

AstroTurfing, 'CyberTurfing' and other online persuasion campaigns

Mark Leiser^[1]

Cite as Leiser M., "AstroTurfing, 'CyberTurfing' and other online persuasion campaigns", in European Journal of Law and Technology, Vol 7, No 1, 2016.

ABSTRACT

AstroTurfing and its online equivalent CyberTurfing not only pose a significant threat to consumers, but democratic discourse. Regulatory measures in place to prevent deceptive marketing do not grasp the nature of the threat. It is argued that people make decisions about consumer purchases by using heuristics - mental shortcuts and other rules of thumbs. When making decisions, consumers and voters are often relying on incomplete and false information spread as part of an AstroTurfing campaign. Digitally mediated platforms are being manipulated by propagators to help spread false messages in order to advance specific agendas. As a result, consumer trust and democratic discourse are both undermined. It is argued that further regulation is required to combat the deceptive practice.

Keywords: Online Persuasion Campaigns; Regulation; Social media; CyberTurfing; AstroTurfing; Commercial speech; Castells; Sunstein; Heuristics; Unfair Commercial Practices Directive; Rationality

INTRODUCTION

This paper examines the phenomenon known as 'AstroTurfing', and specifically its online equivalent 'CyberTurfing' and other forms of online persuasion. It starts from the position that information technology academics have spent so much time analysing subversion, intrusions, and disruption by both State and business actors in private spheres that there has been a failure to conduct a proper examination of how the same actors use various digitally mediated platforms already in existence to advance their own commercial and political interests. The focus of this article is 'CyberTurfing' - the online equivalent of AstroTurfing - a form of stealth marketing designed to appear to be as a kind of bottom-up activism that in reality is powered by someone else behind the scenes. 'CyberTurfing' is the practice by state actors and commercial entities using digitally mediated platforms to facilitate a commercial benefit or to advance a political objective. The theoretical framework for the paper is provided by two very different narratives of the regulation of cyberspace: Castell's networked individualism and Kahneman and Tversky's theory of heuristics in judgement making. The first part of the paper provides a brief history of AstroTurfing, before introducing its online equivalent and other forms of digitally mediated persuasion. The second part of the paper analyses the regulatory and legal framework surrounding CyberTurfing and other forms of online persuasion campaigns. The next section differentiates between two different types of AstroTurfing: that done for commercial gain and that for the purpose of shaping political influence, concluding that under our current regulatory regime the former is largely regulated, yet suffers from regulatory failings while the latter is, for the most part, not. The paper argues that while commercial CyberTurfing is at least on regulators' radar, CyberTurfing campaigns for political benefit are far more disruptive to democratic discourse. When a campaign is undertaken to advance a political agenda, campaign or ideology, any organisations offering these services do so for a commercial benefit and therefore should be subjected to regulatory oversight. Accordingly, the final section proposes a framework for regulating CyberTurfing and a regulatory solution that allows regulators to protect "politically neutral" and "innocent" platforms.

ASTROTURFING

AstroTurfing is a deceptive practice often deployed by marketers to create the false impression that a campaign has developed organically (the grass-roots in AstroTurf are fake^[2]). 'Classic' AstroTurfing involves the use of paid agents to falsely represent popular sentiment surrounding a product or service. As a result, consumers 'follow the herd'. ^[3] Its deployment in advertising and marketing contexts is subject to regulatory oversight. AstroTurfing campaigns have two distinct characteristics: first, it involves the use of deception to disguise the true origins behind what is really an orchestrated campaign and second, a lack of transparency hides the identity of the campaign's backers. It appears to outsiders as existing only through a spontaneous upwelling of public opinion, reducing effort required for rational thinking. 'Modern' forms of AstroTurfing comes in the form of stealth campaigns designed to gain mass acceptance for a commercial benefit or a political ideology. AstroTurfing campaigns deceive people into thinking that support for a claim or a product/service appears to be organic in nature, and has formed through a natural groundswelling of support of otherwise unconnected parties. Due to the lack of transparency found

in any AstroTurfing campaign, its victims are unaware that support has really been manipulated by an unidentified source. Often it is believed that the source of information often comes from a credible, unaffiliated entity, yet in reality the campaign is designed to withhold the true backers of the strategy. As de Figueiredo states, "Strategic behaviour by interest groups can make it seem that individuals are participating in the rulemaking when organized interests are actually pulling the strings."^[4] As the true identity behind the source of the campaign is withheld, any regulation of AstroTurfing has historically been justified on the basis that there has been a deception of otherwise rational actors. Its practice has been referred to and defined in terms of advertising: it is a false form of grass-roots marketing that makes the illusion of grass-roots support in supporting a consumer product. Others define AstroTurfing as a "guise concerning the use of endorsements and testimonials in advertising" through "material connections".^[5] The Canadian Competition Bureau defines AstroTurfing in an advertising context as referring to "the practice of creating commercial representations that masquerade as the authentic experiences and opinions of impartial consumers, such as fake consumer reviews and testimonials".^[6]

Its origin has been traced back to the early 20th Century by sociologist Caroline Lee who identified one of the first AstroTurfing campaigns was a drive to get people to discard their "tin dipper", a metal cup to dip water from the well bucket, to a cheaper renewable waxed paper cup.^[7] The social movement, spearheaded in a small militant pamphlet called the *Cup-Campaigner*, never disclosed that the pamphlet was printed by Hugh Moore, the co-founder of the Public Cup Vendor Company which would rename itself later to 'Dixie Cup Company'.^[8] Not exactly sounding out of place today, the social message within the *Cup Campaigner* urged citizens to "join the fight" to "banish the most prolific medium for spreading disease".^[9]

Throughout the 1990s AstroTurfing's use as a form of public relations and reputation management grew. After health advocates mobilized against the dangers of tobacco use, the tobacco industry counter-mobilized through the National Smokers Alliance (NSA), a public relations created front group funded by the tobacco industry. Part of Microsoft's strategy for defending itself against an anti-trust case^[10] was funding the Americans for Technology Leadership (ATL). The group orchestrated a fake poll and sent letters sympathetic to Microsoft in order to convince lawmakers that there was public support for a sympathetic ruling.^[11] Therefore, it is unsurprising that campaigners and activists, businesses and politicians would transition from relying on the use of letters and 'robocalls' to digitally mediated platforms where the ease and reach has infinite potential. Not only are we living in an ever connected world, propagators can take advantage of the Internet's speed and lower transaction costs to manipulate discourse, advance their political agenda, or increase commercial awareness of a product. The lack of transparency in the online environment allows propagators to circumvent present regulations preventing 'classic' AstroTurfing.^[12] As the source behind an online campaign is not identifiable, the online version of the phenomenon not only disrupts the information seeking *rational* actor^[13] during commercial transactions, but has the potential for disrupting genuine political discourse, empower 'information entrepreneurs' wanting to spread disinformation^[14], and facilitate digital wildfires and online forms of moral panic.

In addition to the multiple ethical issues it raises, AstroTurfing should also be seen as a significant threat to the legitimacy of genuine grassroots movements.

CYBERTURFING

A successful online CyberTurfing campaign shares some characteristics with its offline equivalent. First, the viral manner in which the deception spreads. Once the seed is planted, the message spreads virally, growing exponentially, while hiding the true origins of the source. After the initial seeding, the message travels by contagion. Even with a minimum amount of seeding, small campaigns can turn into mass movements. [15] People simply don't know they are participating in a highly organised, yet deceptive and manipulative campaign. The second characteristic is the deceptive nature of the campaign. As a result, people who read a CyberTurfing message are not aware of the material connection between the misleading message and the source.

AstroTurfing's online equivalent, CyberTurfing, comes in numerous forms: consider the 'folksy' online blog that appeared to be written by 'Jim and Laura' whereby they reported on their experiences driving their recreational vehicle across America spending the night in Wal-Mart parking lots, and reporting how happy Wal-Mart employees appeared to be. The blog did not mention that Working Families for Wal-Mart, an entity formed by Wal-Mart's public relations firm, paid for the entire enterprise. When exposed by Business Week, the PR firm acknowledged an error. As part of a campaign to boost the biotechnology giant Monsanto's image, a consultancy group fabricated emails and created "The Center for Food and Agricultural Research", a fake institute designated for the purpose of attacking Monsanto's critics. [16] This was one of the early corporate responses to the growing role of the internet in encouraging anti-corporate protests. In addition to Monsanto, Microsoft and Wal-Mart, Sony and Belkin have also been caught AstroTurfing consumers online.

When AstroTurfers were able to take advantage of digitally mediated platforms, CyberTurfers found their natural home. Ratkiewicz et al began tracking bots sending up to 10,000 tweets a month. [17] Some accounts were setup and subsequently identified as 'honey pots' to attract followers. These were relatively easy to identify as fake accounts - they had either pre-programmed times to tweet message or periods of heavy usage. They relied heavily on hashtags to connect to other legitimate users or to enter and become part of ongoing conversations. They often used blacklisted URLs and spam words. An obvious characteristic of a bot or fake account was that they actually weren't very social at all. They had few friends, making them easy to identify. Bots can engage users on social media in real time and are capable of using natural language. They can be simplistic in nature or engage in complex conversations. Users of social media sites like Twitter have had unsolicited contact from simple versions of bots and maybe even actually engaged in conversations with users that they thought were real, but were really advanced bots. Some bots can be programmed with such precision to appear to be a sophisticated interlocutor. Some also recognize and respond to user emotions, including frustration and depression. [18] Others can sense patterns in affective expression and can send prompting stimuli to gauge human reaction. [19] Some are capable of deceiving people through a pre-programmed algorithm that fools people into thinking that the bot is capable of displaying human emotions or has,

at the very least, internal mechanisms analogous to human or animal emotions. Bots can be programmed in a variety of ways, including searching and replying to tweets in response to a series of pre-programmed words and/or phrases. With relative low effort, a bot can be programmed to respond and engage with people punctuating their tweets with a Twitter 'hashtag' which helps organize tweets about similar topics. It can be programmed to seek out influential users in various environments by contacting them directly, with the aim of getting the users to share his commercial message. The bot, usually followed by a small number of other users, has little social gravitas in the online environment. Its aim is to spread the message to other users by taking advantage of their online influence. By targeting people with thousands of followers, it can help to facilitate a marketing and advertising campaign, or start a cascade among other social media users that spreads positive, negative, or disinformation.

The manipulation of a digitally mediated platform in order to advance a specific agenda, product, or political ideology advance an agenda is becoming a frequent occurrence. Although AstroTurfing is subject to regulation, in the online environment, CyberTurfing is much more difficult to detect and with potential for far greater consequences. Any attempts to regulate have ramifications for how we use digitally mediated platforms, commerce and consumer law, and fundamental rights of expression.

WHY CYBERTURFING IS A PROBLEM

Bots programmed to facilitate a CyberTurfing campaign take advantage of features of the online environment where they reside and our reliance to deploy heuristics when making judgments. A common error arising across various digital media platforms comes from our reliance on the repetition heuristic. We deploy this mental shortcut believing that more people report a claim, the greater its credibility. [20] In a series of experiments, a piece of favourable information is repeated across several subjects (strengthening its credibility). This is the common knowledge condition. In the hidden profile condition, the same piece of favourable information is held by only one of the subjects (weakening its credibility). Since repetition across sources bolsters credibility, it is no surprise that repeated pieces of information are more persuasive than unique pieces of information in the experiment: the common knowledge effect is not as shocking as it seems. The repetition heuristic is fallible; a more costly requirement to offer and consider reasons for each bit of information is more likely to yield a better group judgement than intuitive tallying of their frequencies of mention; however, one of the hazards is the risk of the increased use of hidden profiles and the resulting common knowledge effect. Information and repetition entrepreneurs use a range of tactics to take advantage of our reliance of these type of heuristics to create either "norm bandwagons" or "norm cascades". [21] Norm bandwagons occur when small shifts lead to large ones, as people join the "bandwagon"; norm cascades occur when there are rapid shifts in norms. Unsurprisingly, propagators have tapped into certain characteristics of the online environment by developing strategies across digitally mediated platforms that take advantage of our reliance on heuristics.

REGULATION OF COMMERCIAL SPEECH

Any interference of speech undoubtedly courts controversy. In the US, current commercial first amendment jurisprudence is in flux and courts have refused to settle on a set of doctrinal rules permitting the banning speech when its content and context would mislead a certain segment of the population. This is largely in part due to the recognition that courts are not experts in psychology and are unable to "intuit the differences between good and bad law". [22] On the other hand, regulators, who may be experts, are faced with developing "prophylactic rules" that rely on hindsight[23], susceptible to behavioural biases[24] and regulatory capture. [25] In order to combat deceptive advertisements and business practices, reputable agencies self-regulate themselves to avoid reputational damage for misleading and deceptive advertising. In the US, federal legislation permits a business that feels that they have been harmed through the deceptive and misleading practices of one of its competitors can bring an action under Section 43(a) of the Lanham Act 1946.[26]

REGULATION OF DECEPTIVE COMMERCIAL SPEECH IN THE US

The main source of direct regulation for advertising is the Federal Trade Commission (FTC) which can initiate its own investigation about deceptive or misleading practices; issue "cease and desist" orders, and impose fines for serious violations. Advertisements whose messages are false or related to committing crimes are easy enough to spot; however, most issues before the FTC deal with advertisements the commission believes may be deceptive or misleading. The FTC applies a three-part definition in order to determine what is deceptive. First there must be a "material representation, omission, or practice" to be assessed. Material means that the advertisement has made a claim that is likely to affect whether the consumer is likely to buy the product or service. Material statements do not have to be made explicitly; for example, a speaker in a television ad dressed in a white lab coat surrounded by medical charts while promoting a health product was found to be a material representation. [27] The fact that the spokesperson was not a doctor satisfied the second part of the FTC's definition, the material representation, omission, or practice must be "likely to mislead a consumer". "Likely" is assessed on a sliding scale, depending on the FTC's view on how sophisticated the advertisement's intended audience is about advertising and media. Some false statements, like puffery [28] and pure opinion, are not deceptive. For example, a certain brand of golf balls that claimed that it would send your "drives into the stratosphere" is not confusing, as no-one "acting reasonably under the circumstances" would believe the 'outlandish' claims about the flight of a golf ball. The FTC tends to believe that the public is quite savvy about advertising techniques and skeptical of their messages. The fact that some gullible or inattentive consumers fall for the deceptive or misleading claims made does not prove the advertisement was deceptive or misleading. Only if the material statements would be misleading to a reasonable consumer will the advertisement lose its First Amendment protection and be subject to punishment by the FTC.[29] This is an objective reality; in theory, an ad can be deceptive even if the misleading claims fool no one and the ad might be non-deceptive, even though some consumers might be misled.

REGULATION OF DECEPTIVE COMMERCIAL SPEECH IN THE UK

In Europe, deceptive commercial speech is regulated by both the Unfair Commercial Practices Directive (UCPD) and national advertising codes. When corporations deploy this type of deception as a marketing strategy, it is considered a misleading commercial practice under the UCPD. The practice is specifically blacklisted as 'falsely claiming or creating the impression that the trader is not acting for purposes relating to his trade, business, craft or profession, or falsely representing oneself as a consumer' is to be considered unfair in all circumstances. In the UK Regulations 12 and 13 of the Consumer Protection from Unfair Trading Regulations 2008 (enforced by the OFT and local authorities) make this a criminal offence, punishable by a fine of up to £5,000 and/or (on indictment) a prison sentence of up to two years. AstroTurfing is regulated for two reasons: it interferes with the *rational* or *average* consumer and it is deceptive and misleading at its core. It adds more transaction costs to consumers searching to make informed decisions. Those seeking to make informed decisions about the products and services they buy are unwillingly being influenced by organized interests pulling the strings behind the scenes.^[30]

The UCPD only protects consumers' (B2C) economic interests (as opposed to moral interests) and does not concern unfair commercial practices which may damage business' economic interests (B2B), for there is at this time no clear consensus in Europe on the harmonisation of unfair competition rules. ^[31] The Directive provides an extensive list of definitions of the key concepts. The consumer is defined as "any natural person who, in commercial practices covered by this directive, is acting for purposes which are outside his trade, business, craft or profession"^[32], while the trader is defined as being "any natural or legal person who, in commercial practices covered by this directive, is acting for purposes relating to his trade, business, craft or profession and anyone acting in the name of or on behalf of a trader"^[33]. A consumer is limited to "a natural person", while a trader must be either a natural or a legal person. The term "product" is defined in a broad matter as "any goods or service including immovable property, rights and obligations"^[34]. The UCPD's Annex I contains a 'black list' covering all media, including the Internet.^[35] Under No. 22 of the list, "falsely claiming or creating the impression that the trader is not acting for purposes relating to his trade, business, craft or profession, or falsely representing oneself as a consumer" shall constitute a misleading practice. Number 11 bans using editorial content in the media to promote a product where a trader has paid for the promotion without making that clear in the content or by images or sounds clearly identifiable by the consumer (advertorial) is also a banned practice under the Directive.^[36]

Although there are significant differences between the way the US and Europe regulate speech, there are some similarities. At the time of the enactment of the UCPD in Europe, in the United States regulating commercial speech was well established. The Federal Trade Commission's regulations, first published in 1980, were updated to address social media and word-of-mouth marketing^[37] including AstroTurfing^[38] in 2009. The FTC's guidance highlighted the need to "distinguish between the honest word of mouth shared among actual consumers from marketing messages spread by controlled consumer endorsers" ^[39]. The FTC also considered the types of reviews and endorsements commonly left by bloggers and professional reviewers under the guidance. If the author of a review received payment

or in-kind payment such as free products, such endorsement must be disclosed.^[40] A Tennessee company named Legacy Learning Systems, Inc. was fined \$250,000 and given further sanctions for hiring affiliate marketers to write positive reviews on a DVD they had released ^[41]. The affiliates did not disclose they were getting paid for the praise they heaped on the educational DVD series called "Learn and Master Guitar". The FTC subsequently ruled the advertisements deceptive and illegal. In a similar case, the FTC reached a settlement with Reverb Communications, a PR firm that posted false reviews by having employees pose as ordinary consumers posting game reviews in Apple's iTunes store, without disclosing that the reviews came from paid employees working on behalf of the game developers. ^[42] The agency also issued a warning to fashion retailer Ann Taylor after the company gave gifts to bloggers, but closed the investigation without taking action. ^[43] In such cases, the authorities considered that "the bloggers concerned were engaging in hidden commercial practices." The abuse of online reputation management is prevalent in the digital era: Yelp, a well-known online review site, has faced a barrage of lawsuits regarding its business practices. Last year, the Wall Street Journal received a Freedom of Information Act ^[44] response from the FTC outlining the fact they had received 2,046 complaints against Yelp between 2008 and 2014.^[45]

Any violations of the FTC's guidelines for appropriate use of social media may prompt an investigation. The FTC started investigatory proceedings against fashion company Cole Haan in 2014 after the company organised a competition on Pinterest, a 'personalised media platform' that allows media to 'pin' things that they like to their user boards. The company offered a cash prize to the Pinterest user that designed the most creative board using the company's brand of shoes. The FTC ruled that the boards were endorsements of Cole Haan shoes and launched an inquiry into the company's failure to instruct contestants to label their Pinterest "pins" to clarify that the display of Cole Haan products was part of a contest to win a cash prize. The investigation was the first instance where the FTC found that entry into a contest was a "material connection" under Section 5 of the FTC Act. The FTC typically reserves its resources to combat more egregious violators, and a small company's inadvertent or minor infraction is not a likely target. But even a minor slap on the wrist can quickly become a public relations nightmare for a franchise system because the FTC publicizes every settlement in a press release. In addition, if similar violations exist across multiple franchises or if the FTC identifies a pattern of violations, an entire franchise system may be subject to heightened scrutiny or investigation.

While the FTC publishes guides and can bring enforcement action against rogue traders, they are not the only actors capable of raising an action. For the most part, the FTC takes a hands-off approach and tends to let industry regulate itself. ^[46] Fines will be issued for non-compliance, but only after a warning has been issued. The FTC will only proclaim a practice unfair if injury to the consumer cannot reasonably be avoided. ^[47] When the Commission does go after a company, it is usually a larger one; according to Mary Engel (Associate Director for Advertising Practices at the FTC), "certain cases involving smaller local or regional businesses may be more appropriately resolved by a state attorney general's office."^[48] At least two states have taken actions against companies for AstroTurfing and violating essential principles of the FTC's guidelines. State investigations of AstroTurfing, defined in Lifestyle Lift as the practice of "preparing or disseminating a false or deceptive

review that a reasonable consumer would believe to be a neutral, third-party review"^[49] increased following the 2009 settlement between the Florida Attorney General's Office and the cosmetic surgery company Lifestyle Lift. ^[50] AstroTurfing is a violation of multiple state laws prohibiting false advertising and illegal and deceptive business practices. The state of Florida took action against Lifestyle Lift, which at the time of the settlement had over forty locations throughout the United States. The company had combatted negative postings on message boards by having employees create false accounts on online message boards and review sites for the purpose of posting positive reviews and comments about their employer and its services. The Florida Attorney General's Office uncovered emails proving that Lifestyle Lift gave employees guidance about how to pose as satisfied customers online. In addition to posting false reviews, these employees also attacked legitimate critical or negative customer reviews. As part of the settlement, Lifestyle Lift agreed to pay \$300,000 in penalties and costs to the State of Florida and cease posting anonymous, false positive reviews about its business. The Attorney General's Office called Lifestyle Lift's practices cynical, manipulative, and illegal and pledged to continue its mission to protect customers from "emerging fraud and deception, including 'AstroTurfing' on the Internet."

In another example of state action to clamp down on fake reviews, in 2013 the Attorney General's Office of the State of New York undertook an investigation called Operation Clean Turf which resulted in a \$350,000 settlement with nineteen companies and an agreement that they would stop selling fake reviews to businesses for posting on online review sites like Yelp and Citysearch. The offending companies were third-party providers of fake reviews that marketed themselves as search engine optimization companies (SEOs) offering online reputation management services. As part of the investigation, representatives from the Attorney General's Office made prospective phone calls to companies advertising themselves as "Search Engine Optimisation Companies" posing as potential clients looking to bolster their online reputations. They wanted help rebutting negative reviews posted on the Internet. A number of the SEOs immediately offered to write false reviews and post them on sites like Yelp and Citysearch, even publicizing the fraud by advertising on Craig's List for people to "post multiple positive reviews on major review [sic] sites." Owners of a spa advertised: "I need someone who is a Yelp expert to post positive reviews...." A nightclub publicized that it was looking for writers to post reviews "without getting flagged". ^[51] In the end, 19 companies were assigned fines. These ranged from \$2,500 to \$100,000, for a total of \$350,000 in penalties. In addition, all of the firms entered into an assurance of discontinuance. ^[52]

In the UK, the Office of Fair Trading (OFT) encountered similar business practices.^[53] The OFT's accused Handpicked Media Ltd of violating Regulation 3(4) (b) under the provisions of Regulation 6 of the Consumer Protection from Unfair Trading Regulations 2008 after an investigation determined Handpicked Media had engaged a network of bloggers to post on niche websites across various sectors that published "online content which promoted the activities of Handpicked Media's clients, without sufficient disclosures in place to make it clearly identifiable to consumers that the promotions had been paid for."^[54] The OFT took the view that not disclosing who was behind promotional tweets was deceptive and misleading to consumers. In another case, the OFT accepted legal undertakings from

MoreNiche Limited regarding potential violations of Regulation of 3(4) (b) and behaviour that constituted a violation of paragraph 11 and paragraph 22 of Schedule 1 of the CPRs. [55]

The Advertising Standards Agency (ASA) investigated the Mars Chocolate Co Ltd brand Snickers' 'You're not yourself when you're hungry' Twitter campaign after it was referred for misleading consumers under 2.1, 2.3 and 2.4 of the CAP Code. Complainants said that a series of out-of-character celebrity tweets were not readily identifiable as a marketing communications. Katie Price and Rio Ferdinand tweeted references to news events that seemed out of step with their normal tweeting profile. The final of a series of tweets referred to Snickers with the hashtag #Spon. In March 2012 the ASA cleared the campaign as the reveal tweet made clear it was a marketing communication - and was the only one to involve the product. [56] Conversely, In June 2012 the ASA upheld a complaint against the Nike #makeitcount campaign under the same provisions of the CAP Code. Manchester United's Wayne Rooney tweeted "My resolution - to start the year as a champion, and finish it as a champion...#makeitcount gonike.me/makeitcount". Footballer Jack Wilshere tweeted "In 2012, I will come back for my club - and be ready for my country. #makeitcount.gonike.me/makeitcount". The ASA found that the tweets from footballers were under the marketers control and did not make clear that they were marketing communications. By November 2013, the ASA was prompted by the problem of bloggers' misleading consumers through promotional writing enough to publish "*Blurring Advertising and Blogs - Why it pays to know the ad rules*". [57] In response to the Highlighted Media case, the Internet Advertising Bureau and the Incorporated Society of British Advertisers issued guidelines on the payment of editorial content to promote brands within social media. [58]

THE CASE FOR FURTHER REGULATING CYBERTURFING

Despite acting as basis for most policy making in the latter half of the 20th Century, human beings are not *rational* actors. [59] Rather, when presented with complicated decisions, we often deploy a series of mental shortcuts to overcome our laziness and lack of ability to access complete information. Kahneman and Tversky's work on cognitive thinking made the case that "people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgemental operations." [60] These heuristics may be efficient but are sometimes inaccurate procedures for solving a problem. Humans are actually quite poor at decision making, regardless of the information available. They expressed concerns that various constituents have proven to be poor at maximising utility; for example, jurors use of mental heuristics when ascertaining facts during trials and legislators acting on their perceived policy preferences of voters. [61] These mechanisms are usually constructed from primitive mental capacities and developed over a lifetime of experiences. They tell us what information to seek out or select in the environment and how to integrate several sources of information to infer the characteristics of events that are not directly available to perception. [62] When we encounter a situation when we need to use judgement, we select a cognitive tool that is suited to the judgement. We use heuristic strategies because we are generally mentally lazy and in most instances, result in good outcomes. However, although heuristics are generally useful, on occasion their use leads to "severe and systematic" errors. [63]

Bots, looking to bring about a commercial benefit, take advantage of the automation and our reliance on heuristics. For example, a common error that arises from the repetition heuristic is the incorrect assumption that the more people that report a claim, the greater its credibility. [64] In a series of experiments, a piece of favourable information is repeated across several subjects (strengthening its credibility). This is the common knowledge condition. In the hidden profile condition, the same piece of favourable information is held by only one of the subjects (weakening its credibility). Since repetition across sources bolsters credibility, it is no surprise that repeated pieces of information are more persuasive than unique pieces of information in the experiment: the common knowledge effect is not as shocking as it seems. The repetition heuristic is fallible; a more costly requirement to offer and consider reasons for each bit of information is more likely to yield a better group judgement than intuitive tallying of their frequencies of mention; however, one of the hazards is the risk of the increased use of hidden profiles and the resulting common knowledge effect. Information and repetition entrepreneurs use a range of tactics to take advantage of our reliance of these type of heuristics to create either "norm bandwagons" or "norm cascades". [65] Norm bandwagons occur when small shifts lead to large ones, as people join the "bandwagon"; norm cascades occur when there are rapid shifts in norms. Unsurprisingly, propagators have tapped into certain characteristics of the online environment by developing strategies across digitally mediated platforms that take advantage of our reliance on heuristics.

MIT technologist Nicholas Negroponte prophesied that a new user-created and formatted method of communication would develop called the "*Daily Me*". [66] The "*Daily Me*" would allow users some autonomy to filter what they would read every day. While some applauded this as a tool of ease and convenience, Castells us to pause and ask, "How will the increasing control of private powers affect democracy"? [67] Castells queried how emerging technologies affect the growing power of consumers to filter what they see? He sees an essential part of the democratic experience as coming from exposure to ideas that one wouldn't encounter on his own and secondly, citizens should have a range of common experiences that make up part of our shared collective knowledge. One form of CyberTurfing seeks to influence the "*Daily Me*" as proprietors of CyberTurfing platforms seek out influential users in the hope that they will re-share the false story.

The "*Daily Me*", in Castell's view would make it possible to filter the Internet's vast streams of accessible information. As people organized their "*Daily Me*", they would select and categorize information that only appealed to their interests. However, as a result they will likely use these filtering mechanisms to reinforce peoples' prejudices and isolation. This type of group polarization would result in the entrenchment of a certain type of ideals reducing the ability to participate in meaningful deliberative democracy. For Sunstein the best way for deliberative democracy to thrive is through a public forum doctrine and for governments to "take steps to ensure that people are exposed to a diversity of views", including unwanted and unplanned experiences. [68] For Sunstein the public forum doctrine "increases the likelihood that people will generally be exposed to a wide variety of people and views" [69]. He warns about the outcome where the "*Daily Me*" results in reinforcement of narrow interests that cultural balkanization and group polarization. [70] Two independent studies

have shown that members of a deliberating group take a more extreme position in the same direction as their inclinations before deliberation began.^[71]

CyberTurfing does not operate in a social vacuum. A campaign interacts emphatically with social processes.^[72] The first of those processes involves the spread of deception across digitally mediated platforms. The second includes the role of heuristics, and in particular, users' use of mental shortcuts when making commercial decisions. CyberTurfing relies on natural human interdependence and the importance of the collective. Agents of CyberTurfing do this by targeting specific, separate individuals. It is the "*Daily Me*" in the reverse. This abuse of the collective is a manifestation of abuse of the modern telecommunications network. The nature of the Internet's architecture has ensured that functional communities exist without being held back by the geography of real space. The Internet is intimately and inseparably connected to social relations and the communities within, from its physical layer which is specific to each country to political and social uses. This "networked individualism" allows for a new kind of community support. Networked individualism is defined by Castells in two ways: through weak and strong ties. Through either kind of connection, the Internet, according to Castells, forms and strengthens bonds between people who share common experience or interest. At the heart of this is social media which facilitates a certain kinship amongst people and interest groups. They share characteristics such as variant levels of user participation and interaction among various users and interest groups and resolve the barrier that the lack of spatial density presents to efficient real world communications. The ease of communication is a benefit of the cheap and efficient platform that social media provides. Furthermore, the problem presented by group boundaries is removed by the links between multiple networks.^[73] As Wellman et al notes, the "Hierarchies are flatter and more recursive":

"Rather than relating to one group, they cycle through interactions with a variety of others, at work or in the community. Their work and community networks are diffuse, sparsely knit, with vague, overlapping, social and spatial boundaries. The technological development of computer networks and the societal flourishing of social networks are affording the rise of networked individualism in a positive feedback loop. Just as the flexibility of less-bounded, spatially dispersed, social networks creates demand for collaborative communication and information sharing, the rapid development of computer-communications networks nourishes societal transitions from group-based societies to network-based societies."^[74]

Castells notes, "The most important role of the Internet in structuring social relationships is its contribution to the new pattern of sociability based on individualism."^[75] This is a natural outcome of the openness of the Internet's architecture, something Castells describes as the source of its main strength.^[76] The architecture of social media can be said to be deployed in a limited and resource light manner. The 140 character limit of the Twitter platform (with its limited features) form a schematic that allow humans to organize and spread simple and complex and intricate knowledge with limited cognitive thoughts. Returning to Castells, Twitter and other social media platforms are organized into two hierarchies: people follow Twitter accounts or make friends on Facebook because either (1) they know the person or account personally, or (2) because they have an interest in the person or account. This interest might be morbid curiosity, or might be through a genuine

interest in the person's account because of a direct connection with a person considered to be a close tie or personal relationship.

Unlike spam, another form of direct marketing, there is though no suggestion that CyberTurfing is a resource-based or utility-based problem, although there are some estimates that bots now make up to 50 per cent of all internet traffic.^[77] It is a problem of deception. CyberTurfing targets influential Internet users directly with an aim to get the user to re-tweet a tweet of propaganda with the aim of creating a cascade. Sunstein defines an availability cascade as a "self-reinforcing process of collective belief formation by which an expressed perception triggers a chain reaction that gives the perception increasing plausibility through its rising availability in public discourse".^[78] He calls these agents "availability entrepreneurs - activists who manipulate the content of public discourse - strive to trigger availability cascades likely to advance their agendas".^[79] Their availability campaigns may yield social benefits, but sometimes they bring harm, which suggests a need for safeguards. CyberTurfing is interference, or harm, a deceit on the rational actor in the online environment.

What distinguishes CyberTurfing from true commercial dialogue is the manner in which the information spreads. Twitter has recently been considered as case study for information diffusion. Galuba's study takes into account user behaviour, user influence, and resource virulence to predict the spread of URLs through the social network.^[80] Morris's study on infectious diseases on Twitter showed that rumours gradually acquired more credibility as more and more network neighbours acquire them. After some time, a threshold is crossed and the rumour is believed to be true within a community.^[81] For Castells, the network society has become a basic unit of modern society - the network is influenced by political, cultural, and economic variables. Power relationships are the foundations of social organization in all societies and throughout history communication information have been fundamental sources of power and counter power; domination and social change. The fundamental battle is over the mind of the people. The way people think determine the fate of the values, norms, and institutions are founded.

Furthermore, research has shown that social media has surpassed traditional media as a trusted source of information.^[82] This trust is paramount. Many younger people don't look for news anymore as the younger demographics rely on the collective wisdom of the herd. Their 'Daily Me' is determined through connection and relationship. News will be sent to them by their network of friends, not by seeking it out themselves. This is due to the tendency of traditional media to be low engagement, but high reach. Conversely, social media is high engagement and low reach. Social media may have a higher *potential* audience, but the rate of visibility to that audience is incredibly low. To compensate for this, social media platforms incorporate sharing into their design, whereas revenue assumptions of traditional media platforms are based on costs of delivering the service -- but because the cost of distribution technologies is significantly higher, business models have been developed that incorporate an equation that calculates how much it is going to cost to reach each user. This trust and sharing facility is rooted too in heuristics. Twitter's architecture is based in a large part on re-tweeting other people's tweets. The practical effect of this element of Twitter's architecture is that a retweet of a person's tweet is

potentially visible to everyone. On Facebook, the 'liking' system and notification box ensure that 'likes' are visible beyond the initial user publication. Both systems contribute to a type of heuristic called the multiple source effect which occurs when people give more credence to ideas that appear validated by multiple sources. [83] The effects of social influence furthermore can be seen in the tendency of large groups to conform to choices which may be either correct or mistaken, a phenomenon sometimes referred to as herd behaviour. Although social proof reflects a rational motive to take into account the information possessed by others, formal analysis shows that it can cause people to converge too quickly upon a single choice, so that decisions of even large groups of individuals may be grounded in very little information. This helps form an information cascade where a small notional belief within a network can contribute to a cascade based on either a reputational or an informational cascade. [84] What is needed to tackle CyberTurfing is a socio-legal-techno triumvirate of regulation to ensure compliance in the online environment. This approach removes the need for the regulator to enter into the domain of regulating speech. If a CyberTurfer is operating for a commercial benefit, then the messenger should be subject to further regulation, rather than the message.

A REGULATORY SOLUTION

Any regulatory action to tackle the problem of CyberTurfing requires a three-pronged solution implemented through social, legal and technological means. One of the key challenges of social media (and for its supporters) is determining how to assess the quality of the information delivered to users. Making users understand the inherent danger in blindly reading and internalising any type of digitally mediated message is a noble, but unachievable, aim; however, informing users of the consequences of republishing an AstroTurf message is not. Socioeconomics assumes people's preferences and predispositions are partly formed by social norms and can change over time. This stands in contrast to neo-classicists who assume users' preferences are given and fixed. [85] The assumption for neo-classicists is that people are free and rational agents. [86] As Etzioni states, "if the preferences themselves are changeable by social and historical factors and processes that the actor is neither aware of nor controls, the actor's behaviour may be non-rational and not free". [87] In socioeconomics; however, norms are internalized through a process during which an actor will learn to follow rules without threats of external sanctions. [88] Social norms can be seen as regulating, constraining, or reinforcing behaviour. [89] As CyberTurfing is a form of online persuasion campaign, neo-classicists must accept that it falls into one of two categories. For the neo-classicists CyberTurfing is a strictly informational form of advertising and should not affect people's preferences. If the alternative was true, neo-classicists would have to admit that people adopt messages of persuasion by internalising the message. Accordingly, users can no longer be considered rational actors that the neoclassical paradigm assumes them to be. After all, the consequences of manipulation depend upon the CyberTurfer's intent, which may well be to exploit rather than to ameliorate and also upon the effectiveness of the CyberTurf in question. For socio-economists CyberTurfing is a form of informational advertising with a persuasive element, such as outlining an opposition opponent's position on a controversial issue. As opposed to the advertiser who is using information and persuasion in their commercial message, the CyberTurfer is using informational and deception in their message to reinforce *existing* social norms. However,

social norms, and in particular, cyber-based norms take a while to become internalized. Norms, according to Eric Posner "constrain attempts to satisfy their preferences".^[90] Furthermore, norms have social meanings and according to Lessig supply benefits and impose costs to individuals and groups.^[91] Sunstein rationalizes norms as either a subsidy or a tax.^[92]

In some social media environments, information is guarded and an individual's assessment of message is based on the strength of the connections between receiver and sender. However, in other social media environments, proprietary platforms are designed to ensure unfettered connections without any filtering of the messages communicated. There are little limitations on the openness of the platform and the recipient may not have any tools as to gauge its validity or reliability. As already discussed at length above, using heuristics can lead to errors in judgement. AstroTurfing campaigns relies on a series of heuristics that may make people to alter their normal decision making process. For example, someone interested in promoting an AstroTurf message wants to establish a false sense of group consensus about a particular idea.^[93] The consensus heuristic is best summed up in the expression, "if other people believe it then it must be true". This bias is commonly present in a group setting where one thinks the collective opinion of their own group matches that of the larger population. Since the members of a group reach consensus, rarely encountering those who dispute it, they tend to believe that everybody thinks the same way. As an extension, when confronted with evidence that a consensus does not exist, people often assume that the others who do not agree with them are defective in some way. There is no single cause for this cognitive bias; the availability heuristic, coupled with self-serving bias, have been suggested as partial underlying factors. Related to this process is the fact that users are more likely to believe a message that they perceive as coming from several independent sources, or from an acquaintance (the common knowledge effect).

Benkler has argued that some systems of accreditation through commons based peer production system could contribute a possible solution. Users "cooperate to provide information, knowledge or cultural goods without relying on either market pricing or managerial hierarchies to coordinate their common enterprise".^[94] In the Open Source Software community, a programmer's reputation is at stake and a good piece of code will earn its developer certain clout. This has the effect of ensuring that online nicknames are fiercely guarded among the community.^[95] In both of these examples, a form of self-regulation has developed among the community. Legitimacy has come from within and the authority has come to guard against the erosion of confidence in the information that is accessed.^[96] For example, the system of customer rankings on Amazon, Benkler suggests, is a way to solve information degradation^[97].

The second phase would be to tackle the problem of CyberTurfing through technological solutions. As May notes, "one of the key challenges for supporters of openness is to find ways of delivering the quality-related quick and ready assessments that previously were delivered via proprietary modes of assuring the origins of information and knowledge. The bounds or limits of openness are likely to vary with the (social) importance of establishing the reliability of the information and knowledge elements of any specific product, service or knowledge."^[98] Spam detection systems often focus on the content of a potential spam

message-for instance, to see if the message contains a certain link or set of tags. To better detect CyberTurfing, regulators could encourage further development of systems designed to examine how messages are delivered rather than its content; for example, looking for "retweeting anomalies" and mapping tweets that rapidly ascend to become "trending topics". Identifying potential targets is paramount of any technical solution as influencers will often be the target of the attack. Gomez-Rodriguez, Leskovec, and Krause have proposed algorithms that can efficiently approximate linkage information based on the times at which specific URLs appear in a network of news sites. [99]Boutet, Kim, and Yonecki examined 1,150,000 messages from 220,000 users and were able to determine with 86% accuracy their political leanings.[100]. Their retweet graph presented a highly segregated partisan structure with party members more likely to make reference to their own party than other parties. The same analysis can be deployed to determine brand loyalty and the likelihood of making a commercial transaction. By identifying a network of users that may be susceptible to attack, monitors should be able to see CyberTurfing as it happens.

Despite the fact that many of the examples of CyberTurfing discussed in this paper are characterized by small diffusion networks, it is important to note that this is the stage at which such attempts at deception must be identified. CyberTurfing detection systems have been "devised to automatically monitor the data stream from Twitter, detect relevant "memes", collect the tweets that match themes of interest, and produce basic statistical features relative to patterns of diffusion." [101] A potential shortcoming of present technological solutions is once an attempt is successful at gaining the attention of the community, CyberTurfing patterns become "indistinguishable from an organic" pattern. The early identification and termination of accounts associated with CyberTurfing is critical.

Finally, watchdogs could deploy legal strategies regulating entities engaged in political marketing for a commercial benefit by amending election law. Commercial entities that orchestrate campaigns on behalf of individual candidates or political parties (or a party list at a London Assembly election) are regulated under Section 75 of the Representation of the People Act 1983 (RPA). Although spending is controlled for the 365 days ending with the date of poll for UK Parliamentary general elections and for four months preceding an election for the other types of regulated elections, currently, party and issue-based campaigning is regulated only at relevant elections. The law could quickly and easily be expanded to cover currently unregulated periods. Secondly, electoral regulators should mandate transparency about the relationship between the marketing company and their political affiliation and backers. These organisations should file disclosure with the regulator as done in the US with the UK's Electoral Commission. This means that governments may bypass both political opposition and judicial scrutiny of state actions as any plan to require a regulator to curtail freedom of speech would no doubt immediately encounter political opposition. Therefore, a possible solution would be for the electoral commission to engage in co-regulation. [102] In practice, this would involve political marketing groups agreeing to a code of practice - enforceable by the UK Electoral Commission - mandating that the court (empowered by legislation) may order that social media sites close the accounts of those practicing deceptive marketing. The purpose of the regulation would be two-fold. The political marketer would be provided immunity from civil liability for acting (in)appropriately(?) and political parties are removed from discussions on accountability for

the decision to block or suspend accounts (taken by the SNS) or materials used to justify blocking (compiled by the Electoral Commission). In turn, political marketers could forestall legislative mandates that might be more onerous by entering into a timely voluntary arrangement; could present themselves as protecting the democratic integrity of the election process, and could also place the onus for the decisions on the UK Electoral Commission. By imposing a legal duty to disclose this information to regulators, a criminal charge could be made under Section 3 of the Fraud Act, if the entity practicing CyberTurfing in a political context was operating for a commercial benefit.

It is expected that corporations operating within the UK's borders refrain from deceptive behaviour, and this is reflected in the current consumer protection and commercial regulatory framework. The regulatory framework doesn't address organizations disguising commercial speech motivated by economic gains as political speech. Regulating CyberTurfing in this manner circumvents regulating social media. This is because, contrary to Marshall McLuhan's phrase, CyberTurfing is about the medium *and* the message.^[103]

CONCLUSION

The Competition and Market Authority estimated that £23 billion a year of UK consumer spending is potentially influenced by online reviews.^[104] The review sector forms an important part of consumer decision-making processes with 68% of respondents finding online reviews of hotels and travel operators finding 'much more important' or 'a little more important' than other sources of information.^[105] A number of economic papers cited by the CMA report suggests that reviews play a crucial part of consumer decision making.^[106] There is a growing concern at the CMA that some review practices prevent consumers from making the most optimal choice when selecting a product or service. Some of the manipulation of consumer online review systems results in loss of business custom.^[107] Although the manipulation of online review systems has triggered an investigation after consultation by the Competition Market Authority^[108], the terms of the report only cover one form of CyberTurfing. But it is a start of a long and overdue analysis of the way digitally mediated platforms are being manipulated in order to help spread political and commercial propaganda and advertising through deceptive practices. Will the CMA investigation expand itself to examine further forms of exploitative practices? Or will the problem of CyberTurfing remain unaddressed for now, leaving commercial actors to their own devices to manipulate consumers and advance their commercial agendas? The next steps in the review by the CMA present an opportunity to 'fix' some of the problems found in the online review system, but a lack of the recognition of the problem by other national regulators is frustrating and dubious. Without any oversight and regulatory controls, with more knowledge gained everyday about our overreliance on heuristics and mental shortcuts when making judgements, our ability to make *rational* commercial decisions is being compromised. Until action is taken, it may be that the one with the best bot wins.

[1] Mark Leiser is a PhD student at the University of Strathclyde.

[2] The term itself appears to have been coined in 1985 by then Texas Senator Lloyd Bentsen, who noted that the mountains of letters he received about legislation on insurance originated with insurers: "A fellow from Texas can tell the difference between grass roots and AstroTurf... this is generated mail." Alex Wade January 9, 2011. "Good and bad reviews: The ethical debate over 'astroturfing'". The Guardian. 07/09/2013.

[3] See Çelen, B., & Kariv, S. (2004). Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review*, 484-498; Godinho de Matos, M., Ferreira, P., Smith, M. D., & Telang, R. (2016). Culling the Herd: Using Real-World Randomized Experiments to Measure Social Bias with Known Costly Goods. *Management Science*.

[4] John M. de Figueiredo, "Law E-Rulemaking: Bringing Data to Theory at the Federal Communications Commission" *Duke Law Journal*, Vol. 55, No. 5, Thirty-Fifth Annual Administrative Law Issue: The Role of the Internet in Agency Decision making (Mar., 2006), pp. 969-993

[5] FEDERAL TRADE COMMISSION, 16 CFR Part 255, Guides Concerning the Use of Endorsements and Testimonials in Advertising
<http://ftc.gov/os/2009/10/091005revisedendorsementguides.pdf>

[6] The Deceptive Marketing Practices Digest, Canadian Competition Bureau, Accessed 07/08/2015 Available at: [http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/vwapj/cb-digest-deceptive-marketing-e.pdf/\\$FILE/cb-digest-deceptive-marketing-e.pdf](http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/vwapj/cb-digest-deceptive-marketing-e.pdf/$FILE/cb-digest-deceptive-marketing-e.pdf)

[7] "AstroTurfing" *New Scientist* 193.2590 (2007): 48. *Computers & Applied Sciences Complete*, Accessed 10 Nov 2012, Available at: <https://www.newscientist.com/article/mg19325902-300-the-word-astroturfing/>

[8] Caroline W Lee, "The Roots of Astroturfing," (2010).

[9] *Ibid*

[10] *United States v Microsoft Corporation* 253 F.3d 34 (2001)

[11] Joseph Menn and Edmund Sanders, *LA Times*, 23 August 2001, Lobbyists Tied to Microsoft Wrote Citizens' Letters, Available At <http://articles.latimes.com/2001/aug/23/news/mn-37472>, Accessed 30/03/2016

[12] Castells refers to the anonymity provided to users in the online environment as its "main strength". See Manuel Castells, "Communication, Power and Counter-Power in the Network Society," *International journal of communication* 1.1 (2007) at p. 27.

[13] For an extended analysis of the problem with rational actors in the online environment, see See Mark Leiser, "The problem with 'dots': Questioning the role of rationality in the online environment" (ID: 1145952 DOI:10.1080/13600869.2016.1145952) Journal: *International Review of Law, Computers & Technology* Estimated Publication date - 02 May 2016 (Online) , 30 Sep 2016 (Print)

[14] For a general overview in the context of conspiracy theories, see Sunstein, C. R., & Vermeule, A. (2009). Conspiracy theories: Causes and cures*. *Journal of Political Philosophy*, 17(2), 202-227. See also Rojecki, A., & Meraz, S. (2016). Rumors and factitious informational blends: The role of the web in speculative politics. *new media & society*, 18(1), 25-43; Sunstein, C. R., & Hastie, R. (2015). *Wiser: Getting beyond groupthink to make groups smarter*. Harvard Business Press; Kuran, T., & Sunstein, C. R. (1999). Availability cascades and risk regulation. *Stanford Law Review*, 683-768 and Bikchandani, S., Hirschleifer, D., & Welch, I. (1993). *The blind leading the blind: Social influence, fads and information cascades*. UCLA Working Paper, October.

[15] Eric Zuesse, "Final Proof The Tea Party Was Founded As A Bogus AstroTurf Movement", Accessed 23/01/2016, Available at:http://www.huffingtonpost.com/eric-zuesse/final-proof-the-tea-party_b_4136722.html Accessed 30/03/2016

[16] The Bivings Group, a PR firm, built a reputation for manipulating internet discussions in order to promote the interests of its corporate clients. They explain their methods in an article entitled *Viral Marketing: How to Infect the World*: "There are some campaigns where it would be undesirable or even disastrous to let the audience know that your organisation is directly involved... Once you are plugged into this world, it is possible to make postings to these outlets that present your position as an uninvolved third party... Perhaps the greatest advantage of viral marketing is that your message is placed into a context where it is more likely to be considered seriously." G. Goodale and A. Tully, "Who's that selling at your (online) door?", 100 77 *Christian Science Monitor*, 08827729, (17 March 2008); J. Goldberg,, "Selling Wal-Mart", 83 6 *New Yorker*, 0028792X (2 April 2007); A. Krotoski, "New Sony viral marketing ploy angers consumers", *The Guardian* (11 December 2006) [guardian.co.uk http://www.guardian.co.uk/technology/gamesblog/2006/dec/11/newsonyviral](http://www.guardian.co.uk/technology/gamesblog/2006/dec/11/newsonyviral) (last visited January 15, 2010); See letter of apology by the President of Belkin on the company's website, available at <http://www.belkin.com/pressroom/letter.html> (last visited January 15, 2010).

[17] Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Patil, S., Flammini, A., & Menczer, F. (2011, March). Truthy: mapping the spread of astroturf in microblog streams. In *Proceedings of the 20th international conference companion on World wide web* (pp. 249-252). ACM.

[18] J. Klein, Y. Moon and R.W. Picard (2002), "This Computer Responds to User Frustration" (2002) 14 *Interacting with Computers* 119 Available at: <ftp://whitechapel.media.mit.edu/pub/tech-reports/TR-501.pdf> Accessed 18/08/2015 and R. W. Picard, "Toward computers that recognize and respond to human emotion" (2000) 39:3 *IBM Systems Journal* 705, Available at:

<http://www.research.ibm.com/journal/sj/393/part2/picard.html>> Accessed at: 18/08/2015

[19] Picard & Klein, "Computers that Recognize and Respond", *ibid.*; Rosalind W. Picard & Jocelyn Scheirer, "The Galvactivator: A Glove that Senses and Communicates Skin Conductivity" (Paper presented to the 9th International Conference on Human-Computer Interaction, New Orleans, August 2001), <<ftp://whitechapel.media.mit.edu/pub/tech-reports/TR-542.pdf>>; Ashish Kapoor, Yuan Qi & Rosalind W.

Picard, "Fully Automatic Upper Facial Action Recognition" (Paper presented to the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, October 2003), <<ftp://whitechapel.media.mit.edu/pub/tech-reports/TR-571.pdf>>.

[20] Kuran, T., & Sunstein, C. R. (1999). Availability cascades and risk regulation. *Stanford Law Review*, 683-768.

[21] Sunstein, "Social Norms and Social Roles.", Page 909. See also Pp. 71-73 Timur Kuran, Private Truths, Public Lies: The Social Consequences of Preference Falsification (Harvard University Press, 1995).

[22] See Russell N. Laczniak & Sanford Grossbart, An Assessment of Assumptions Underlying the Reasonable Consumer Element in Deceptive Advertising Policy, 9 J. PUB. POL'Y & MKTG. 85, 95 (1990) (suggesting that experts are poor at evaluating puffing speech). See generally W. Kip Viscusi, Jurors, Judges, and the Mistreatment of Risk by the Courts, 30 J. LEG. STUD. 107 (2001) (examining the hindsight bias in how juries and judges analyze risk).

[23] *Ibid.*

[24] Russell N. Laczniak & Sanford Grossbart, An Assessment of Assumptions Underlying the Reasonable Consumer Element in Deceptive Advertising Policy, 9 J. PUB. POL'Y & MKTG. 85, 95 (1990) (suggesting that experts are poor at evaluating puffing speech). See generally W. Kip Viscusi, Jurors, Judges, and the Mistreatment of Risk by the Courts, 30 J. LEG. STUD. 107 (2001) (examining the hindsight bias in how juries and judges analyze risk).

[25] Jon Hanson & David Yosifon, The Situation: An Introduction to the Situational Character, Critical Realism, Power Economics, and Deep Capture, 152 U. PA. L. REV. 129, 202-06 (2005) (reciting the "basic story" of the mechanism of regulatory capture).

[26] Bauer, Joseph P., A Federal Law of Unfair Competition: What Should Be the Reach of Section 43(A) of the Lanham Act? (1984). *UCLA Law Review*, Vol. 31, pp. 671-753, 1984; *Intellectual Property Law Review*, Vol. 17, p. 227, 1986. Available at SSRN: <http://ssrn.com/abstract=1399527>

[27] The classic example is *FTC v. Colgate-Palmolive Co.* 380 U.S. 374 (1965)

[28] Puffery is defined an acceptable amount of lying found in trade advertisements and marketing. See Hoffman, David A. "The best puffery article ever." *Iowa Law Review* 91 (2006): 1395.

[29] For an example of a case that sits in the middle, see *FTC v Oreck Corporation* Available at: <https://www.ftc.gov/news-events/press-releases/2011/04/ftc-settlement-requires-oreck-corporation-stop-making-false>

[30] Another example of political AstroTurfing can be found in the story about a group called the Swift Boat Veterans, a 527 organization. A 527 group is a type of U.S. tax-exempt organization organized under Section 527 of the U.S. Internal Revenue Code (26 U.S.C. § 527). A 527 group is created primarily to influence the selection, nomination, election, appointment or defeat of candidates to federal, state or local public office. There are no upper limits on contributions to 527s and no restrictions on who may contribute. There are no spending limits imposed on these organizations; however, they must register with the IRS, publicly disclose their donors and file periodic reports of contributions and expenditures.

[31] Member States have decided to harmonise certain aspects of unfair competition, while in others, they have discarded this possibility. Directive 2006/114/EC of the European Parliament and of the Council of 12 December 2006 concerning misleading and comparative advertising applies to business-to-consumer but also business-to-business transactions. Likewise, Directive 2005/29/EC does not preclude Member States from extending its application also to businesses through national law.

[32] Article 2 (a) UCPD

[33] Article 2 (b) UCPD.

[34] Article 2 (c) UCPD. For other definitions, see Article 2, especially "business-to-consumer commercial practices", "to materially distort the economic behaviour of consumers", "professional diligence" and "transactional decision".

[35] European Commission, "The unfair commercial practices Directive. questions and answers", 12 December 2007, Available at http://europa.eu/rapid/press-release_MEMO-07-572_en.htm Accessed 01/04/2016

[36] Office of Fair Trading, Guidance on the Consumer Protection from Unfair Trading Regulations 2008, Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/284442/oft1008.pdf, Accessed 01/04/2016

[37] The aptly named Word of Mouth Marketing Association (WOMMA) defines "word of mouth marketing"

"Giving people a reason to talk about your products and services, and making it easier for that conversation to take place. It is the art and science of building active, mutually

beneficial consumer-to-consumer and consumer-to-marketer communications."

<http://womma.org/womm101>

(last visited Oct. 1, 2009)

<http://www.ftc.gov/os/2009/10/091005endorsementguidesfnnotice.pdf>

[38] Discussed further below

[39] Guides Concerning the Use of Endorsements and Testimonials in Advertising 16 CFR Part 255 <http://www.ftc.gov/os/2009/10/091005endorsementguidesfnnotice.pdf>

[40] Ibid, § 255.5 of the Guides

<http://www.ftc.gov/os/2009/10/091005revisedendorsementguides.pdf>

[41] In the Matter of Legacy Learning Systems, Inc., a corporation, & Lester Gabriel Smith, individually & as an officer and director of corporation FTC File No. 102 3055 Available at: [ftc.gov/os/caselist/1023055/110315llscmptexa.pdf](http://www.ftc.gov/os/caselist/1023055/110315llscmptexa.pdf) Accessed 01/04/2016

[42] *FTC vs. Reverb Communications*, FTC File No. 0923199 Available at https://www.ftc.gov/sites/default/files/documents/federal_register_notices/reverb-communications-inc.analysis-proposed-consent-order-aid-public-comment/100901reverb.pdf Accessed 01/04/2016

[43] Metaxas, P. T. (2009). Web spam, social propaganda and the evolution of search engine rankings. In *Web Information Systems and Technologies* (pp. 170-182). Springer Berlin Heidelberg.

[44] Freedom of Information Act, 5 U.S.C.A. § 552

[45] Angus Loten, *Yelp Reviews Brew a Fight over Free Speech vs. Fairness*, Wall Street Journal, 2 April 2014, Available at http://online.wsj.com/news/articles/SB10001424052702303847804579477633444_768964, Accessed 01/04/2016

[46] Haynes, A. W. "Online Privacy Policies: Contracting Away Control over Personal Information?" *Penn State Law Review*. 2007.

[47] Nehf, James P. "Shopping for Privacy Online: Consumer Decision-Making Strategies and the Emerging Market for Information Privacy." *University of Illinois Journal of Law, Technology & Policy*. 2005.

[48] Steinmetz, Katy. "The Audacity of Hype Feds States Fight Fake Etail Reviews Comments." *Time*. 24 Sep. 2013. Available at <http://www.swampland.time.com>, Accessed 07/09/2015

[49] A.G. Schneiderman Announces Agreement With 19 Companies To Stop Writing Fake Online Reviews And Pay More Than \$350,000 In Fines, Available

at: <http://www.ag.ny.gov/press-release/ag-schneiderman-announces-agreement-19-companies-stop-writing-fake-online-reviews-and> Accessed 26/03/2016

[50] See Lifestyle Lift available at http://www.ag.ny.gov/media_center/2009/july/july14b_09.html (last visited January 15, 2010) and Cl. Cain Miller, "Company Settles Case of Reviews It Faked," *The New York Times* (15 July 2009), available at http://www.nytimes.com/2009/07/15/technology/internet/15lift.html?_r=1 (last visited January 15, 2010)

[51] Kyle-Beth Hilfer, "Why Astroturfing and Fake Online Reviews Are Illegal", October 8, 2013 [Accessed 11/08/2015] Available at <http://maximizesocialbusiness.com/astroturfing-fake-online-reviews-illegal-9985/>

[52] An assurance in writing that a person will not engage in the same or in similar conduct in the future.

[53] Under the provisions of the Enterprise and Regulatory Reform Act 2013, the Competition and Markets Authority (CMA) was established on 1 April 2014 combining many of the functions of the OFT and the Competition Commission and superseding both. Department for Business, Innovation and Skills. 8 January 2013. Retrieved 31 January 2016

[54] Case Reference: CRE-E-25932 Accessed: 07/08/2015 Available at: <http://webarchive.nationalarchives.gov.uk/20140402142426/> and http://www.offt.gov.uk/OFTwork/consumer-enforcement/consumer-enforcement-completed/handpicked_media/

[55] Case reference CRE-E/26745 Available at <http://webarchive.nationalarchives.gov.uk/20140402142426/> and <http://www.offt.gov.uk/OFTwork/consumer-enforcement/consumer-enforcement-completed/affiliate-marketing/> Accessed 07/08/2015

[56] ASA Ruling on Mars Chocolate UK Ltd, 097/03/2012 Complaint Ref: A12-185389 Accessed 11/08/2015 Available at: https://www.asa.org.uk/Rulings/Adjudications/2012/3/Mars-Chocolate-UKLtd/SHP_ADJ_185389.aspx#.Vcmx0vntlBd

[57] Blurring advertising and blogs - why it pays to know the ad rules, (13 November 2013) Available at: <https://www.asa.org.uk/News-resources/Media-Centre/2013/Blurring-advertising-and-blogs.aspx#.VcstMPntlBc> Accessed 12/08/2015

[58] Guidelines on the Payment for Editorial Content to Promote Brands within Social Media Q&A Available at: <http://www.iabuk.net/sites/default/files/IAB%20ISBA%20Guidelines%20on%20the%20Payment%20for%20Editorial%20Content%20QA.pdf> Accessed 12/08/2015

[59] See Note 12, *Supra*.

- [60] Sunstein, C. R. (2005). Moral heuristics. *Behavioral and brain sciences*, 28(4), 531-541.
- [61] Hastie, R., Penrod, S., & Pennington, N. (1983). *Inside the jury*. The Lawbook Exchange, Ltd.
- [62] Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.
- [63] Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124-1131, at pp. 1127 and Frank, R. H. (1990). *Rethinking rational choice. Beyond the marketplace: Rethinking economy and society*, 53-87.
- [64] Note 19, *Surpa*
- [65] Sunstein, C. R. (1996). Social norms and social roles. *Columbia law review*, 96(4), 903-968 at page 909. See also Kuran, T. (1997). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press. (Harvard University Press, 1995) at 73-77.
- [66] Negroponte, N. (1996). *Being digital*. Vintage.
- [67] Sunstein, C. (2007). *Fragmentation and Cybercascades*. Republic. com, 51-88.
- [68] Cass R Sunstein, *Republic. Com 2.0* (Princeton University Press, 2007).P.32
- [69] *Ibid* .
- [70] Cass R Sunstein, "Deliberative Trouble? Why Groups Go to Extremes," *The Yale Law Journal* 110.1 (2000). But see Mike Godwin, *Cyber Rights: Defending Free Speech in the Digital Age* (The MIT Press, 2003).
- [71] See Roger Brown, *Social Psychology*, 81, No. 6 (December 2001): 1028 and specifically 223-24. and more generally Cass R Sunstein, *Going to Extremes: How like minds unite and Divide* New York: Oxford University Press, 2009
- [72] Note 19, *Supra*
- [73] Manuel Castells, (2000). *The rise of the network society*, (2nd Ed). Oxford: Blackwell
- [74] Barry Wellman, Anabel Quan-Haase, Jeffrey Boase, Wenhong Chen, Keith Hampton, Isabel Díaz and Kakuko Miyata, "The Social Affordances of the Internet for Networked Individualism," *Journal of Computer-Mediated Communication* 8.3 (2003).
- [75] Manuel Castells, "Communication, Power and Counter-Power in the Network Society," *International journal of communication* 1.1 (2007).
- [76] *Ibid* ., p.27

[77] John Patrick Pullen, "How Bad Bots Are Destroying The Internet", 28 May 2015, Available at <http://time.com/3899612/internet-bots/> Accessed 01/04/2016

[78] Note 19, *Supra*

[79] *Ibid* .

[80] Galuba, W.; Aberer, K.; Chakraborty, D.; Despotovic, Z.; and Kellerer, W. 2010. Out-tweeting the Twitterers - Predicting Information Cascades in Microblogs. In 3rd Workshop on Online Social Networks (WOSN).

[81] Morris, S. 2000. Contagion. *Rev. Economic Studies* 67(1):57-78.

[82] MILLENNIALS AND NEWS, FACT-CHECKED, Available at <http://www.ypulse.com/post/view/millennials-and-news-fact-checked>, Accessed 01/04/2016; See also Flanagin, A.J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviours on the perceived credibility of web-based information. *New Media & Society*, 9, 319-342; Johnson, T. J., & Kaye, B. K. (2004). Wag the blog: How reliance on traditional media and the Internet influence credibility perceptions of weblogs among blog users, *Journalism and Mass Communication Quarterly*, 81, 622-642

[83] Herbert C Kelman, "Compliance, Identification, and Internalization: Three Processes of Attitude Change," *The Journal of conflict resolution* 2.1 (1958).

[84] See David Easley and Jon Kleinberg, *Networks, Crowds, and Markets*, vol. 8 (Cambridge Univ Press, 2010). For further information on notional beliefs, see Noah E Friedkin and Eugene C Johnsen, *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*, vol. 33 (Cambridge University Press, 2011).

[85] George Stigler and Gary Becker (1997) claim that preferences are fixed. "[O]ne does not argue over tastes," they reason, "for the same reason that one does not argue over the Rocky Mountains-both are there, will be there next year, too, and are the same to all men" (Stigler & Becker, "De Gustibus Non Est Disputandum," 67 *American Economic Rev.* 76, 76.

[86] Charles Edward Lindblom, *The Intelligence of Democracy: Decision Making through Mutual Adjustment* (Free Press New York, 1965).

[87] Amitai Etzioni, "Social Norms: Internalization, Persuasion, and History," *Law and Society Review* (2000).

[88] Lawrence Kohlberg, "Stages of Moral Development," *Moral education* (1971).

[89] Lawrence Lessig argues that four constraints regulate human behaviour: "The law, social norms, the market, and architecture-and the "regulation" of this dot is the sum of these four constraints."

[90] Eric A Posner, "Law, Economics, and Inefficient Norms," *U. Pa. L. Rev.* 144 (1995). at 1699.

[91] Lawrence Lessig, "The Regulation of Social Meaning," The University of Chicago Law Review (1995).at 1044

[92] Sunstein, "Social Norms and Social Roles." at 1939.

[93] Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011, July). Detecting and Tracking Political Abuse in Social Media. In ICWSM.

[94] Benkler, Y., & Nissenbaum, H. (2006). Commons-based peer production and virtue*. *Journal of Political Philosophy*, 14(4), 394-419 and Benkler, Y., & Nissenbaum, H. (2006). Commons-based peer production and virtue*. *Journal of Political Philosophy*, 14(4), 394-419..

[95] Ingo, H. (2006). Open life: The philosophy of open source. Lulu. com.

[96] Braman, S. (2006). Tactical memory: The politics of openness in the construction of memory. *First Monday*, 11(7).

[97] Benkler, Y. (2006). The wealth of networks: How social production transforms markets and freedom. Yale University Press at pp76-77.

[98] Christopher May, "Bounded Openness: The Future Political Economy of Knowledge Management," European Intellectual Property Review 33.8 (2011).

[99] Gomez-Rodriguez, M.; Leskovec, J.; and Krause, A. 2010. "Inferring networks of diffusion and influence", In Proc. 16th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD), 1019-1028.

[100] Boutet, A., Kim, H., & Yoneki, E. (2013). What's in Twitter, I know what parties are popular and who you are supporting now!. *Social Network Analysis and Mining*, 3(4), 1379-1391.

[101] Note 1, *Supra*.

[102] Senden, L. A. (2005). Soft law, self-regulation and co-regulation in European law: where do they meet? *Electronic Journal of Comparative Law*, Vol. 9, No. 1, January 2005 Available at SSRN: <http://ssrn.com/abstract=943063>

[103] McLuhan, M., & Fiore, Q. (1967). *The medium is the message*. New York, 123, 126-128.

[104] CMA's call for information report into online reviews and endorsement (June, 2015) at Para 1.3

[105] *Ibid* , at Para 3.15.

[106] Anderson and Magruder (2012), 'Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database', *Economic Journal*; Anderson (2012), *The Impact of Social Media on Lodging Performance*, Cornell University Center for

Hospitality Research; Chevalier and Mayzlin (2006), 'The Effect of Word of Mouth on Sales: Online Book Reviews', *Journal of Marketing Research*.

[107] THE CMA report makes three major empirical claims: consumers find online review systems a valuable part of purchasing decisions; consumers also felt that purchases matched their expectations after reading online reviews, and that to counteract negative reviews, some companies will engage numerous tactics - ranging from engaging customers directly to search engine optimization to online reputation specialists.

[108] Competition and Markets Authority (2015) Online reviews and endorsements report on the CMA's call for information Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/436238/Online_reviews_and_endorsements.pdf, Accessed 30/01/2015