

Chapter 28

Trolling Is Not Just a Art. It Is an Science: The Role of Automated Affective Content Screening in Regulating Digital Media and Reducing Risk of Trauma

Jonathan Bishop

Centre for Research into Online Communities and E-Learning Systems, Wales, UK

ABSTRACT

This chapter seeks to explore the role media content ratings play in the age of “Internet trolling” and other electronic media issues like “sexting.” Using ANOVA to validate a four-factor approach to media ratings based on maturity, the chapter finds the ability of a person to withstand various media content, measured in “knol,” which is the brain’s capacity to process information, can be used to calculate media ratings. The study concludes it is feasible to have brain-computer interfaces for PCs and kiosks to test the maturity of vulnerable persons and recommend to parents/guardians or cinema managers whether or not to allow someone access to the content they wish to consume. This could mean that computer software could be programmed to automatically censor content that person is likely to be distressed or grossly offended by. Public policy issues relating to these supply-side interventions are discussed.

INTRODUCTION

The convergence of media content is posing challenges in terms of protecting the vulnerable while also protecting free speech. Regulating abusive online media content, such as Internet trolling, is not suited to ‘before-the-fact’ model of current film rating agencies even if the objective rules and judgements used by those agencies are. Basing rating on age appropriateness is also ineffective for the online world, as people will have different

maturity for different content at different ages with some being more prone to offense than others.

Internet trolling as a concept has transformed in definition in recent years from classical trolling, which was the posting of messages in a friendly way, to Anonymous Trolling, which is posted to harm others (Bishop, 2014b; Phillips, 2011). Internet trolling messages posted to entertain others can be seen as ‘kudos trolling’ and those designed to harm others can be seen as ‘flame trolling’ (Bishop, 2012b). But this does

DOI: 10.4018/978-1-4666-6324-4.ch028

Trolling Is Not just a Art. It Is an Science

not mean all flame trolling is 'bad' and should be punished, nor does it mean all kudos trolling is 'good' and should be allowed. In some forums on the Internet, flame trolling is encouraged, such as the criticism of politicians, bankers, or other people who may be part of a group with dislikeable qualities. The consensual nature of these forums it could be argued should not mean their abusive comments should be prosecutable (Starmer, 2013). Equally not all kudos trolling is designed to be in the interests of people. For instance, a type of online community user called a chatroom bob, will often post friendly comments in order to seduce others. They may be a pervert looking to coax naked pictures out of the person, or a sex predator trying to groom a child or other young person (Bishop, 2012c; Jansen & James, 1995; Jansen & James, 2002). A recent example of this was the case of Daniel Perry. Daniel Perry was a 17-year-old man who was tormented into killing himself as part of an online plot to extort money from him. Perry was a popular teenager from Dunfermline, Fife, who took his own life after being targeted by a group of Internet trolls who exploited him for their own gain. Perry took his own life after a Skype conversation with a person he was led to believe was a girl the same age as him, where he shared sexualized videos of himself online known as 'sexting' or 'getting naked on cam' (GNOC). Suddenly a gang then hijacked the chat and threatened to show the video to his family and friends, unless he paid them menaces (i.e. money). Known as 'Nigerian Chatroom Bobs,' due to an urban myth that it is mostly people in Nigeria who extort money from people online, these trolls will use all means to gain the confidence of someone and then extort money from them, which in Daniel Perry's case was attempted through blackmail.

It is therefore necessary to have a more technical way of looking at Internet trolling and other online misdemeanours, such as through linguistic or other forms of studying media (Bishop, 2014b; Hardaker, 2013; Hardaker, 2010). This could in-

volve making it easier to regulate online content, so that mature users know what to expect, and Internet security software providers can better produce software with parental controls to avoid the corruption of the minds of those lacking in maturity (Haravuori, Suomalainen, Berg, Kiviruu, & Marttunen, 2011; Roche, 2012). Such people may not have been exposed to severe or traumatic content in the past and as a result be less able to cope with it (Dutta-Bergman, 2006). It may therefore be appropriate to refer to those Internet trolling messages which are prosecutable as electronic message faults (EMFIs) and those which are not prosecutable as electronic message freedoms (EMFrS).

There appears in the UK to be a cycle of youth justice where young people are at one point seen unfavourably and at another there is increased concern for them (Bernard, 1992; Weijers, 1999). Some have said that whilst there is a developing international literature exploring youth participation in community arts activities, to date relatively little attention has been paid to issues surrounding young people's decision-making within participatory arts projects (Rimmer, 2012). A recent study in Wales has shown how the significant involvement of young people in deciding the outcome of community arts projects rather than simply being led by controlling adults who have their own outdated ideas has seen positive outcomes for communities (Bishop, 2012a). Understanding the effect of the arts and other cultural pursuits on the engagement of persons within society in general is proving to be a particular challenge for such a complex range of activities, sites and settings for arts participation (Gilmore, 2012). Online film reviews and accompanying film ratings have been shown to be significant predictors of both aggregate and weekly box office revenues (Huang & Yen, 2013). It is also the case that consumer generated film ratings have a direct effect on sales (Lee, 2012). This would suggest there is significant merit in basing media ratings on the individual needs and capabilities of consumers.

Young people are usually always presented as up-to-date with the latest technology. Whether they are boffins in the 1980s (Morrow, 2001), the hackers in the 1990s, the millennials in the 2000s or the digital teens in the 2010s (Bishop, 2014a; Cavagnero, 2012), they will at one point need protecting from abusers of the technology and at another be the cause of all the abuse with that technology. It has become clear that whether or not media sources cause crime or behavioural problems that the presence of a stimulus that activates traumatised parts of the brain can affect the sensory perceptions of the persons exposed to them (Bishop, 2012c; Dutta-Bergman, 2006).

BACKGROUND

The rating of media content has a different regime in most jurisdictions in the world. According to (Smartt, 2011), in the United Kingdom it is the British Board of Film Classification (BBFC) that is responsible for the rating of media content, such as TV, DVD, video games, etc. They say that each year the BBFC rates over 10,000 media sources and in 2010 alone, of 407 films censored there were five that were cut (1.2 per cent), which compared to 1983 where 123 out of the 514 (24 per cent) were cut. (Smartt, 2011) also indicated that the BBFC have the competence to rate a piece of audio-visual media content where it falls into a number of categories. These include where the content is; criminal behaviour, illegal drugs, violent behaviour or incidents, horrific behaviour or incidents, human sexual activity, regardless of whether it is a film or video game (Wilcox, 2011). The difficulty with some online content, such as that arising out of cyberbullying, its that it is not possible to pre-screen it like one can with traditional media, but it can be equally as shocking (Katz, 2012). One notable example of discrepancies between the regulation of film and Internet content is the case of *DPP v Collins* [2006] 1 WLR 2223. In the Collins case the defendant was found to be guilty of sending grossly

offensive messages via a public communications network through persistent phone calls to their local Member of Parliament's staff referring to "foreigners" whom Collins said should "go back to their own country." One should ask therefore, what is the reason why racism exposed in an imaginary world through film, which causes offence to the same magnitude as that by public communications networks, is any more acceptable? Why is hearing something reposted on Twitter for instance more prosecutable than equally bigoted representations of the same vile language shown in films? On this basis, one might argue that the BBFC, considered by (Smartt, 2011) to be one of the most highly regarded and trusted non-governmental organizations in the world, should be given the remit to rate Internet-based media content.

It is likely that BBFC, who have a lot of experience balancing what is simply offensive and what is a GOIOM in the case of film, could add a lot to regulating online media content, but the stronghold of basing ratings on age, are totally unsuited to the digital age we are in. Today, young people, or digital teens as they are sometimes called, are able to mature in their use of media content a lot faster than some of their ancestors were able to. There therefore needs to be a new rating system that goes beyond age, as it should perhaps be considered unacceptable to expose vulnerable people to certain content, just because they happen to be of an arbitrary age where most people have greater maturity. Indeed, the case of *DPP v Connolly* [2008] 1 W.L.R. 276 found that even among adult professionals, some are likely to find certain images (in this case aborted foetuses) to be GOIOM (in this case by pharmacists), whereas others, such as abortion surgeons, would be unlikely to see it as a GOIOM message. The regulation of media content therefore needs to adapt to go beyond age and provide greater information to the public on the actual content of the file or other media content they are accessing, and not simply assume if they are at the age of majority that it is unlikely they will find it to not be of a GOIOM nature.

THE EFFECT OF MEDIA AND THE ARTS ON CRIME AND VICE VERSA

The causal linkages between media and crime have long been spoken about. In the 1970s the alleged dangers were from television (Holland, 1972; Levi, 1979), and in the 1990s it was from 'video nasties' (Petley, 1994). Today the social-ill in our societies, so it is said, are caused by online social networking services. Every twenty years – a generational gap – there seems to be a moral panic around the affect the media has on crime and behaviour in general (Bishop, 2014b; Halloran, 1970; Schroeder, 1996). One can see that in the 1970s the television reached saturation point, by the 1990s video games had, and in the 2010s, it seems that the Internet of Things has reached that state. People now have access to the Internet and online media content from a range of devices (Pearce & Rice, 2013). This includes tablet PCs, netbooks, smartphones as well as the traditional PC and specific television and video games devices (Bishop & Mannay, 2014).

In the 1990s politicians lined up to attack the video game manufacturer, Sega, for one of its innovative full-motion-video games, 'Night Trap,' which was criticized for its brutality (Schroeder, 1996). The Conservative Member of Parliament for Birmingham Edgbaston at the time was Jill Knight, known for banning discussing homosexuality in schools and her views on wanting to outlaw Irish Republicans, "This is a new generation of videos, nastier than ever before," she said of Night Trap, "I am extremely concerned to hear of this extraordinary new direction computer games are taking. We should consider legislation against such games because they encourage people to maim, mutilate and murder." The then Member of Parliament for St Helens North and former Labour Party official, John Evans, even called for the Home Office to investigate the game: "The concept is absolutely horrifying," he said.

"How can we allow such a dreadful thing to be freely available to youngsters?" In the 2010s, the Member of Parliament for Liverpool Walton, Steve Rotherham, was vociferously calling for action against Internet trollers in the same way, "Laws of the land need to be constantly updated to reflect social and technological advancements", he said. "My intention is to see a greater conviction rate for those guilty of this vile practice."

MEASURING THE EFFECTS OF MEDIA CONTENT ON A PERSON'S BRAIN

The extent to which someone can withstand or effectively use media sources can be measured in 'knol.' In the context of the human mind, knol technically refers to the amount of pressure on parts of the brain that acts as a restriction to synaptic flow and thus the plasticity of the mind to efficiently process the requests made of it consciously by the actor, or forced on them by the environment. Those pressures, or 'impressions,' imposed on the actor may include traumatic event such as abuse of violent videos on television, which have the effect of reducing their ability to process information, measured as a 'knol.' The most researched part of the brain in relation to the calculation of knol is the prefrontal cortex (Bishop, 2011b; Bishop, 2012c; Bishop & Goode, 2014). This part of the brain is believed is responsible for many of the social interactions humans have with one another (Baarendse, Counotte, O'Donnell, & Vanderschuren, 2013; Beadle, 2009; Blair, 2007; Dickey et al., 2001; Maiza et al., 2008; Mushiake et al., 2009; Sala et al., 2011; Tanji & Hoshi, 2008). It is hypothesized that if the prefrontal cortex has a high or optimal knol of 0.81 then the flow of information, or neuro-response plasticity, is likely to be greater than the level of involvement required to utilize that part of the brain (Bishop, 2011b).

Equation 1 Calculating a Phantasy

Equation 1 presents the means by which a phantasy is calculated. A phantasy can be defined most easily as a memory that has an effect on how quickly the brain can perform certain tasks. It can have an inhibiting effect, such as if it is a traumatic memory, or an enabling effect, such as if it is the memory of a loved one. A phantasy is made up of two ‘cognitions,’ which are essentially the meanings we have attached to those phantasies, which are both linked together in some way by us.

$$p_i = \left(\frac{((x + x_1)) * (y + y_1) - \bar{z}j}{c} \right)$$

An example that will be used through this section is someone with a social orientation impairment such as autism drawn from (Bishop, 2011b). An SOI who suffered a traumatic assault as an infant might hold a belief that they were harmed by someone (y=3) and then every time someone holding characteristics resembling that person a detachment of 6 is attached to them creating a Flustered phantasy roughly measured in Equation 1.

Using Equation 1 gives a phantasy (p_i) of -5, which on its own creates a Pression (P) of 47, which when the Baseline (B) for hours to receive tax credits is added (48-16) to the maximum recommended force (F) of 48.

Equation 2 Calculating a Pression

The next stage is to compute the ‘Pression’ for each participant using Equation 2. The Pression (P) is the extent to which a number of phantasies inhibit or enhance performance at a particular point in time.

$$P = \left(\left(\sum_{i=1}^n p_i \right) / 5 \right) + F$$

Equation 3 Calculating a Knol

Equation 3 below shows that the final stage is to calculate the knol, which is the level of possible brain productivity for each person in performing a particular task where certain phantasies become activated as a result of the process represented in Equation 2. Essential to this is the baseline (B) figure.

$$k = \frac{P}{(F + B)}$$

The B figure reflects the number of hours a person is working or is comfortable working (e.g. often 16 for a person with a disability) and the F figure the maximum they should be expected to work (i.e. 48 according to working time rules in the European Union). Keeping with the example of the SOI, dividing the Pression calculated in Equation 2 by the value of B and F combined gives a knol (k) of 0.5, which below the recommended of 0.81 and the maximum prefrontal cortex potential. This shows that the person with the SOI is not at the optimum capacity that is safely possible.

AN INVESTIGATION INTO THE EFFECT OF MEDIA CONSUMPTION ON EMFT PROPENSITY: CALCULATING ‘KNOL’ FROM ATTITUDES TO MEDIA CONTENT

To understand how attitude to media content, in this case the arts, have an impact on knol, this section will apply neuroeconomic equations to demonstrate the variation in knol and how it is calculated. The purpose of this study therefore will be to test out these theories relating to knol using empirical data.

Methodology

This project made use of secondary data from an ICM Research study (Couldry, Markham, & Livingstone, 2005) and freedom of information requests into the number of incidents of trolling recorded by the police in the UK between 2009 and 2011. Phase One of the ICM research project comprised of detailed qualitative work across six regions of England. The diaries of 37 participants' media consumption were analysed, initial and subsequent interviews were conducted with those respondents, and focus group interviews were conducted with diarists. Phase Two involved a telephone survey of 1,017 people, conducted by ICM Research across the United Kingdom that aimed to produce conclusions on the detailed issues about consumption and citizenship raised in Phase One.

Preliminary Data Coding and Analysis

The variables in the secondary dataset needed to be re-computed to match those in the equations. One of these was the trolling magnitude scale (Bishop, 2013a; Bishop, 2013b), which is for measuring the gravity of an act of trolling. The computation made the TM values relative to the number of hours spent on the Internet. This was because it was assumed that the longer one is on the Internet the more likely one would be abusive towards others, even if unintentionally (Cassidy, Jackson, & Brown, 2009). On that basis, a trolling magnitude (TM) of 1, which is Playtime, where people are more likely to act in the heat of the moment, was assigned to all the participants who used the Internet 0 to 1 hours. A TM of 2 (i.e. Tactical) was assigned to people who used the Internet for 1 to 3 hours. A TM of 3 (i.e. Strategic) was assigned to people who used the Internet for 3 to 5 hours, and finally, a TM of 4 was assigned to people who used the Internet for between 5 and 7 hours, or greater than 7 hours. This study is important

because it will allow for not only the automated pre-screening of Internet content but by liking knol to trolling magnitude will make it easier to prove injury towards someone by someone else in relation to the offences linked with trolling magnitude (Bishop, 2013a; Bishop, 2013b). In this case it would be when engaging with media in general or art in particular. In Equation 1 the variable F reflects the normal force a human can handle in an average week, measured in hours, which is taken to be 48, based on the European Union's Working Time Directive [2003/88/EC]. In the case of this study, phantasies were computed through the use of a type of cognition called a detachment (rated from 0 to 6) and a plan (rated from -2 to +2). A detachment is an attitude towards a person that makes one anxious and which the mind tries to suppress. A plan is a willingness, or lack of, to perform a particular action.

The detachments were calculated through linking six regions in the UK where data on reports of electronic message faults between 2009 and 2011 were known. This was done by the region with the lowest number of EMF incidents being 0 and the highest number of incidents of flame trolling being 6. This resulted in Yorkshire and the Humber being ranked 0, with 42 incidents and the East of England being ranked 1, with 114 incidents. Wales was ranked 2 with 135 incidents, Scotland 3 with 174 incidents, and South West England 4 with 578 incidents. Northern Ireland was ranked 5 with 1485 incidents and South East England 6 with 1903 incidents.

The three plans used in the study were, 'People like me don't get involved in the arts and culture,' 'There are other things besides arts and culture that I prefer to do in my leisure time,' and 'It's expensive to get involved in arts and culture.' Using the lookup table in Bishop (2011b) for the x_1 and y_1 variables, Equation 1 was used to produce three phantasies from combining the detachment with the associated plan for each participant.

Participants were asked to state whether they worked full-time, part-time, were a homemaker,

registered unemployed, a full-time student, or not working at all. The number of hours worked were harmonized based on official government figures to reflect the 'potential' as opposed to 'actual'. Those working full-time were assigned to 30 hours and those working part-time were assigned to 24 hours. These are the hours required for the respective rates on the UK's Tax Credits system. A homeworker was assigned 35 hours, which is the number of hours required to claim Carers Allowance and a person who was not working or seeking work was assigned 0 hours. Someone registered unemployed and seeking work was assigned 15 hours, which is the maximum one can work and still receive out-of-work benefits, and a full-time student was assigned 21 hours, which is the number of hours one needs to be studying to qualify as a full-time student to receive a reduction in council tax.

The baseline figure, which is a person's individual production possibility in terms of workable hours is then added to the maximum a person can work in an healthy environment, which is 48 hours as signified by F. Adding the baseline and the force together and dividing them by the Pression calculated by Equation 2 then produces a persons speed and performing in a particular setting. In terms of phantasies containing detachments, 0.81 is an ideal figure for someone achieve as they are likely to be able to withstand most abusive content due to media content of a GOIOM nature being suppressed. A knol based on a detachment phantasy score of 0.98 or over is not ideal as this could put a lot of pressure on the prefrontal cortex of the brain which is quite vulnerable. And a knol of below 0.5 is below also not ideal as the phantasy in question is likely to be affecting the person's mental wellbeing and participation in society. Where this reaches around 0.6 then there is an increasing chance the person would not be as productive as they would be at either 0.81 or 0.50.

Results

Using a One-Way ANOVA, the data was then analysed to see whether there was a significant difference between the knol of those assigned to the 4 groups, to make it worthwhile proceeding with further analysis of difference between the groups.

The ANOVA was successful. The degree of freedom numerator for the dataset of 975 participants was 3 and the degrees of freedom denominator was 971. This gave a CV of 2.08380. As the F for the dataset was 1.145 (0.9388) this was not an ideal outcome, but further inspection of the dataset showed it was the best outcome possible. An eight-factor model gave an F of 0.568 and CV of 1.71672 (a difference of 1.14872) and a five-factor model, as currently used by the police gave an F of 0.891 and CV of 1.94486 (a difference of 1.05386). This shows the four-factor model to be most appropriate for determining media ratings based on knol and number of hours spent on the Internet.

Table 1 shows the number of persons assigned to each trolling magnitude, the Mean knol (k) for each group and the upper and lower knol for those groups using the means from the ANOVA. As can be seen, there is a negative relationship between trolling magnitude and knol. That is, as trolling magnitude increases then knol decreases. This can be interpreted as meaning that in order to troll at a higher magnitude that more effort is needed in the brain to do so. It could also mean that those who have been traumatized to the extent their knol is very low, may be at more risk of trolling at a higher magnitude. The upper and lower bounds of knol in Table 1 could be used to develop models, such as based on neural networks, to understand the breaking points for different people in terms of stimulus and response to develop the automated affective content screening discussed in the next section.

Trolling Is Not just a Art. It Is an Science

Table 1. Relationship between trolling magnitude (TM) and knol (k)

TM	N	Mean k	L Bound k	U Bound k	Description
1	90	0.7265	0.6926	0.7604	The prefrontal cortex is operating nearing its most optimal (0.81) means a person is likely to act in the heat of the moment when their knol is at its greatest.
2	285	0.7061	0.6884	0.7239	The prefrontal cortex has decreased from its optimal state by 0.2 points, meaning a decrease in the value of knol increases self-sanctioned trolling activity.
3	519	0.7094	0.6960	0.7228	The prefrontal cortex becomes further sub-optimal as the magnitude of a trolling offence increases, which might explain why the most determined of trolls have social orientation impairments.
4	81	0.6843	0.6527	0.7160	The prefrontal cortex is at its most sub-optimal when the person has to put a lot of effort into their trolling.

AUTOMATED AFFECTIVE CONTENT SCREENING: TOWARDS THE DISTRESS IN THE MIND TEST

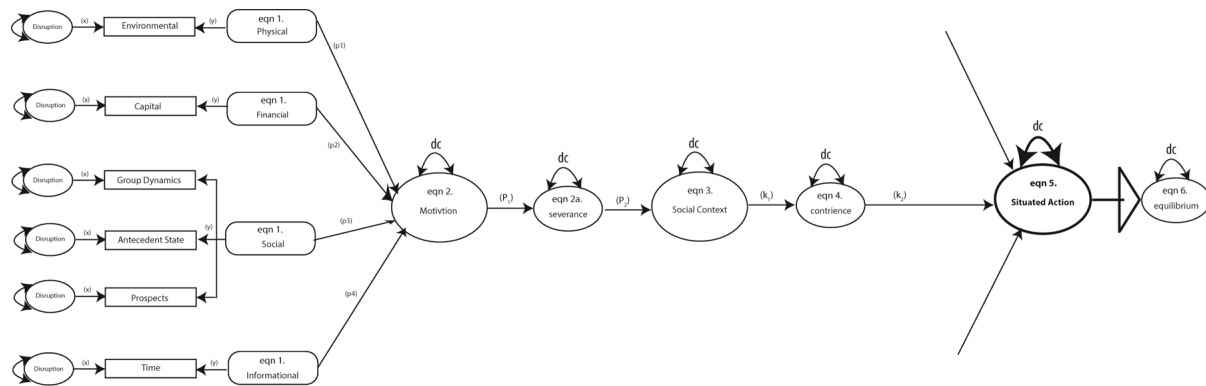
Now that this study has shown that there is a link between length of time on the Internet – exposure to the media – and also someone’s brain processing capacity (i.e. knol) and the trolling magnitude scale, this makes it possible to suggest intervention and methods for assessing someone’s risk of being harmed by certain media content. Using the model from (Bishop & Goode, 2014), which is in Figure 1, the data derived in the earlier section will be used to develop a generalizable framework for determining the risk of content, such as trolling, to people perceiving that stimuli.

In terms of legislation and case law in Great Britain, there is support for varying media content ratings based on the individual who wishes to use them. The case of DPP v Connolly [2008] 1 W.L.R. 276 identified that whether something is of a GOIOM nature is dependent on how much the person being exposed to the media content can withstand it. In this case a number of pharmacists were found to have been ‘grossly offended’ when sent a picture of an aborted foetus. The court said however that if the communication had been sent to an abortion surgeon, who would not be distressed by the images, then it would be no different from

any other political message. On this basis, it is essential ratings systems be changed from being based on age, to being based on something around maturity. Table 2 presents a matrix of various means for assessing media content to determine the maturity needed to consume it. It suggests that maturity is a factor of the strategic thinking required to maximize use of the information in the media content in question without being influenced in such a way that it causes harm to oneself or others. This applies equally to someone who might be extremely vulnerable, such as an infant or a person with an intellectual disability, as well as someone of moderate maturity, who might not find explicit content comforting to watch. Using Table 2, it is possible to easily distinguish between that which could be grossly offensive, indecent, obscene or menacing (GOIOM) at a TM of 1 and that of a TM of 4. Equally it shows which content could cause harassment, alarm or distress, and how this differs with a TM of 2 compared to a TM of 4.

Table 2 incorporates much of the data analysed in the study, including the upper and lower bound knol scores from the study participants shown in Table 1. It can be seen that the higher the average knol the lower the magnitude of trolling. This is because it takes a lot of effort (i.e. involvement) to conduct the higher magnitude forms of trolling, which brings knol down. Table 2 also shows how

Figure 1. A model for measuring the flow of information through the human brain and body



measuring someone’s knol – such as in response to a flame or other stressing stimuli – can make it possible to determine the risk to someone of being exposed to such a stimulus.

AN APPLICATION OF THE DISTRESS IN THE MIND TEST

Consider a child that has been told by their parent that they will see them at their school play. There is nothing explicit in what their parent said. However, if that parent did not turn up at that play the child may be grossly offended. This could be quite traumatic for them if they feel their parent has betrayed them. If the parent keeps missing school plays or similar events where the child wants to make them proud, their phantasy from the original ‘betrayal’ will strengthen and their knol will increase, which will make them not only less able to be grossly offended over ‘minor’ upsets, but also less able to notice the positives in situations where they may be at risk of being upset in a similar way.

The level of strategic thinking ability also relates to Trolling Magnitude. Table 2 shows the lower and upper bounds of knol and how they relate to the maturity of different audiences to consume and otherwise use content without suffering HAD or perceiving it as GOIOM. Someone who is at

the Playtime level of strategic thinking will be like an infant, who could be easily led astray even though they think they are being clever. It also includes people who would get caught up in the moment, such as people who were intellectually challenged, or regularly victimized. The Tactical level of strategic thinking applies to those who are able to understand concepts like intertextuality and be able to be humoured while also understanding when something humorous could be offensive in the wrong context. The Strategic category of strategic thinking relates to content where the person would have to be able to differentiate those practices and behaviours that harm others and should not be adopted from those which others might get gratification from whilst in other situations other people might be discomforted. Finally, the domination form of strategic thinking requires an advanced level of understanding so that it is possible to know what could severely harm others and also cause harm to oneself. Someone of this maturity in strategic thinking should be able to understand murder plots in media, or be able to see graphic scenes of sexual and other violence, without in any way being corrupted by them. People with a low knol, who are made anxious by this content, should know in advance whether they would be at risk. This content is most suited to people who have a high knol in all situations, including violence, so that it in effect ‘goes in one

Trolling Is Not just a Art. It Is an Science

Table 2. Proposed media content classification matrix for determining maturity using strategic thinking level and potential gravity

TM	Rating (Mean k)	Strategy L Bound k (U Bound k)	Potential Gravity (DSS)	Description / Examples
1	Exempt (0.7265)	Playtime 0.6926 (0.7604)	Grossly Offensive, Indecent, Obscene or Menacing (Low Flow, High Involvement)	Apply DPP v Connolly [2008] 1 W.L.R. 276. Some people with a high knol who have not been subject to much trauma might find educational or informational content GOIOM whereas others wouldn't. For instance, pornographic or sexual content may disturb or corrupt the mind of a child, but not a teenager.
2	Universal (0.7061)	Tactical 0.6884 (0.7239)	Harassment, Alarm or Distress (Med Flow, High Involvement)	Apply Chambers v DPP. A person with a lower knol than required for Exempt should not be considered to cause HAD if the content does not make them feel apprehension. For instance, a victim of traumatic abuse should be made to re-live that abuse by being exposed to similar content from a source of information expected to be suitable for vulnerable persons.
3	Parental Advisory (0.6960)	Strategic 0.6960 (0.7228)	Harassment, Alarm or Distress (Med Flow, Low Involvement)	Apply DPP v Collins [2006] 1 WLR 2223. Some people might find something causes HAD if they are the group the content is referring to. Therefore content should be made in such a way an independent third-party could assess whether something offensive, which may not offend everyone, would cause HAD to a minority audience they are responsible for who might have a moderately low knol in some cases, but much lower in others.
4	Explicit (0.6843)	Domination 0.6527 (0.7160)	Grossly Offensive, Indecent, Obscene or Menacing (High Flow, Low Involvement)	Apply DPP v Connolly [2008] 1 W.L.R. 276. A person with a low knol, should be able to expect and tolerate explicit material which they might find GOIOM if they are using a dedicated source containing explicit content which others may find to be GOIOM. One should expect to be grossly offended on a 'hard-core porn' website, but not a 'soft porn' website, for instance. However, a person with one of the lowest knol might be GOIOM if the explicit content relates to the traumatic event that may be lowering their knol.

ear and out the other.' Such people have a clear value system and are unlikely to have suffered any serious abuse. While it is more likely such people will be over the age of 18, where most people will be at less risk of corruption from a high knol, it may not always be the case. Those people at 18 years-old, who may not have had any form of personal, social or health education, or citizenship education, may not be aware of the risks and dangers of the society in which we live. Developing a system of regulating media based on these levels of maturity is essential.

In Bishop (2011a), a system is described that is able to read affective information in a person and recommend a particular course of action to another person. Using the rating system below with this technology, it could be possible for cinemas to be installed with kiosks that assess the risk of viewing

a particular film by a particular individual. This could be done through assessing the emotion in their eyes and face, or through a brain-computer interface, which the person wears while watching those clips from the film that trigger distressing emotions. In terms of Internet trolling, a web-cam or brain-computer device could be used by a parent or guardian to calibrate the Internet security software to only display websites which are suited to the maturity of their child. It could also be used as a permanent Internet of Things device, which can be like an audio-video baby monitor that warns parents or guardians when their children are viewing distressing content. The brain-computer device could even act as a 'chastity band' to ensure that sex predators are unable to access content that could give them inappropriate thoughts or encourage them to perform undesirable acts.

IMPLICATIONS AND FUTURE RESEARCH DIRECTIONS

The data used in this study on Internet trolling was from police records for 2009 to 2011 where they recorded harassment and related offences concerning the Internet. Since 2011 there has been a huge explosion in the reporting of Internet trolling offences, with high profile cases where the public have demanded action be taken and those law enforcement authorities have bowed to that pressure. A future study may be needed to confirm that the findings are still valid using post-2011 data. The study made a link between length of time on the Internet, geography and attitudes towards the arts to derive the figures for determining media ratings and risk to consumers of accessing specific media. Future research may want to perform a regression analysis to confirm the validity of this. The modelling devised in this study need not only apply to the regulating of media content through its analysis, but could actually be used to annotate other environments where distressing material could exist (such as chatrooms and social media) as well as the abstraction of information from computer systems (e.g. for forensic linguistic use) so that it is possible to test the effect of a set of stimuli on a particular person – such as to test whether a suspect recalls a crime they deny committing. This can be done with EEG or MRI scans being analysed to derive the variable values for the equations, such as in terms of flow and involvement.

DISCUSSION

Media content is regulated in different ways in different legal jurisdictions, although in general they tend to follow an out-of-date age-related system, which may mean someone who is at the age of majority is unlikely to know whether a particular film will contain content they would find grossly offensive, indecent, obscene or menacing (GOIOM). Few democratized jurisdictions have

applied the same strict regulatory regime for video and broadcasting as they do for the Internet. The notable exception is Great Britain, where there has been higher reported convictions for the sending of media content that is seen to be GOIOM via a public communications network like the Internet than with regards to commercial communications networks like film, television or video games.

The United Kingdom's media classifications authority – The British Board of Film Classification (BBFC) – has the power to request that any media falling within its remit be cut or censored in other ways in order to meet a certain single rating. The BBFC is a highly respected organisation because it uses objective criteria, unlike organisations like the Advertising Standards Authority which are more reliant on the subjective judgements of officials, even if these in no way reflect reality. It would therefore make sense for the BBFC to cover the regulation of all media content, regardless of frontier. With technology as advanced as it is, it may be possible for film broadcasters, or video on demand websites, to show different versions of the media content to different audiences, depending what is most appropriate for people based on their projected knol in relation to the file or other media content. An organisation like the BBFC would we well placed to devise criteria for what content should be censored by an algorithm and which should not.

Most practicably, using a simple webcam and software kit it could be possible for parents or guardian to be able to assess from using a recommender system that is based on emotion recognition which particular rating for media they believe is appropriate for their child or a vulnerable adult. They could also have more detailed recommendations for specific films. The algorithm could also warn parents when their children are 'getting naked on cam' (GNOC) or taking part in other forms of sexting.

It has long been argued that there is a causal link between media consumption and criminal activity. By presenting a media ratings system which is based on the magnitude of Internet troll-

Trolling Is Not just a Art. It Is an Science

ing offences, this chapter has shown how these can be used to assess whether there are any such links. By linking severity of an Internet trolling offence to the extent to which a person can withstand being subjected to that intensive of GOIOM media content, then it should be easier to show whether someone reacted in a particular way as a result of being exposed to that stimulus in an overpowering way.

ACKNOWLEDGMENT

The author would like to acknowledge all those reviewers who provided comments and suggestions on earlier drafts of this chapter. Special thanks are due to Stephanie Lee and Niren Basu of Swansea University's Institute for Life Science for providing the motivation to prepare this chapter as a basis for a future research collaboration. This research was in part funded by a Economic and Social Research Council research grant (# RES-143-25-0011). The data use in the study was collected by ICM Research (Couldry et al., 2005) and various police forces in the United Kingdom.

REFERENCES

- Baarendse, P. J., Counotte, D. S., O'Donnell, P., & Vanderschuren, L. J. (2013). Early social experience is critical for the development of cognitive control and dopamine modulation of prefrontal cortex function. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 38(8), 1485-1494. doi:10.1038/npp.2013.47
- Beadle, J. N. (2009). *The neuroanatomical basis of empathy: Is empathy impaired following damage to the ventromedial prefrontal cortex?* ProQuest.
- Bernard, T. (1992). *The cycle of juvenile justice*. Oxford, UK: Oxford University Press.
- Bishop, J. (2011a). Assisting human interaction (AU/GB2011/2011266844 ed.). GB: PCT/GB2011/050814.
- Bishop, J. (2011b). *The role of the prefrontal cortex in social orientation construction: A pilot study*. Paper presented to the British Psychological Society's Sustainable Well-being Conference. London, UK.
- Bishop, J. (2012b). Scope and limitations in the government of wales act 2006 for tackling internet abuses in the form of 'Flame trolling'. *Statute Law Review*, 33(2), 207-216. doi:10.1093/slr/hms016
- Bishop, J. (2012c). Taming the chatroom bob: The role of brain-computer interfaces that manipulate prefrontal cortex optimization for increasing participation of victims of traumatic sex and other abuse online. In *Proceedings of the 13th International Conference on Bioinformatics and Computational Biology (BIOCOMP'12)*. BIOCOMP.
- 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(3), 301-318. doi:10.1080/13600869.2013.796706
- Bishop, J. (2013b). The effect of deindividuation of the internet troller on criminal procedure implementation: An interview with a hater. *International Journal of Cyber Criminology*, 7(1), 28-48.
- Bishop, J. (2014a). Digital teens and the 'Antisocial Network': Prevalence of troublesome online youth groups and internet trolling in great Britain. *International Journal of E-Politics*.
- Bishop, J. (2014b). 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), 7-24. doi:10.1504/IJWBC.2014.058384

- Bishop, J., & Goode, M. M. H. (2014). Towards a subjectively devised parametric user model for analysing and influencing behaviour online using gamification: A review and model. In J. Bishop (Ed.), *Gamification for human factors integration: Social, educational, and psychological issues*. Hershey, PA: IGI Global. doi:10.4018/978-1-4666-5071-8.ch005
- Bishop, J., & Mannay, L. (2014). Using the internet to make local music more available to the south wales community. In J. Bishop (Ed.), *Transforming politics and policy in the digital age*. Hershey, PA: IGI Global. doi:10.4018/978-1-4666-6038-0.ch005
- Bishop, J. (2012a). Lessons from the emotivate project for increasing take-up of big society and responsible capitalism initiatives. In P. M. Pumilia-Gnarini, E. Favaron, E. Pacetti, J. Bishop & L. Guerra (Eds.), *Didactic strategies and technologies for education: Incorporating advancements* (pp. 208-217). Hershey, PA: IGI Global.
- Blair, R. J. R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, 11(9), 387–392. doi:10.1016/j.tics.2007.07.003 PMID:17707682
- Cassidy, W., Jackson, M., & Brown, K. N. (2009). Sticks and stones can break my bones, but how can pixels hurt me? students' experiences with cyber-bullying. *School Psychology International*, 30(4), 383–402. doi:10.1177/0143034309106948
- Cavagnero, S. M. (2012). Digital teens: An investigation into the use of the web by adolescents. In P. M. Pumilia-Gnarini, E. Favaron, E. Pacetti, J. Bishop, & L. Guerra (Eds.), *Didactic strategies and technologies for education: Incorporating advancements* (pp. 129–136). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-2122-0.ch012
- Couldry, N., Markham, T., & Livingstone, S. (2005). *Media consumption and the future of public connection*. London, UK: London School of Economics and Political Science.
- Dickey, C., Shenton, M., Voglmaier, M., Niznikiewicz, M., Seidman, L., & Frumin, M. et al. (2001). Prefrontal cortex in schizotypal personality disorder males and females. *Biological Psychiatry*, 49, 336S.
- Dutta-Bergman, M. J. (2006). Community participation and internet use after september 11: Complementarity in channel consumption. *Journal of Computer-Mediated Communication*, 11(2), 469–484. doi:10.1111/j.1083-6101.2006.00022.x
- Gilmore, A. (2012). Counting eyeballs, soundbites and 'plings': Arts participation, strategic instrumentalism and the london 2012 cultural olympiad. *International Journal of Cultural Policy*, 18(2), 151–167. doi:10.1080/10286632.2011.577283
- Halloran, J. D. (1970). Television and delinquency.
- Haravuori, H., Suomalainen, L., Berg, N., Kivirusu, O., & Marttunen, M. (2011). Effects of media exposure on adolescents traumatized in a school shooting. *Journal of Traumatic Stress*, 24(1), 70–77. doi:10.1002/jts.20605 PMID:21268117
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research. Language, Behaviour. Culture (Canadian Ethnology Society)*, 6(2), 215–242.
- 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), 58–86. doi:10.1075/jlac.1.1.04har
- Holland, D. (1972). Television and crime—a causal link. *Auckland UL Rev.*, 2, 53.

Trolling Is Not just a Art. It Is an Science

- Huang, A. H., & Yen, D. C. (2013). Predicting the helpfulness of online Reviews—A replication. *International Journal of Human-Computer Interaction*, 29(2), 129–138. doi:10.1080/10447318.2012.694791
- Jansen, E., & James, V. (1995). *NetLingo: The internet dictionary*. Oxnard, CA: Netlingo Inc.
- Jansen, E., & James, V. (2002). *NetLingo: The internet dictionary*. Oxnard, CA: Netlingo Inc.
- Katz, A. (2012). *Cyberbullying and E-safety: What educators and other professionals need to know*. Jessica Kingsley Publishers.
- Lee, Y. J. (2012). Exploring the economic values of online user ratings with advertising spending on box office sales. In *Proceedings of the 11th International Workshop on E-Business*. Academic Press.
- Levi, R. P. (1979). Violence on television: An old problem with a new picture. *North Carolina Law Review*, 58, 97.
- Maiza, O., Razafimandimby, A., Delamillieure, P., Brazo, P., Beau cousin, V., & Lecardeur, L. et al. (2008). Functional deficit in medial prefrontal cortex: A common neural basis for impaired communication in chronic schizophrenia, first episode of schizophrenia and bipolar disorders? *Schizophrenia Research*, 98, 39–39. doi:10.1016/j.schres.2007.12.083
- Morrow, V. (2001). Young people's explanations and experiences of social exclusion: Retrieving bourdieu's concept of social capital. *International Journal of Sociology and Social Policy*, 21(4/5/6), 37-63.
- Mushiake, H., Sakamoto, K., Saito, N., Inui, T., Aihara, K., & Tanji, J. (2009). Involvement of the prefrontal cortex in problem solving. *International Review of Neurobiology*, 85, 1–11. doi:10.1016/S0074-7742(09)85001-0 PMID:19607957
- Pearce, K. E., & Rice, R. E. (2013). Digital divides from access to activities: Comparing mobile and personal computer internet users. *The Journal of Communication*, 63(4), 721–744. doi:10.1111/jcom.12045
- Petley, J. (1994). In defence of video nasties'. *British Journalism Review*, 5(3), 52–57.
- Phillips, W. (2011, December). LOLing at tragedy: Facebook trolls, memorial pages and resistance to grief online. *First Monday*, 16(12), 11. doi:10.5210/fm.v16i12.3168
- Rimmer, M. (2012). The participation and decision making of 'at risk' youth in community music projects: An exploration of three case studies. *Journal of Youth Studies*, 15(3), 329–350. doi:10.1080/13676261.2011.643232
- Roche, S. W. (2012). *Internet and technology dangers*. Dragonwood.
- Sala, M., Caverzasi, E., Lazzaretti, M., Morandotti, N., De Vidovich, G., & Marraffini, E. et al. (2011). Dorsolateral prefrontal cortex and hippocampus sustain impulsivity and aggressiveness in borderline personality disorder. *Journal of Affective Disorders*, Schroeder, R. (1996). Playspace invaders: Huizinga, baudrillard and video game violence. *Journal of Popular Culture*, 30(3), 143–153.
- Smartt, U. (2011). *Media & entertainment law*. London: Routledge.
- Starmer, K. (2013). *Guidelines on prosecuting cases involving communications sent via social media*. London: Crown Prosecution Service.
- Tanji, J., & Hoshi, E. (2008). Role of the lateral prefrontal cortex in executive behavioral control. *Physiological Reviews*, 88(1), 37. doi:10.1152/physrev.00014.2007 PMID:18195082
- Weijers, I. (1999). The double paradox of juvenile justice. *European Journal on Criminal Policy and Research*, 7(3), 329–351. doi:10.1023/A:1008732820029

Wilcox, A. (2011). Regulating violence in video games: Virtually everything. *J.Nat'L Ass'N Admin.L. Judiciary*, 31, 253.

KEY TERMS AND DEFINITIONS

Electronic Message Fault: A message posted on the Internet or via electronic means that is either unfair or wrongful to a person who is offended by it.

Grossly Offensive: A word to describe an electronic message fault that leaves a person feeling apprehension because the message was targeted at them or a group of which they are part.

Internet Trolling: The posting of provocative or offensive messages on the Internet.

Knol: A knol is a unit of measurement for the ease to which neurological information can pass between the synapses of the brain to access specific functions.

Phantasy: A persistent memory that effects the ability of the brain to access functions located where that memory was formed to protect access to that functioning following an overflow of information to that part of the brain.

Pression: The accumulative pressure put on the brain to process information following a person's mind bringing a lot of variables into play at once.

Trolling Magnitude Scale: The Trolling magnitude Scale (TMS) is a measurement of the severity of Internet trolling in a given situation.