elliott, 1996). Such contradic-
135 tions are rarely orientated to by speakers and listeners. Research suggests that written contradictions can remain similarly undetected by readers (Otero & Kintsch, 1992).

In case of trolling, the action orientation of talk is to achieve disruption and annoyance within the online forum. As discussed above, this definition has been drawn from examples of trolling identified in online fora either by analysing archived data from a
140 given source, or through interactions with members of an online space. However, the approach taken has generally asked posters to specifically identify trolls (Shachaf & Hara, 2010), or to draw statements concerning the definition of trolling directly from a corpus of online interactions where trolling is explicitly defined (Hardaker, 2010), in order to develop a definition of trolling as used by those who are affected by the phe-
145 nomenon. Furthermore, such studies have tended to focus upon a single online space as their source of data, rather than drawing upon a number of online sources to compile meanings.

Discursive psychologists have shown that the meaning of terms used can be fluid both within and across situations, rather than fixed. Meanings of categories and la-
150 bels are typically negotiated during interactions (Condor, 2000; Condor et al., 2006; Condor, 2006). Posters to online fora therefore may be assumed to show a similar inconsistency in their use of the term 'troll'. Posters to online fora may adopt a fluid understanding of the term.

1.1.1. Research Question

155 The aim of the current paper is to explore the differing uses and definitions of the phenomenon of trolling online, for members of online fora. Drawing upon a number of different fora, the research question will explore how online users mobilise and make use of the category of 'troll'. Analysis of online forum members responses to the publication of Buckels et al. (2014) will explore the differences and similarities in
160 meaning between speakers and across fora. It will be will shown that, similarly to any other rhetorical category, there are tensions inherent in posters accounts of the nature and acceptability of trolling (Edwards, 1998).

2. Method

2.1. Data Selection

165 Data for this project was drawn opportunistically from publicly available online forum threads hosted by generalist websites. No data was be elicited for the specific purpose of this study. Rather, data comprises that which is already freely available on the internet. Such data can be said to be 'naturally occurring' – not elicited for the purpose of research but rather represents a 'naturalistic record', spontaneously generated
170 by users of target websites (Griffin, 2007; Potter & Hepburn, 2007).

Data was drawn from online sources which had responded to the publication of a journal article on the topic of 'trolling', authored by Buckels et al. (2014). These sources comprised of comments threads for online newspapers and discussion forums in which a members of these online communities had either posted their own summary
175 and initial commentary concerning the paper by Buckels et al. (2014), or they had provided a hyperlink to another summary of the paper elsewhere online.

Comment threads were found via an internet search using a popular and well-known online search engine. The title of the paper by Buckels et al. (2014) was used as the search term entered into the search engine. This search was conducted shortly after
180 the publication of the trigger article, on the 20th of March, 2014. As such, analysis is of any data which had been posted to, and moderate by, the host site by this date.

2.1.1. Inclusion criteria

Comments threads were selected for inclusion in this study if they contained a sufficient number of posts such that members of the forum have the opportunity to interact with each other online. Therefore, any thread with more than five posts on the topic, after the initial posting concerning the paper was deemed to provide enough material to comment on from a research perspective.

Responses were limited to English language articles only, though no geographic restrictions were imposed. Thus, data comprises of articles originating from the USA, Canada and the UK. This lack of region specificity is not considered problematic, as the diaspora of online communities is similarly unrestricted, and individuals posting to online forums may be located anywhere in the world.

This resulted in 15 comment threads being included in the analysis. These comment threads were drawn from either online newspaper comment sections (from a mixture of tabloid and broadsheet newspapers) or from online community forums (either educational or technology forums). The fewest number of posts-per-forum was 8, while the highest number of posts-per-forum was 128. The fewest number of individuals posting on each forum was one, although this was because each individual posted anonymously, either without being required to sign in or to select a unique nickname. As such, each poster thus was assigned the same username by the website. The true number of posters to this website is therefore unknown. The highest number of individuals posting on each forum was 59. Table 01, below, summarises some further characteristics of the dataset.

Activity	Count
<i>n</i> Threads	014
<i>n</i> Posters	307
<i>M</i> Posters-per-thread	022
<i>M</i> Posts-per-thread	045
<i>n</i> Posts	625
<i>M</i> Posts-per-user	002

Table 1: Average characteristics of the data

Comments in threads online may be voted ‘up’ or ‘down’ – thus affecting the order
205 of presentation to the audience. For this study, comments were not filtered sequentially
by date submitted to the host website, but rather were filtered ‘naturally’. That is,
comments were taken in the order they were published and presented and organised on
the websites according to whether they had been voted ‘up’ or ‘down’ in the rankings
and at the time of data collection.

210 Within psychological research it is generally accepted that behaviours which are enacted
in a public space, and which are freely observable, do not require consent when
those behaviours have not been specifically elicited by the researchers (Ethics Committee
of the British Psychological Society, 2009). As data for this study is drawn from
online forums, it is considered to meet these criteria and so consent from participants
215 was not sought.

As data was drawn from publicly available online fora, it can be assumed that
posters do not have a reasonable expectation of privacy. Furthermore, use of the data
for not-for-profit research purposes constitutes fair use.

The decision was taken not to further anonymise the data by changing posters
220 names or otherwise obfuscating their identities, as presenting extracts from online fora
and comment threads in support of the analysis also renders the data discoverable on-
line though a simple search.

2.1.2. *Data analysis*

Analysis was approached from a social constructionist perspective – that meaning
225 is produced through talk and interaction (Andrews, 2012; Burr, 2003). Data was analysed
using a discursive psychological approach (Lamerichs & Te Molder, 2003; Potter,
2012) to explore how meanings and understandings are negotiated by contributors in
an asynchronous online forum. Specifically, data was analysed in order to explore how
users of online forums negotiate the meanings of terms such as ‘troll’ and ‘trolling’
230 during interaction.

As the gender of posters is not generally identifiable from their usernames (except
in occasions where posters use an evidently gendered username, or provide a profile
picture), all posters will be referred to as ‘he’. This is not to imply that all posters are

male. Rather it is a convention adopted for the convenience of this papers authors.

235 3. Analysis

Posters make a wide variety of comments concerning the nature of ‘trolls’ and of ‘trolling’. The nature of these posts reveals posters mobilise a variety of repertoires in constructing ‘trolling’ as a separate activity from ‘being a troll’. Further, these repertoires reveal a tension between trolling as a sometimes acceptable behaviour, and trolls as sometimes undesirable elements.

This analysis is organised around the four repertoires which were identified in posters online messages. These repertoires are: 1) that trolls are easily identifiable, 2) nostalgia, 3) vigilantism and 4) that trolls are nasty. A final theme follows these repertoires – that of identifying trolls. This final section shows that despite protestations to the contrary, identifying trolls is not a simple and straight-forward task.

3.1. *Repertoire 01: Trolls are easily identifiable, trolling is acceptable*

A common repertoire amongst all sources examined is that ‘trolls’ are easily identifiable to the regular posters of the online space. This identifiability is constructed in a number of ways. This repertoire is used by posters to protect and reclaim the online space – to position trolls as unsuccessful in their attempts to disrupt online communications. As posters can readily and easily identify whom the trolls are, trolls can be easily and appropriately responded to. This will be examined in the following extracts.

It is important to note that when posters talk explicitly of trolls, they acknowledge the difficulty in identifying online forum members who would fit this designation. However, when online users are talking about trolls in more general terms, it becomes implicitly asserted that trolls are easily identifiable.

Extract 01, below, is drawn from the comments section of a British broadsheet newspaper’s online version. The post appears roughly halfway through the range of comments made by other readers of the newspaper. The poster, ‘Arfur Sixpence’ makes this remark without particular reference to any other individual involved in the comment thread.

3.1.1. Extract 01 Arfur Sixpence, *The Independent*

1 In the good ol' days when the comments on this site were entered via
2 the Disqus system, there was a wonderfully embittered and sullen little
3 troll calling itself 'olympic'. No capital 'o' such was the troll's humility.
4 Myself and many others took great pleasure in baiting the troll to see if it
5 would dance and, oh, how it danced!

Extract 01 demonstrates both the posters confidence in being able to correctly identify trolls, as well as the negative evaluations which posters attach to the nature of trolls. Extract 01 also demonstrates a common theme in posters avowed responses to trolls – that of 'trolling the troll'.

10 At lines 02–03, Arfur flags the existence of a troll known to the readers of the Newspaper's online forum. This troll is described as being 'wonderfully embittered and sullen'. These personality traits stand in contrast with the 'dark tetrad' identified by researchers (Buckels et al., 2014). The troll's third characteristic is given, with irony, as being 'humility'. The trolls characteristics and intentions then are not presented here
15 as being Machiavellian or malicious, but rather as being driven by characteristics which might reasonably be taken as character deficits within the individual (Bishop, 2013b).

Arfur makes an appeal to consensus, presenting this troll as being readily identifiable to forum users in general – rather than positioning his ability to spot trolls as being something which requires skill or specialist knowledge. At line 04, Arfur states that
20 having successfully identified the troll, both himself "and many others" responded to the troll in order to 'make it dance'.

This reference to 'making the troll dance' is itself worthy of note, as it serves to further disempower the troll. The troll is positioned not as an agitator, not as someone who controls the online conversation in order to irritate, annoy or otherwise disrupt on-
25 line communications. Rather, the troll is someone who is manipulated and controlled collectively by the forum users. In this way, the troll is presented as unsuccessful, a figure identified, controlled and exploited by other forum users (which may be compared with the criteria for defining a troll, outlined by Hardaker, 2010).

Extract 02, below, demonstrates once more posters assertions that they are able to

30 identify whom the trolls are in online communities. The extract is taken from an online British tabloid newspaper, which specialises in technology news. Once again, this post is taken from roughly halfway through the responses to the article.

3.1.2. *Extract 02: Anonymous Coward, The Register*

- 1 Funny...
- 2 Given the dismissive drubbing the swivel-eyed trolls are usually dished
- 3 out round here, I'd always assumed their apparently insatiable appetite for
- 4 more made them masochists.

5 Extract 02, above, implicitly reiterates the notion that trolls are readily identifiable in online communities. The poster, Anonymous Coward, asserts that trolls who attempt to post to these comment threads for this newspaper are themselves met with antagonistic comments (line 02). In order to meet trolls with such comments, posters would need to know whom the trolls are. As such the implication is that posters to these
10 threads are able to identify whom the trolls are.

Anonymous Coward also presents trolls as having negative personality characteristics. Once again, these characteristics are different from those identified by researchers. Trolls are described as being 'swivel-eyed' (line 02). Furthermore, in maintaining their presence online in spite of being ill-treated by regular forum users, trolls are further
15 asserted to be masochistic (line 04). This latter comment is less likely to be a genuine assertion concerning the nature of trolls, but rather a comment upon the ability of the members online space to self-police. It is not posters to this newspapers comment threads who are the victims of trolling, but the trolls who are victims of these posters!

The final extract in this section comes from the comments section of an online blog.
20 Once again, this extract comes from a point around halfway through the comments section. What is presented here is not the complete post by BeccaM, rather only the final paragraph of their post.

3.1.3. Extract 03: BeccaM, AmericaBlog

1 Anyway, you're absolutely right: The top solution is not to feed the trolls.
2 But also a given forum, such as this one, really does need its team of mod-
3 erators who *actively and vigorously* go after the trolls who step across the
4 line. What AmericaBlog gets right, other sites such as YouTube, Face-
5 Book and HuffPo don't.

This extract endorsed two of the remedies for trolling posited in the literature – do not feed the trolls (line 01) (Binns, 2012) and the role of moderators in policing the forum (lines 02–03) (Lampe et al., 2014). It is important to note however that earlier in the thread on AmericaBlog, posters have discussed how hard it can be to identify
10 trolls, and to differentiate them from genuine posters. This rhetoric rapidly falls by the wayside in favour of the simple solution advocated here – don't feed the trolls (line 01).

This admonition is surprisingly rare across all sources examined. This may be because the injunction 'don't feed the trolls' requires an explicit orientation to whom the
15 trolls are. When this solution is posed by other online posters, it is often problematised by other posters challenging the conditions under which one should and should not feed the trolls. Nevertheless, the *précis* which precede the comment sections sometimes include this instruction as a closing remark, and at least one thread included 'don't feed the trolls' in its title.

20 However, this solution is predicated upon knowing whom the trolls are. While identifying trolls is not, in this extract, presented as problematic, it is a skill attributed in particular to a specialist subset of forum members – the moderators (lines 02–03). Thus, action regarding trolls is divided into two types. General forum members should simply 'not feed the trolls' – i.e. they should avoid responding to trollish posts. Forum
25 moderators on the other hand should actively pursue trolls (line 03), to police the online community.

Even here however, there is a tension present. Moderators are stated to pursue trolls who step over the line – that is, trolls who go too far (lines 03–04). This implies that there is a tolerable, if not acceptable amount of trolling which may be at the least

30 expected. It is only once trolling exceeds this limit that action need be taken.

While extracts 01–03 commonly imply that trolls can be identified, no posters specify *how* trolls can be identified. Nevertheless, identifying trolls is not presented as a special ability of these posters, but rather as a general ability possessed of any reasonably-minded forum user.

35 Posters do ascribe personality characteristics to trolls though. These characteristics, as we have seen, as variously posited as personality deficits (e.g. bitterness, extract 01; masochism, extract 02). The repertoire of identifiability then may be argued to be built upon the inference of whom the trolls are, based upon their display of negative personality characteristics through the trolls online activities.

40 In being easily identifiable, posters identified as being trolls are also presented as being typically unsuccessful in their attempts to disrupt online communications. Instead, posters identified as being trolls attract derision themselves, and provide amusement – being baited into ‘dancing’ (extract 01) rather than providing irritation to the regular users of the forum.

45 Trolling then is not necessarily a single, negative activity which corrupts the forum. Rather, when trolling is unsuccessful in provoking online forum members, it can act as a form of amusement for them. The next repertoire to be discussed – that of nostalgia – posits trolling not as something negative, but rather as something which provided entertainment for all members of the online community, and which is regarded fondly.

50 3.2. *Repertoire 02: Nostalgia*

The repertoire of nostalgia builds upon that of identifiability to posit trolls as harmless while also differentiating between past and contemporary types of trolling.

55 Across all sources considered, posters have attempted to construct ‘trolling’ as an acceptable practice. As we have seen in extract 01, line 01, one means by which posters attempt to accomplish this is through appeals to nostalgia, and to a ‘better class of trolling’, which was present in the past. This appeal to nostalgia is used to present trolls as largely harmless and ineffective on the forums rather than as an inconvenience or annoyance.

In extract 01, Arfur Sixpence's reminiscences about a troll who used to blight the forums is flagged as being 'the good old days' (line 01). While this designation is most likely attached to the mode of appending comments to newspaper articles (the 'Disqus' system), part of the enjoyment of this era is the presence of the 'wonderfully embittered troll', Olympic (lines 02–03). The enjoyment this troll provided for posters such as Arfur was in baiting and teasing the troll – trolling the troll (line 04). This is presented as being done for the amusement of the posters, to make the troll 'dance' (See also extract 02, lines 02–03).

Another poster to the Register's comment thread further exemplified the repertoire of 'nostalgia'. Extract 04, below is posted in reply to another poster's comments concerning different categorisations (types) of troll.

3.2.1. *Extract 04: Anonymous Coward, The Register*

1 Re: Troll types

2 Trolling the old-school way is not as much fun anymore. It used to be
3 a Devil's Advocate kind of deal. Instead of juicy bites and an apparent
4 debate these days you just get Call Of Duty level illiterate responses.

5 People who are referred to as trolls these days just seem to be those who
6 are being generally offensive and taking cheap shots. No skill at all, and
7 without appreciation of the skill, where's the fun?

Anonymous Coward begins the extract by establishing a temporal division between trolling – the traditional 'old school' approach (line 02) and the modern, 'illiterate' abuse (line 04). This traditional approach to trolling is 'fun' – a notion repeated at the end of the forum post (line 07). The label 'fun' implies that trolling is not necessarily a vindictive, mean-spirited action intended to cause harm. Rather, it is presented as having been a skilled interaction requiring at least the facsimile of debate. This is contrasted both with the current state of the art of trolling ('illiterate', line 04 and 'offensive' 'cheapshots', line 06). This also contrasts with the (tacit) public notion of trolling as a negative practice, exclusively concerned with disrupting online communications.

It is important to note that there is an ambiguity in extract 04. Whom trolling is ‘fun’ for is unclear – it may be fun for those who commit the trolling, as they craft
20 an skillful facsimile of a conversation in order to entrap their conversational partners. Conversely, trolling may be fun for others participating in the thread, who are able to appreciate the ‘skill’ (lines 06–07) of the trolls.

Trolls then, are not only easy to identify, but this very fact disempowers them in their ability to provoke online community members. But members of the online com-
25 munity lament the change from ‘sophisticated’ trolling to ‘simple abuse’.

A further function of baiting the trolls is to expose them for what they are. This is elucidated further in the section on vigilantism, below.

3.3. *Repertoire 03: Vigilantism*

Despite the presence of the admonition ‘don’t feed the trolls’ (extract 03), and the
30 assertion that trolls are easily identifiable (see extracts 01–02). The next three extracts illustrate the importance and function of being able to identify whom the trolls are in online forums – to be able to take more direct action against them, rather than simply not feeding them. It is also important to identify whom the trolls are, so that action is not taken against genuine forum members. This is exemplified through the repertoire
35 of *vigilantism*.

The repertoire of vigilantism serves to construct an appropriate response to trolls, and to provide a means to defend online spaces. It is used to counteract the effects of trolls on the online forum. Vigilantism – trolling the trolls – is constructed as a legitimate action which can be taken by non-trolls. Couching their behaviour in terms
40 of vigilante action allows posters to engage in counter-normative behaviours without necessarily having to adopt (or admit to having) the identity of ‘being a troll’.

This repertoire has been suggested with extract 01, where Arfur Sixpence describes baiting a troll for the amusement of all members of the online community (lines 04–05). The following extracts expand upon this notion, transforming the notion of ‘baiting
45 trolls’ from simply trying to annoy, and into exposing their nature and unpalatable behaviour.

Extract 05 is once again taken from an online British tabloid newspaper, which specialises in technology-reporting. This post comes early in the thread.

3.3.1. *Extract 05: ShellUser, The Register*

1 What does this make me when I start trolling the trolls (sometimes done to
2 expose them or call them out) and when doing so actually take pleasure in
3 the fact that some goofball buries himself completely in his own lies; right
4 up to a point where he's caught in his own web? More sadistic than the
5 sadists or a rightful "hero" who's using proportional "virtual violence"?

In common with the above extracts, ShellUser implies that trolls are known individuals, and that once a troll has been identified, they can be engaged with in their own terms – 'trolling the trolls' (line 01). The poster here treats the identification of trolls as unproblematic – as something which can be taken for granted. As with extracts 01
10 and 02, the poster here claims to spend time aggravating online trolls. Such behaviour is explicitly flagged as 'trolling the trolls' (line 01).

Extract 05 begins with a rhetorical question (line 01), intended to problematise the notions of trolls and trolling. This action is necessary to allow ShellUser to avoid having to adopt the identity of 'being a troll' himself, as will be demonstrated.

15 Having admitted to engaging in trolling, ShellUser asserts that he derives pleasure from this action. While identifying trolls is presented in this extract as commonplace, taking pleasure in trolling them is presented as newsworthy. Trolling actually causes pleasure (lines 02–03). The rhetorical question is positioned around this notion of pleasure, in order to flag this concept as surprising.

20 The subjects of ShellUsers behaviour are once again (tacitly) ascribed a number of negative characteristics. They are described as 'goofballs' and 'liars' (line 03) and, it is implied, 'sadists' (lines 04–05). In presenting trolls as having these negative characteristics, ShellUser can be seen to legitimate his own behaviour towards them, further emphasising the newsworthiness of his taking pleasure in trolling them in return.

25 ShellUser seeks to differentiate his activity from that of genuine trolls through orientation to the purpose of his action. While genuine trolling is concerned with disruption of an online space, and engaging users in fruitless debates (Binns, 2012; Bishop,

2012a), ShellUser's activity does not share this intent. Rather, ShellUser aims to expose trolls by unravelling their own arguments (at least, sometimes – line 01). ShellUser finishes this extract by redefining his behaviour away from trolling, as well as closing the rhetorical question the extract began with. ShellUser repositions his actions as 'heroic' (line 05), since his targets are deserving (line 03).

In this way, ShellUser manages to avoid having to adopt the identity of an online troll, despite fitting the profile of a troll and admitting to enjoying engaging in activities intended to aggravate other posters. In common with extracts 01 and 02, above, ShellUser can be understood as implying (if not stating) that there are times when 'trolling' is an acceptable behaviour, and something which in-group members may engage in against legitimate out-group targets. ShellUser is able to engage in trolling, without becoming a troll.

Extract 06 is taken once again from an online British broadsheet newspaper. Extract 06 comes two posts after extract 01, and is defending Arfur Sixpence from the question by another poster as to whether his words admission of making a troll 'dance' would make Arfur himself a troll. In the extract, Opusfra can be seen explicitly refuting the appellation.

3.3.2. *Extract 06: Opusfra, the Independent*

- 1 NO - baiting a troll is an honourable and important function. We have to
- 2 expose trolls to their own true nature to make them see just how hideous
- 3 they are. That is not trolling, just holding up a mirror.

As part of presenting 'trolling' as acceptable, posters have implied that one function of trolling is to attack and expose others who engage in malicious trolling (extracts 01, 02 and 05). Thus, this form of trolling can be seen as a heroic act of vigilantism (as in the previous extract). Here, teasing trolls is presented as actively desirable – honourable and important (though the poster still repositions this action as baiting rather than trolling – line 01).

That trolls are legitimate targets of such action is reiterated at line 03, where Opusfra explicitly rejects the notion that he himself might be a troll. Instead, trolling or baiting trolls is repositioned as simply reflecting the unpleasantness of others.

Trolling against legitimate targets is repositioned as both ‘honourable’ and ‘important’ rather than ‘aggressive’ or ‘illiterate’ (extract 04). Trolling trolls is presented as
15 being an activity vital to the healthy of the forums, as it serves as a means of reflecting malicious trolling back on the posters, so that they may realise who they are. The extract finishes with the poster explicitly challenging the definition of this action as ‘trolling’. Instead, it is presented as being ‘simply’ reflecting back the behaviours of genuine trolls.

20 Extract 07 comes from the same comments section of the online blog as extract 03. The extract comes from the beginning of the comments section however, being the fourth comment after the published summary of the main article.

3.3.3. *Extract 07: David Riker, American Blog*

1 I’m a troll. I admit it. But my ”victims” are bullies themselves. Remember
2 the term ”flaming?” Well, I flame a lot of people who, for example, spout
3 stubbornly hateful and vicious portrayals of President Obama. I flame
4 bigots and climate change deniers. I flame science deniers and anti-gay
5 halfwits. These people aren’t dainty, vulnerable opinionators. They are
6 hardened, irrational haters. They deserve all the ridicule that I pubically
7 dump on their heads. Sometimes, the only reply you can give to these
8 people is a good, hard slap on the face. Are they sometimes offended?
9 GOOD.

10 Extract 07 begins with an apparent admission of the posters own status as a troll (line 01). This may be considered especially surprising, as the poster seems to be using a plausibly real name, rather than obvious pseudonym. Having admitted his status, this poster immediately seeks to first justify his actions, and then to mitigate the label of ‘troll’.

15 David Riker seeks to distance himself from the category ‘troll’ at line 02, in mobilising the term ‘flaming’ to describe what he does (in a similar manner to earlier posters repositioning their actions as ‘baiting’ rather than trolling). Interestingly, flamers here are presented not as a subtype of troll (as with ‘flame trolls’ – Bishop, 2013a) but as a lesser category of action which may be considered more acceptable than trolling.

20 David Riker introduced the concept of flaming with a rhetorical question (lines 01–02). This serves to downgrade his actions from the initially cited ‘trolling’ and instead claiming this less contentious action as the appropriate label.

David Riker further seeks to justify his actions through an appeal to the legitimacy of his targets, as with extract 05. David Riker asserts that the ‘victims’ of his trolling are themselves bullies (line 01). He further legitimates this target group by placing the term
25 ‘victims’ in inverted commas – a common literary device used to indicate a knowing uncertainty about the legitimacy of the term used. Thus, while acknowledging that these people are the targets of his flaming, David Riker manages to bring into question the notion that these people should be regarded as victims.

30 These victims then are presented as having a variety of undesirable characteristics, temperaments and attitudes. Their negative characteristics are presented in a three-part list (Jefferson, 1990) as being stubborn, hateful and vicious (line 03). Each element of this list may be seen to escalate the severity of the supposed victims own inappropriate attitudes. Stubbornness becomes hatefulness, which in turn becomes viciousness.

35 This three-part list detailing the characteristics of those who are trolled is itself embedded in a larger three-part list detailing examples of people whom David Riker flames. They include people who dislike the president of the United States of America, people who are bigots and deny climate change, and finally people who deny science generally and whom are ‘anti-gay halfwits’ (lines 04–05). Each element of this second
40 list is constructed of at least two clauses, giving the impression of a wide variety or deserving targets for this posters ire.

The legitimacy of David Riker’s targets is finalised at line 05, this time embedding the three-part list into a dyadic extreme case formulation (EFC – Pomerantz, 1986), which in turn draws upon the cognitive thesaurus (Potter & Edwards, 2003). These
45 ‘victims’ are ‘hardened, irrational haters’. EFC’s are typically used by speakers to state the strongest form of available argument, in order to reduce the probability of dispute, or to emphasise the legitimacy or illegitimacy of a set of actions. Each element in the list represents an escalation of the previous element, and in doing so makes the category being mobilised – legitimate target rather than illegitimate victim – seem more
50 plausible.

The cognitive thesaurus is used to present illegitimate targets for trolling as simply having opinions, or being ‘dainty, vulnerable opinionators’ (line 05), liable to take disproportionate offence should their opinions be disagreed with. Legitimate targets however are presented as being irrational, and this lack of rationality is linked to their emotional state of ‘hate’ (Edwards, 1999).
55

Extract 07 ends as did extract 05. David Riker likens his online actions to a form of physical violence – ‘a good, hard slap’ (line 08). This is presented as a form of capital punishment and as with extract 05, is presented as proportional and appropriate.

Throughout this post, the nature of the targets of David Riker’s trolling is more elaborated upon than is the nature of his trolling, indicating that this poster is more
60 concerned with justifying whom he trolls, than his is in defending his actions as a troll. If his victims can be presented as legitimate targets, then his actions will no longer need defending.

Posters are often seen asserting that they are able to troll the trolls, in order to expose them (extract 05, lines 01–02). This might be understood as a form of ‘gamification’ itself (Bishop, 2012a), with posters attempting to engage the trolls in fruitless debate, in order to frustrate their would-be attackers purpose and instead expose them for what they are (extract 05, line 02). Posters seem unaware of the possibility that this instead feeds the trolls agenda, as taking time to bait the trolls serves to disrupt the
70 thread conversation, thus allowing the trolls aims to be achieved.

Trolling here is presented as being a potentially positive act, though it is still not one which can be easily and readily admitted to. The positive purpose trolling serves is that of ‘holding up a mirror’ (extract 06) to undesirable elements on the Internet, revealing their own awfulness to them, and of policing the forum, in order to discourage their
75 behaviour (though at the cost of establishing a norm of civility) (Lampe et al., 2014).

3.4. Repertoire 04: Trolls are nasty

Despite posters previous assertions that trolling can be acceptable when directed against deserving targets (extracts 05–07), and that trolling was once an art form to be appreciated (extract 01), we see in extract 07 that speakers may still exhibit difficulty
80 adopting the label ‘troll’ themselves. The term ‘troll’ is still a negatively loaded phrase,

and one which posters are careful to distance themselves from.

This is exemplified in the tacit repertoire of trolls being having antisocial characteristics. The extracts presented so far have presented ‘trolls’ as being ‘sullen and embittered’ (extract 01), swivel-eyed (extract 02), illiterate, offensive and unskilled (extract 85 04). Trolls are sadistic liars (extract 05), hideous (extract 06), irrational bigots (extract 07). The unpleasant nature of trolls is further elaborated upon in extract 08, below.

This extract is taken once again from the same blog as extracts 03 and 07. The extract presented here has been truncated, presenting only the opening and closing segments which deal with the nature of trolling.

90 3.4.1. *Extract 08: Monophylos Fortikos, AmericaBlog*

1 I’ve done trollish things, hon. I’ll admit that right now, I’ve said insulting
2 things with the deliberate intention of needling someone and provoking
3 them to make an angry response. That is why your remark was trollish.
4 Properly speaking, being a troll isn’t about calling names or hurling in-
5 sults; it’s about baiting someone into losing their temper, which is some-
6 thing that can be done without names or insults.

7 ...

8 And, yeah, I’ve done it myself and will probably continue to do it when-
9 ever I’m angry or depressed or otherwise in a state of mind not to think
10 about the consequences of my behavior. I’m not proud of it, though. You,
11 on the other hand, don’t seem much to care.

As with extract 07, above, Monophylos Fortikos begins with an admission of being a troll. Unlike extract 07, the only mitigation provided is that the behaviour is ‘trollish’ rather than being fully committed to being a troll (line 01).

15 In alignment with the definitions of trolling given in academic literature, ‘trollish behaviour’ is defined here as an intentional act – insulting other individuals online for the sole purpose or angering them (line 02–03). Further elaboration is given between lines 04–06, in that trolling does not just involve insults. ‘Proper’ trolling is about manipulating another individual into losing their temper. This may be compared with

20 extract 04, above, which bemoaned the loss of the art of trolling. While causing deliberate offence is an important element of trolling, it is not the only part. Further, 'simple' name calling and insulting other individuals is not considered an essential part of the art form.

25 While Monophylos does not attempt to downgrade his trollish behaviour, he does attempt to excuse it. The behaviour is still considered negative, but it is presented as being beyond this posters active control. Monophylos reiterates their trollish behaviour at line 08, before elaborating upon the conditions under which he engages in trolling through the appeal to the emotional thesaurus. These conditions are presented as a three part list, with an open completer (Jefferson, 1990). This list draws upon discourses of 30 emotional states (Edwards, 1999) to posit two specific mental states which may be taken to affect the posters decision-making ability (anger and depression). The open completer is used here as a catch-all term for any state which absolves the poster of culpability for their action, as well as making explicit the purpose of the initial two states – Monophylos would 'not be in a state of mind' to appreciate the consequences 35 of his actions. Thus we see that while this poster has behaved in a trollish way, they are not responsible for such behaviour and thus may still be considered a valued member of any online community. Trolling is a compulsion, engaged in when Monophylos is not able to control his emotional state and thus in turn, not disposed to act rationally.

40 Monophylos further mobilises the emotional lexicon at line 09, when he makes reference to *pride*. That trolling is not something to be proud of further flags this kind of behaviour as something undesirable. This may be seen as contradicting earlier extracts however, where trollish behaviour was presented as a humorous aside (extract 04), something to wax nostalgic about (extract 06) or something which may be used to exert control over 'genuine trolls' (extract 07).

45 This extract then allows the poster to behave in a trollish manner without becoming a troll themselves. From the extracts presented so far then, we see that while trolling may be acceptable under certain circumstances, being a troll is not acceptable. It is an accolade which posters go to great lengths to avoid.

3.5. Identifying trolls

50 One implication of the repertoires discussed above is that members of the online community are able to identify the trolls in their midst. Once a troll is identified however, how do online community members identify these individuals to each other? The process of publicly identifying another poster as being a troll is not so straight forward. As has been demonstrated above, being a troll is still perceived as holding negative
55 connotations. As such, despite earlier protestations, it is not easy to explicitly state that another poster is a troll. Posters are reluctant to directly accuse each other of being trolls, even when said posters are making contentious statements.

This pattern of behaviour is exemplified in extract 08, above. In this extract, Monophylos responds to another online user, whom he suspects of being a troll. However,
60 Monophylos is not able to simply accuse this other poster of being a troll – despite the implication in extracts 01–03 that trolls are known individuals, and are easily identifiable. Rather, Monophylos moves through a number of steps in order to be able to challenge the trollishness of this other forum user.

Monophylos begins with an apparent admission to having done trollish things himself (line 01), along with a definition of what ‘trollish things’ might include (lines
65 02–06). As has been shown above, Monophylos also takes measures to ensure that, having identified as trollish, he is not further identified as a troll.

Having tentatively adopted this trollish identity, Monophylos then goes on to accuse another poster of also engaging in trollish activity (line 03). In this context, we
70 can understand Monophylos’ admission of trollishness as an attempt to claim category entitlement (Potter et al., 1993). Monophylos is able to identify trollishness in the other poster at line 03, because Monophylos engages in trollishness themselves (line 01).

Being trollish, rather than being a troll is a less contentious accusation to make against another online poster. This label allows both posters plausible denial – both
75 could claim to have been taken the wrong way in what they posted, or to have misunderstood the other. However, Monophylos escalates his claims against the other poster towards the end of the extract, at lines 10–11. Once again, this is accomplished through the appeal to the emotional discourse. What shields Monophylos from being placed in the category ‘troll’ is his mental state when engaging in trollish activities, and

80 his acknowledgement that such action is not something to be proud of. Such emotional states are not applied to the poster to whom Monophylos directs his remarks. This other poster, seems not to care about the consequences of their actions. As this other poster has implied disinterest in the negative consequences of their trollish action, they may be taken as being a ‘troll’.

85 Extract 09 provides a final illustration of the difficulty exhibited by forum posters in identifying – publicly – other posters who may be trolls. The extract is taken from a forum dedicated to discussing science related information. This post comes towards the end of the thread.

3.5.1. *Extract 09: Trippy, SciForums*

- 1 Your posts are loaded with passive aggressive innuendo and double mean-
- 2 ings. I’m not the only one that sees it, and I’m growing tired of it.

In extract 09, Trippy addresses another poster on the same thread. At this point in the interaction, each poster has been interacting from some time. Trippy now feels
5 confident enough to move towards accusing this other poster of engaging in trollish behaviours – here defined as passive-aggressive innuendo (compare with trollishness being defined as attempting to provoke anger, extract 08, and with the definitions cited in the introduction). At line 02, Trippy makes an appeal to consensus – he is not the only one who can detect the innuendo. Consensus is dropped however when Trippy
10 asserts that he (only) is growing tired of this behaviour.

What is important about both extracts 08 and 09 is that despite trolls being readily identifiable and known (see extracts 01–03), neither poster here is willing to explicitly and directly accuse another forum user of being a troll. Rather, they are accused in engaging in disruptive online behaviours, which may be considered ‘trollish’.

15 **4. Discussion**

Far from being a clearly defined and well understood terminology, trolling is a complex activity. The above analysis demonstrates that the category ‘troll’, and its

associated activity of 'trolling' are disputed, contentious phrases. Their meanings are actively negotiated during the course of asynchronous online interactions.

20 The meaning of the term 'troll' is co-constructed by posters to the online fora and comment threads according to the posters rhetorical needs at the time of their posting. Both trolling and trolls may be constructed as either positive or negative (though not both at the same time).

Past research, and media interest, has most often focused on the negative aspects
25 of trolling. Posters here identify a type of trolling which is consistent with previously identified types of 'flame trolling' (extract 08) (Bishop, 2012b). Posters claim such flame trolling is easy to identify, and similarly easy to address.

A second type of trolling, while still retaining the characteristics of flame trolling, is presented as 'acceptable trolling'. This acceptable trolling is has been shown to further
30 divide into two broad types. Acceptable trolling may be a form of flame trolling conducted by others, albeit ineptly. This inept trolling is easy to identify for posters, and thus fails to annoy or to draw into pointless debate. Rather, this inept trolling provides a source of amusement and nostalgia rather than aggravation and annoyance. Trolling of this nature lacks the characteristics identified by Hardaker (2010), as although trolls
35 may still display dishonesty in seeking to hide their motivations for posting to the forum, and display aggression in attempting to disrupt the nature of online interactions, they are not successful as these individuals are easily identified as trolls by the savvy members of the online community.

Acceptable trolling is also presented as an activity which the ingroup condones.
40 This may be seen in the case of nostalgia (extracts 01; 04). Nostalgic trolls are easily identifiable, yet no action was taken to censure them or remove them from the online space. Indeed, when the trolls ceased their activity of their own accord, posters discourse suggests that these trolls are missed. It may therefore be inferred that trolling had been in some way 'normative' or 'desirable'.

45 This 'nostalgic troll' aligns well with the classic notion of the online troll. The nostalgic troll is presented as being petty, small-minded and delighting in causing mischief for mischief's sake. However, the nostalgic troll is an ultimately harmless annoyance, disempowered by how easy it is for the members of the online space to identify their

intentions and respond appropriately. This troll, despite having the intention to disrupt
50 the online community, is presented as inept and incompetent, unable to properly rile the
sophisticated members of the online community. They are not successful. As such, the
troll is more amusing than aggravating, and is welcomed as an unintentional member
of the online space. This troll is fondly remembered, and missed.

Acceptable trolling is also presented as an activity which posters themselves ac-
55 tively engage in. Throughout the repertoire of vigilantism (extracts 05–07), posters
claim to ‘troll the trolls’ in order to expose the trolls. Posters themselves then engage in
troll like behaviours, while attempting to avoid being categorised as trolls themselves.

Broadly speaking, whether posters represent trolling as positive or negative de-
pends upon the business at hand. When trolling is being done by people who are not
60 regular posters to the online space, it is presented negatively, even when posters are
claiming it has nostalgic value. This kind of trolling is acknowledged as being con-
ducted for the purpose of disrupting online interactions and communications – ‘flame
trolling’ (extract 08). The business at hand then is not that of defending the ingroup but
rather that of attacking the outgroup. Trolls are those who transgress group norms and
65 threatening the smooth running of online interactions.

When trolling is being done by individuals who are regular posters to the online
space, it is presented more positively. Being a troll however is still presented as some-
thing negative, a categorisation which must be avoided even while posters are appar-
ently confessing to being a troll. Ingroup trolling is acceptable as these posters are
70 trolling for good, in order to expose genuine trolls through their web of deceit and lies
and thus to discourage further misbehaviour from this group.

Here, trolling is rendered acceptable through its inauthenticity. Trolling by mem-
bers of the ingroup is not intended to disrupt online communications or lure others into
fruitless debate. Rather, this trolling is intended to expose others who troll maliciously.
75 Trolling is represented as a legitimate means of retaliation and retribution against those
who would seek to disrupt the online community. While this account of acceptable
trolling does meet all four (negative) criteria of trolling identified by Hardaker (2010),
it does not retain the negative connotations and consequences of the action. Rather, it
is presented as acceptable through posters orientation to the legitimacy of their targets

80 – genuine trolls.

How trolls are represented by members of the online community seems to depend upon the nature of the business at hand, and the context in which posters are writing. In the early stages of each forum thread, posters are defending against an outside threat – in this case, academics who seek to define the nature of the online community (as a place riddled with trolls). As such, the nature of the threat is not the trolls themselves, but rather the notion that the online community is overrun with trolls. Here, to cast trolls as negative entities also serves to problematise the posters to the online forum, as a number of them have admitted to engaging in trolling as a strategy for controlling the undesirable trolls.

90 Later in the threads, the threat moves from academic discourse and towards the trolls themselves as trolls are cast as a disruptive influence on the forum and online community. It is at this point that posters present trolls as being somehow pathological or undesirable.

There are a number of implications from these discourses. One such issue concerns the identification of genuine, disruptive trolls, who are members of the outgroup, and members of the ingroup who just happen to be engaging in troll-like activities. There are further implications for strategies for troll management.

While posters are keen to state (or imply) that trolls are easy to identify (extracts 01–03), their interactions with other members of the online community show that it is a delicate negotiation to be able to begin to explicitly accuse another poster of being a troll (extract 09). Identifying whom the genuine trolls are may be further complicate by the notion that some forms of trolling (vigilantism) are acceptable (extracts 05–07), and some forms of troll (nostalgia) may be considered desirable (extract 04).

In all fora, posters assert that they are able to identify trolls. However, no poster provides a comprehensive account of how trolls can be identified. This might be considered especially unusual, as posters do claim at various points that there is a simple solution to the ‘problem’ of trolling – that of ‘do not feed the trolls’ (extract 03. See also Binns, 2012). However, from looking at the patterns of behaviour exhibited online, we see that posters to online forums are reticent to actually accuse other members of their online communities of being trolls, rather instead spending many turns of conversation

engaging with such individuals before finally challenging them on the genuineness of their trolling behaviours. It is this ambiguity which contributes to the ‘success’ characteristic of trolling.

However, in presenting trolls in a positive light, online community members create a problem for themselves. Namely, genuine trolls themselves have been cast as essentially harmless and inept despite their malicious intent. Trolling itself has been presented as a legitimate activity, performed in response to the presence and attempts of inept trolls. Extract 08 shows how posters manage this concern by creating a third version of the troll. In this version, the troll is not harmless, and is not a valued member of the community. Rather, the troll is dangerous, pathological and disruptive. This final version of the troll is negatively valued, disruptive and not a valued member of the online community. However, they are necessary to justify posters earlier animosity, displayed towards trolls.

These discourses of trolling hold implications for troll management. The simple resolution of ‘don’t feed the trolls’ is abandoned in favour of ‘troll the trolls’ (extracts 05–07). This action may be argued to be ‘gamified’ by the posters themselves, as they derive amusement and pleasure from such action.

The repertoire of ‘vigilantism’ demonstrates that ‘trolling the trolls’ is already considered a legitimate action online, with the reward being that genuine trolls rather than online forum members will be the ones who are aggravated, or drawn into fruitless argumentation.

Problems may also be found with the strategy of cultivating a behavioural norm of civility within the online forum (Lampe et al., 2014). In trolling trolls, established posters will be seen to behave badly by new and old members of the forum alike. The likely outcome then would be to establish a norm of uncivil behaviour amongst all members of the online space.

5. Conclusion

Following assumptions of discursive psychology, categories are rarely supposed to have a single, fixed meaning. Thus the meaning of the terms ‘troll’ and ‘trolling’ are

140 taken as constant and fixed but rather as variable throughout the course of an inter-
action, and between different fora. Other papers researching this topic have assumed
that there is one fixed meaning for ‘troll’, albeit there are also many sub-classifications.
Furthermore, this paper combines data from a number of online sources, following the
ebb and flow of the meaning of these terms within users discussions within a number
145 of fora, rather than just withing one source.

From this analysis, it can be seen that neither the category ‘troll’ nor the action
of ‘trolling’ has a single, fixed meaning. Rather, a discursive analysis has revealed a
tension in posters constructions of trolling.

Trolls and trolling may be presented as undesirable. Such trolls are presented as
150 possessing a number of negative characteristics. They are ‘sullen and embittered’ (ex-
tract 01), swivel-eyed (extract 02), sadistic liars (extract 05) and illiterates, taking of-
fensive, unskilled cheap shots at other forum users (extract 04). Trolls are hideous
(extract 06), irrational bigots (extract 07).

Yet trolling can be considered acceptable too, under certain circumstances. Plea-
155 sure may be derived from baiting trolls (extract 01), as trolls are legitimate targets of
online abuse (extract 06). They should be given a ‘dismissive drubbing’ (extract 02)
in order to expose their true nature (extract 05). Such is an honourable and important
activity (extract 06).

These divisions however may not necessarily be genuine sub-varieties of trolls and
160 trolling. Rather, they represent different available repertoires of trolling, deployed
rhetorically in order to accomplish specific actions in talk – defending the ingroup
or attacking the outgroup, as the needs of the interactional context demand. Modern
trolling is presented as simply ‘nasty’, while old-fashioned trolling is sophisticated and
elegant. Despite these distinctions, vigilantism – trolling trolls – is presented as being
165 honourable, noble and necessary.

5.1. Implications

The major contribution of this paper has been to show that trolling, as with all
discursive categories, shows a fluidity of meaning both within and between users of
online spaces. The same acts of trolling may be presented either positively or nega-

170 tively, depending upon the aims of the posters, and the avowed intent of the trolling actions.

The analysis also shows implications for the management of trolling both for online spaces and for legislators. Traditionally recommended techniques for troll management may prove problematic or ineffective, due to the meanings and functions attached to trolling. Ignoring trolls is not feasible when members of the online space use trolling as a means of attaching trolls. For this reason too, trolls cannot simply be banned from online spaced. This also inhibits the establishment of good behaviour online.

Legislation action cannot be take either, for example in simply making trolling behaviours illegal, as on occasions such behaviours may be ambiguous, represented as desirable, prosocial or otherwise not counter normative. Hosts of online spaces may also struggle to establish norms of civility or to employ proper moderation as a means of troll management, where users of online spaces may be deploying online incivility as a means of troll management.

References

- 185 Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The “nasty effect:” Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, *19*, 373–387. doi:10.1111/jcc4.12009.
- 190 Andrews, T. (2012). What is social constructionism. *The Grounded Theory Review*, *11*, 39–46. URL: <http://groundedtheoryreview.com/2012/06/01/what-is-social-constructionism/>.
- Bauman, S., Toomey, R. B., & Walker, J. L. (2013). Associations among bullying, cyberbullying, and suicide in high school students. *Journal of Adolescence*, *36*, 341–350. doi:10.1016/j.adolescence.2012.12.001.
- 195 Binns, A. (2012). DON’T FEED THE TROLLS! Managing troublemakers in magazines’ online communities. *Journalism Practice*, *6*, 547–562. doi:10.1080/17512786.2011.648988.

- 200 Bishop, J. (2012a). The psychology of trolling and lurking: the role of defriending and gamification for increasing participation in online communities using seductive narratives. In H. Li (Ed.), *Virtual community participation and motivation: Cross-disciplinary theories* (pp. 160–176). Hershey, PA: Information Science Reference. doi:10.4018/978-1-4666-0312-7.ch010.
- 205 Bishop, J. (2012b). Tackling internet abuse in great britain: Towards a framework for classifying severities of ‘flame trolling’. In *The 11th International Conference on Security and Management (SAM’12)* (pp. 16–19).
- Bishop, J. (2013a). The art of trolling law enforcement: A review and model for implementing ‘flame trolling’ legislation enacted in Great Britain (1981–2012). *International Review of Law, Computers & Technology*, 27, 301–318. doi:10.1080/13600869.2013.796706.
- 210 Bishop, J. (2013b). The effect of de-individuation of the internet troller on criminal procedure implementation: An interview with a hater. *International Journal of Cyber Criminology*, 7, 28–48. URL: www.cybercrimejournal.com/Bishop2013janijcc.pdf.
- 215 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, 7–24. doi:10.1504/IJWBC.2014.058384.
- Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, 67, 97–102. doi:10.1016/j.paid.2014.01.016.
- 220 Burr, V. (2003). *Social Constructionism*. London: Routledge.
- Butler, D., Kift, S., & Campbell, M. (2009). Cyber bullying in schools and the law: Is there an effective means of addressing the power imbalance. *eLaw Journal: Murdoch University Electronic Journal of Law*, 16, 84–114. URL: <http://heinonline.org/HOL/LandingPage?handle=hein.journals/elajrnl16&div=7&id=&page=>.
- 225

- Cole, K. K. (2015). "It's like she's eager to be verbally abused": Twitter, trolls, and (en) gendering disciplinary rhetoric. *Feminist Media Studies*, 15, 356–358. doi:10.1080/14680777.2015.1008750.
- Condor, S. (2000). Pride and prejudice: Identity management in english people's talk about 'this country'. *Discourse & Society*, 11, 175–205. doi:10.1177/0957926500011002003.
- Condor, S. (2006). Public prejudice as collaborative accomplishment: Towards a dialogic social psychology of racism. *Journal of Community & Applied Social Psychology*, 16, 1–18. doi:10.1002/casp.845.
- Condor, S., Figgou, L., Abell, J., Gibson, S., & Stevenson, C. (2006). 'they're not racist...' Prejudice denial, mitigation and suppression in dialogue. *British Journal of Social Psychology*, 45, 441–462. doi:10.1348/014466605X66817.
- Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification using game-design elements in non-gaming contexts. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 2425–2428). ACM. doi:10.1145/1979742.1979575.
- Edwards, D. (1998). The relevant thing about her: Social identity categories in use. In c. Antaki, & S. Widdicombe (Eds.), *Identities in talk* (pp. 15–33). London: Sage.
- Edwards, D. (1999). Emotion discourse. *Culture & Psychology*, 5, 271–291. doi:10.1177/1354067X9953001.
- Elliott, R. (1996). Discourse analysis: exploring action, function and conflict in social texts. *Marketing Intelligence & Planning*, 14, 65–68. doi:10.1108/02634509610131171.
- Ethics Committee of the British Psychological Society (2009). Code of ethics and conduct: Guidance published by the ethics committee of the British Psychological Society. URL: <http://www.bps.org.uk/what-we-do/ethics-standards/ethics-standards>.

- Galán-García, P., de la Puerta, J. G., Gómez, C. L., Santos, I., & Bringas, P. G. (2014). Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying. In *International Joint Conference SOCO13-CISIS13-ICEUTE13* (pp. 419–428). Springer. doi:10.1007/978-3-319-01854-6_43.
- 255
- Griffin, C. (2007). Being dead and being there: research interviews, sharing hand cream and the preference for analysing ‘naturally occurring data’. *Discourse Studies*, 9, 246–269. doi:10.1177/1461445607075340.
- 260
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research*, 6, 215–242. doi:10.1515/JPLR.2010.011.
- Herring, S., Job-Sluder, K., Scheckler, R., & Barab, S. (2002). Searching for safety online: Managing” trolling” in a feminist forum. *The Information Society*, 18, 371–384. doi:10.1080/01972240290108186.
- 265
- Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research*, 14, 206–221. doi:10.1080/13811118.2010.494133.
- Jefferson, G. (1990). List construction as a task and resource. In G. Psathas (Ed.), *Interaction Competence* (pp. 63–92). Washington, DC: University Press of America.
- 270
- Lamerichs, J., & Te Molder, H. F. (2003). Computer-mediated communication: From a cognitive to a discursive model. *New Media & Society*, 5, 451–473. doi:10.1177/146144480354001.
- Lampe, C., Zube, P., Lee, J., Park, C. H., & Johnston, E. (2014). Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly*, 31, 317–326. doi:10.1016/j.giq.2013.11.005.
- 275
- Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, 3, 229–235. doi:10.1111/j.1467-9280.1992.tb00034.x.
- 280

- Pomerantz, A. (1986). Extreme case formulations: A way of legitimizing claims. *Human Studies*, 9, 219–229. doi:10.1007/BF00148128.
- Potter, J. (2003). Discursive psychology: Between method and paradigm. *Discourse & Society*, 14, 783–794. doi:10.1177/09579265030146005.
- 285 Potter, J. (2012). Discourse analysis and discursive psychology. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 119–138). Washington, DC, US: American Psychological Association. doi:10.1037/13620-008.
- 290 Potter, J., & Edwards, D. (1999). Social representations and discursive psychology: From cognition to action. *Culture & Psychology*, 5, 447–458. doi:10.1177/1354067X9954004.
- Potter, J., & Edwards, D. (2003). Sociolinguistics, cognitivism and discursive psychology. *International Journal of English Studies*, 3, 93–109. doi:10.6018/48531.
- 295 Potter, J., Edwards, D., & Wetherell, M. (1993). A model of discourse in action. *American Behavioral Scientist*, (pp. 383–401). doi:10.1177/0002764293036003008.
- Potter, J., & Hepburn, A. (2007). Life is out there a comment on griffin. *Discourse Studies*, 9, 276–282. doi:10.1177/1461445607075348.
- 300 Potter, J., & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour*. London: Sage.
- Ruiz, C., Domingo, D., Micó, J. L., Díaz-Noci, J., Masip, P., & Meso, K. (2011). Public sphere 2.0? The democratic qualities of citizen debates in online newspapers. *The International Journal of Press/Politics*, (p. 1940161211415849). doi:10.1177/1940161211415849.
- 305 Shachaf, P., & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science*, 36, 357–370. doi:10.1177/0165551510365390.

- Shin, J. (2008). Morality and internet behavior: A study of the internet troll and its relation with morality on the internet. In *Society for Information Technology & Teacher Education International Conference* (pp. 2834–2840). volume 2008. URL:
310 <http://www.editlib.org/noaccesspresent/27652/46654/>.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, 7, 321–326. doi:10.1089/1094931041291295.
- Thacker, S., & Griffiths, M. D. (2012). An exploratory study of trolling in online video gaming. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)*, 2, 17–33. doi:10.4018/ijcbpl.2012100102.
315
- The Select Committee on Communications (2014). *Social media and criminal offences*. London: The Stationery Office Limited.
- Virkar, S. (2014). Trolls just want to have fun: Electronic aggression within the context of e-participation and other online political behaviour in the united kingdom. *International Journal of E-Politics (IJEP)*, 5, 21–51. doi:10.4018/ijep.2014100102.
320

Trolling the trolls: Online Forum Users Constructions of the Nature and Properties of Trolling

Dr. Bryn Alexander Coles*, Melanie West*

Department of Psychology and Counselling, Newman University, Bartley Green, Birmingham, B32 3NT, United Kingdom

Many thanks are due to Dr. Genovefa Kefalidou of the Human Factors Research Group, the University of Nottingham for her invaluable comments during the drafting of this paper.

*Corresponding author

Email address: a.coles@newman.ac.uk (Dr. Bryn Alexander Coles)