**Online Misinformation and ‘Phantom Patterns’: Epistemic Exploitation in the Era of Big Data**

**Abstract:** In this paper, we examine how the availability of massive quantities of data i.e., the “Big Data” phenomenon, contributes to the creation, spread, and harms of online misinformation. Specifically, we argue that a factor in the problem of online misinformation is the evolved human instinct to recognize *patterns*. While the pattern-recognition instinct is a crucial evolutionary adaptation, we argue that in the age of Big Data, these capacities have, unfortunately, rendered us vulnerable. Given the ways in which online media outlets profit from the spread of misinformation by preying on this pattern-finding instinct, we conceptualize the problem that we identify as a morally objectionable form of “epistemic exploitation.” As we argue, the consumer of digital misinformation is often exploited by having her pattern-recognition instinct used against her. This exploitation is morally objectionable because it deprives her of an epistemic good to which she has a right. This epistemic good is the integrity of the pattern-recognition instinct itself which, we argue, is a capacity that allows us to participate in uniquely human goods. While our primary goal is to bring attention to this form of epistemic exploitation, we conclude by briefly evaluating some general solutions to the growing problem of online misinformation.

**1. Introduction**

Recently, in response to concrete political realities, philosophers have begun to pay much more attention to questions about the nature and effects of “fake news” (Rini 2017; Gelfert 2018), “post-truth” (McIntyre 2018), and “conspiracy theories” (Cassam 2019; Dentith 2019). A considerable portion of this conversation concerns the role that the Internet, and especially social media sites such as Facebook and Twitter, play in the creation and spread of various forms of online misinformation (e.g., O'Connor and Weatherall 2018; Alfano et al. 2020). In this paper, we aim to contribute to this ongoing conversation by highlighting the ways in which the availability of massive quantities of data i.e., the “Big Data” phenomenon, contributes to the creation, spread, and harms of online misinformation. Specifically, we argue that *one* source for the creation and spread of falsehoods that propagate on the Internet is the human instinct to find, recognize, and process patterns. While we review some recent multidisciplinary research suggesting that the pattern-recognition instinct is a crucial evolutionary adaptation, we argue that in the age of Big Data, the pattern-recognition instinct has to a significant degree become a liability. It is plausible that human beings evolved by natural selection to detect salient patterns in their environment; however, as we emphasize, the explosion of information in recent years has drastically increased our propensity to fall prey to what Smith and Cordes (2020) call “phantom patterns”, i.e., misleading patterns. Given the ways in which social media outlets profit from the spread of sensationalistic or outright false content, which prey on this pattern-finding instinct, we conceptualize the problem that we identify as a morally objectionable form of “epistemic exploitation.”

The consumer of digital misinformation in the form of phantom patterns (or conspiracy theories posited to explain phantom patterns) is, we argue, a victim of epistemic exploitation in virtue of meeting two conditions: she is taken advantage of by having her pattern-recognition instinct used against her, and this advantage is taken in an unfair way because it deprives her of an epistemic good.[[1]](#footnote-1) This epistemic good is the pattern-recognition instinct itself which, we argue, acts not only as a tool for survival but is also a capacity that allows us to participate in uniquely human goods such as languages, art, religion, mathematics, and literature. When consumers enter the market for digital data, the flood of (mis)information results in the misfiring of the pattern-recognition instinct, rendering this once-reliable capacity misleading and untrustworthy. Because the digital data market’s exchange of goods results in the loss of this epistemic good for the consumer, we argue that the market exploits consumers, and plausibly results in a violation of what Watson (2018) calls their “epistemic rights”. We conclude by considering some solutions that been proposed to deal with the growing problem of misinformation in the age of Big Data.

**2. The Pattern-Recognition Instinct**

*2.1 What Is a “Pattern”?*

Our ability to effectively navigate our environment in matters both mundane and significant depends heavily on our more general capacity to recognize and make use of *patterns*. Motorists need to recognize and respond successfully to traffic patterns in order to arrive safely to their destination. We all recognize patterns in how foods smell or taste and respond accordingly when those patterns are disrupted, e.g., we throw the food out if it smells odd. We notice patterns in facial expressions and their social or emotional meaning, e.g., a furrowed brow, tense jaw, and flared nostrils is a readily recognizable pattern that suggests anger. In general, the ability to detect patterns of human behavior in the context of interpersonal exchanges is crucial for moving about our complex social world. In mastering a language, we begin to notice grammatical and lexical patterns, e.g., in Spanish words that end in -o tend to be masculine, whereas words that end in -a tend to be feminine. Clearly, given this short list of examples, the term “pattern” is employed in a wide variety of domains to refer to a diverse collection of phenomena. It is probably impossible to spell out some informative set of individually necessary and jointly sufficient conditions for what counts as a “pattern.” Along with Resnick (1994, p. 202), we also “know of no developed philosophical account of patterns.” Still, there are a number of commonalities between many instances of patterns that are worth highlighting.

 Most important, patterns tend to be phenomena that are characterized by some non-trivial element of *repetition*. We would not describe some singular, unique, non-recurring event, such as the sinking of the RMS Titanic in 1912 as a pattern. However, if between the years of 1910 and 1920, twenty-five large, luxury ocean-liners all sank in the North Atlantic after being struck by icebergs, then we would very likely describe this series of events as a *pattern*. When the cat meows in front of the door, waits for the door to be opened by its owner, and then goes outside every morning, it is the repeated nature of this sequence of events that makes it a pattern. Not all patterns need to consist in repeated sequences of some event-type, however. An assembly of objects at a single moment in time can also constitute a pattern. We might recognize and admire the geometric patterns of ancient Mayan art, or the arithmetical pattern in a sequence of numbers, e.g., <2, 4, 8, 16, 32, 64, 128, 256…>. Here, there is a repetition, but the repetition is *synchronic,* i.e., at the same time, rather than *diachronic*, i.e., across distinct times. Some patterns are perhaps better characterized by *resemblance relations*, rather than repetition. For example, one might consider it a pattern if a child was born on July 11th, at 7:11 AM, weighing 7 pounds and 11 ounces. The date of birth, time of birth, and birth weight are all *similar* in some respect, featuring the numbers “7” and “11”. Sometimes patterns are simply said to consist in “one or more objects...that stand in various relationships” (Resnick 1994, p. 203), where the pattern arises holistically on the basis of the objects that occupy various “positions.” For example, the style of art known as “Pointillism” uses patterns of tiny dots of varying colors to create large impressionistic paintings (e.g., *A Sunday Afternoon on the Island of La Grande Jatte* by Georges Seurat).

 Many of the examples of patterns introduced above (but not all) can be usefully analyzed as universally quantified conditionals with *ceteris* *paribus* clauses. For example, the pattern that we notice when driving a car, whereby cars that put their left-blinker on tend to turn left, can be formalized as: “Other things being equal, if a car has its left-blinker on, then that car will soon turn left.” Of course, there is quite a bit of controversy about how to understand *ceteris* *paribus* clauses in a way that avoids rendering such statements vacuous (Reutlinger et al. 2019). Notwithstanding such concerns, given some ceteris paribus conditional statement, we often have a good idea of which sorts of factors make it so that conditions are not equal, based on background knowledge, the contextual details of the situation, etc. For instance, we know that the above conditional will not hold in the next instance if we just saw the driver turn left but then fail to turn their left blinker off. Some instances of patterns are less usefully described as universally quantified conditionals. The repeated use of some geometric form, e.g., equilateral triangles in a decorative quilt, constitutes a pattern, but it would be awkwardly described as a universally quantified conditional. Likewise, patterns that report salient synchronic similarities may not correspond to some conditional statement that relates distinct event-types.

*2.2 The Evolution of the Pattern-Recognition Instinct*

 Even if it is difficult to provide a definitive conceptual analysis of what a pattern is, it is nonetheless clear that pattern-recognition and pattern-processing are crucial cognitive skills. In an article from *Frontiers of Neuroscience*, which synthesizes empirical research from a wide variety of disciplines, including neuroscience, psychology, and human evolution, Mattson (2014) makes a compelling case that “superior pattern processing is the essence of the evolved human brain” (p. 1). Now, of course, pattern-processing is not entirely unique to human beings. Non-human animals engage in pattern-processing all the time, which includes abilities to: i) “construct cognitive maps of the physical environment”, ii) “distinguish individuals of the same species, and their emotional state”, and iii) use “gestures to capture the attention of, and to communicate a desired response from, other individuals” (ibid.). However, when humans evolved from their earlier hominid ancestors, their cerebral cortex greatly increased in size and complexity, and as a result “novel pattern processing capabilities emerged” (ibid.). The sorts of pattern-processing abilities that characterize the evolution of human intelligence include: i) creativity and invention, which led to the construction of tools and technology, ii) spoken and written languages, which enabled humans to communicate detailed information about the environment to each other, iii) reasoning and rapid decision-making, iv) imagination and mental time travel, which permits the articulation and exploration of hypothetical scenarios, etc. Many of these abilities are clearly connected with and likely enhanced by the fact that “humans evolved as highly social animals” (Mattson 2014, p. 7). Crucially though, “an evolutionarily rapid expansion of pattern processing capabilities *is the major reason that the human brain has capabilities considerably beyond those of lower species*” (Mattson 2014, p. 2; emphasis added). In addition to the pattern-processing abilities that non-human animals possess, the superior pattern-processing of humans, which allows for the more advanced abilities mentioned above, is very likely a crucial evolutionary adaptation. In a very real sense then, we human beings are “hard-wired” to notice and process patterns.

 Even though we likely evolved to discern and make use of patterns in the world, it must be emphasized that the pattern-recognition instinct is not perfect, nor should we expect it to be. As Pigliucci (2010) points out, “contrary to popular belief, natural selection is not an optimizing process”; often natural selection “makes mistakes and is inefficient, yielding whatever outcome is good enough for survival and reproduction.” Rather than a process that tends toward optimization, natural selection should be regarded as a “satisficing process” (Simon 1979). An ability to detect poisonous foods that yields some false positives is more likely to be selected for than one that is perfectly calibrated. Likewise, while perhaps humans would have higher levels of *fitness*, were they to have perfectly reliable pattern-processing abilities, the mere fact that some trait would be adaptive does not mean that some population of organisms will tend toward that trait. The historical contingency of some prior adaptation can act as a constraint on what is biologically possible in the present for some species, leading to what Dennett calls “*forced moves in Design Space*” (1996, p. 129). Fitelson and Sober (1998, p. 120) provide a vivid illustration of this point:

The reason zebras don’t have machine guns with which to repel lion attacks is not that firing machine guns would have been less adaptive than running away; the trait didn’t evolve because it was not available as a variation on which selection could act ancestrally.

Among three possible solutions, *A*, *B*, and *C* to some given design problem, it might be the case that *A* would be most adaptive for the species, but because of some prior adaptation that is now fixed and irreversible, only the less adaptive options, *B* and *C* are available.

 In addition, it is worth mentioning the elementary point that a trait selected for in some ancestral environment may become maladaptive if environmental conditions change. In particular, although there is strong evidence to suggest that superior pattern-processing underlies the evolution of human intelligence and the concomitant adaptive success of the human species, it is worth considering the possibility that environmental conditions are now such that the pattern-recognition instinct has, in significant ways, become a liability. One way in which the pattern-recognition instinct could become a severe liability is if the environment were to become polluted with many instances of misleading patterns, which our evolved pattern-processing mechanisms mistakenly identify as genuine.

**3. Online Misinformation and Phantom Patterns**

*3.1 The Phantom Pattern Problem*

 In their recent book, *The Phantom Pattern Problem: The Mirage of Big Data* (2020), economist Gary Smith and statistician Jay Cordes provide some compelling reasons to be alarmed by the prospect of our pattern-processing abilities misfiring. The last few decades have witnessed the arrival of the so-called “Big Data” movement, which can be characterized by the rapid growth and development of new technologies that permit the collection, integration, and analysis of massive quantities of data. The character of this data-collection is unprecedented, not only its volume and variety, but also in the velocity at which such data is being compiled (Alfano and Klein 2019, p. 2).

 For some Big Data enthusiasts, this new era of data-processing promises to bring about lasting changes to the way that we “live, work, and think” (Mayer-Schönberger and Cukier 2013). For example, with the emergence of advanced computational tools, it has been suggested that we will soon see profound changes to the way that science is done. According to some commentators, we will soon see the emergence of a new “paradigm” (Hey et al. 2009) of “data-intensive” science, which is said to differ markedly from traditional conceptions of scientific methodology. The new data-intensive paradigm is intended to be largely “data-driven” rather than “hypothesis-driven” (Mazzocchi 2015), depend less on theoretical background assumptions (Kitchin 2014), and involve “a move away from the age-old search for causality”, where instead the focus will be on correlations, which in many contexts are “good enough” (Mayer-Schönberger and Cukier 2013, p. 14). Perhaps as the new paradigm of data-intensive science continues to evolve, scientists will eventually be replaced by algorithms, at last allowing the data “to speak for themselves.”

 It is against this backdrop that Smith and Cordes (henceforth, “S&C”) raise the “phantom pattern problem”, a serious epistemic predicament that undermines the provocative prognostications of Big Data enthusiasts. With our newfound abilities to gather, process, and analyze more data, we now have the corresponding ability to uncover more patterns than ever before. On its surface, being able to uncover more patterns sounds like an incontrovertible good. As discussed in the previous section, it is through our pattern-recognition abilities that we can predict, understand, and manipulate the world to satisfy our practical goals. However, not all patterns are genuine. Some patterns are what S&C refer to as “phantom patterns”, that is, “coincidental correlations [that] are useless for making predictions” (2020, p. 45). According to S&C, the primary reason that phantom patterns have little to no predictive value is that in order for a pattern to be meaningful, it must have “an underlying causal explanation” (p. 23). But correlations such as that which existed between “avocado prices in San Francisco in 2015 and Google searches for the Virgo zodiac sign” clearly lack any plausible causal explanation, and so such a correlation is best regarded as “a fleeting coincidence that is useless for making predictions” (p. 25).

 In their book, S&C argue, by recourse to a multitude of examples from sports, finance, gambling, astronomy, medicine, etc. that phantom patterns are pervasive in the age of Big Data. As many of the examples in the book illustrate, even expert data scientists (p. 26), financial consultants (p. 35), medical researchers (p. 58), astronomers (p. 68), and economists (p. 145) can find themselves bewitched by phantom patterns. For instance, consider the case of the Urban Institute, “a highly regarded Washington think-tank”, which in 2006 proposed that America was experiencing an “iCrime Wave” based on an observed correlation between an increase in reported murders and an increase iPod sales between 2004 and 2006 (p. 1). As S&C point out, this correlation (i) was based on only three years of data, (ii) was no stronger than the correlation between ice cream sales and violent crime in the same period, and (iii) vanished shortly thereafter as the murder rate dropped in 2007. While the correlation looked to be meaningful, and one that the Urban Institute presumably wanted to explain by positing that violent crime was being committed in order to acquire iPods, it turned out that this was a *phantom* pattern.

*3.2 The Paradox of Big Data*

 Consequently, what we have begun to witness in the last few years is a phenomenon to which S&C refer as the “paradox of big data” (2020, p. 104): meaningless patterns invariably arise in large data sets, and since the number of meaningful patterns is likely fixed, as we gather more data, the probability that any randomly selected pattern is meaningful radically decreases. Thus, from the Bayesian point of view, the prior probability that any randomly selected correlation uncovered through data-mining is meaningful should be regarded as extraordinarily low (Smith and Cordes 2020, pp. 212-13). Mere correlation in the absence of causal knowledge is, as it turns out, not “good enough.” Similar conclusions have been arrived at by Calude and Longo (2017), who have demonstrated mathematically that “spurious correlations largely prevail, that is, their measure tends to one with the size of the database” (p. 602). What seemed to be an undeniable epistemic good, namely access to more information on the basis of which one could potentially discern more correlations in the data, turns out in the end to be an impediment to uncovering *meaningful* patterns. As Calude and Longo astutely remark, “paradoxically, the more information we have, the more difficult is to extract meaning from it. Too much information tends to behave like very little information” (2017, p. 600). Plausibly, the fact that phantom patterns increase exponentially given the increase in size of the data set causally contributes to the staggering fact that over 85% of Big Data projects carried out by businesses fail (Smith and Cordes 2020, p. 102).

 Of course, most people do not work in Big Data analytics, and so one might think that the problem that we have highlighted here, while important for the epistemology of data science, has less relevance to those concerned about the spread of online misinformation. However, we should resist the temptation to think that the paradox of Big Data is an isolated problem. First, the phantom patterns that are uncovered by data analysts are often reported in the popular press. For example, one of the phantom patterns examined by S&C, namely the correlation between Bitcoin prices and Google searches for Bitcoin, which ultimately turned out to be spurious (2020, pp. 127-30), was enthusiastically discussed in outlets such as *Business Insider*, *MarketWatch*, *Investopedia*, and *Yale News*. Thus, it is quite easy for the phantom patterns uncovered by Big Data analysts to get uptake among those who know very little about how sophisticated data science methods work. Second, a more informal version of the paradox of Big Data manifests itself in our daily interactions online, owing to the vast amount of information that is now available on the Internet. According to Dave Reinsel, senior vice president, of the International Data Corporation, a prominent market research group, the amount of digital data created or replicated in 2020 was on the order of 64 trillion gigabytes (or 64 zettabytes); moreover, the amount of digital data that will be created during the next five years will be “greater than twice the amount of data created since the advent of digital storage.”[[2]](#footnote-2) The ordinary person’s ability to access massive quantities of data online is truly unprecedented in the history of the human species, but this unprecedented access to data can very easily lead users’ of the Internet to fall prey to phantom patterns. Just as dramatically increasing the size of the data set will dramatically increase the number of spurious correlations, so too, the more general increase in the amount of information available to people in the Internet age dramatically increases one’s chances of falling prey to phantom patterns. There is thus an important, but neglected, connection between the Big Data phenomenon and the recent upsurge in online misinformation.

*3.3 Phantom Patterns and Conspiracy Theories*

 Consider some recent examples of noteworthy incidents of online misinformation involving phantom patterns. In the summer of 2018, congresswoman Marjorie Taylor Greene wrote a lengthy Facebook post in which she speculated that members of the Rothschild family intentionally started the California summer wildfires. The evidence provided was a short list of coincidental distant relations between personnel at Rothschild Inc. and various other California-based corporations. Greene wrote, “I’m posting this in speculation because there are too many coincidences to ignore, and just putting it out there from some research I’ve done stemming from my curiosity.” Even though her data set was small, and the correlations easily explained in other ways, Greene thought she saw a pattern best explained by positing a conspiracy involving the Rothschilds. Clearly though, the data that Greene sought to explain was a phantom patten, one which likely required no explanation at all.

 Now, to be sure, high-profile purveyors of conspiracy theories frequently have political or financial motives, which almost certainly play a major role in public endorsements of conspiracy theories. In the final analysis, these other causal factors should not be ignored. Perhaps these factors were the primary impetus behind Greene’s Facebook post. As some commentators pointed out[[3]](#footnote-3), positing that the Rothschilds were conspiring toward wicked ends trades on age-old anti-Semitic tropes, often involving a “belief in Jewish conspiracy”, which have recently been linked to feelings of a lack of control in the political realm (Kofta et al. 2020). Nevertheless, the case involving Greene and the California wildfires is instructive for our purposes. Often, part of the explanation for belief in unjustified conspiracy theories is that seemingly unrelated pieces of information are being strung together into some phantom pattern, either by ordinary individuals themselves or those with a large public platform, which is then explained by positing some conspiracy theory, usually involving shadowy actors with nefarious purposes. The increasingly popular QAnon phenomenon, which refers to “the bizarre assemblage of far-right conspiracy theories that holds that [former] U.S. President Donald Trump is waging a secret war against an international cabal of satanic pedophiles” (Amarasingam and Argentino (2020, p. 37) involves the detection and attempt to explain phantom patterns in precisely this way. Followers of the QAnon conspiracy theory are often led to outlandish and unjustified beliefs—e.g., the claim that Hillary Clinton was running a child-sex ring in the basement of a pizza shop in Washington D.C.—largely by misperceiving a pattern in a set of banal and unrelated pieces of data—“‘drops’ of information from an anonymous insider calling themselves Q” (Aliapoulios et al. 2021, p. 1). Without the unprecedented access to information afforded by the Internet, outlandish conspiracy theories such as the QAnon theory would likely not be formulated, let alone be believed by a concerningly large number of people.

Of course, not all online misinformation consists of phantom patterns that are then explained by some unsubstantiated conspiracy theory. Sometimes misinformation simply consists in a simple, factually incorrect statement, e.g., “Tony Hawk died in 2016.” Even so, misinformation in the form of phantom patterns or unsubstantiated conspiracy theories designed to explain phantom patterns is a common enough phenomenon to be worth focusing on.

 While the above examples are rather extreme, we should not make the mistake of thinking that phantom patterns are something that only those antecedently susceptible to far-right conspiracy theories must worry about. For example, Naomi Wolf, a popular feminist author and former adviser to Bill Clinton, was recently suspended from Twitter for promoting outlandish conspiracy theories about COVID-19 vaccines in posts that were often riddled with cases of phantom patterns.[[4]](#footnote-4) Furthermore, as many theorists have already discussed, the spread of online misinformation is intimately connected to the dynamics by which websites and apps, in particular social media, display content to their users. In a recent analysis of fake news as a noxious market, Fritts & Cabrera (forthcoming) argue that there is a complex market for fake news in which consumers, content creators, site owners, and political figures all benefit from the dissemination of false or misleading headlines. This market financially incentivizes every step of the consumption of fake news—content creators profit by selling their articles, site owners profit by selling advertising space, politicians benefit by allying with various sources of online misinformation, consumers benefit by having access to media they find interesting, and online misinformation sites benefit from their allegiances with various political figures (Fritts and Cabrera forthcoming, p. 7). Similarly, Benkler et al. (2018, pp. 77-8) discuss this same sort of phenomenon in the context of both social media platforms and traditional media outlets, such as Fox News, which they refer to as the “feedback loop” of “outrage media” and other kinds of misinformation. Even if users are displayed content that is not strictly false, the *type* of content and the *frequency* with which it is displayed can lead to a kind of “technological seduction” (Alfano et al. 2018), paving the way for a greater vulnerability to phantom patterns.

*3.4 A Species of Information Overload*

 The phantom pattern problem can be folded into the more general ethical concern about the rise of information technology, namely the problem of “information overload.” According to Himma (2007), who provides lucid conceptual analysis of the notion, we can think of information overload in the simplest terms as “hav[ing] access to more information than is conducive to human wellbeing” (p. 266).[[5]](#footnote-5) The concept of “information overload” is a negatively valenced instance of what Bernard Williams (1985) would call “a thick ethical concept.” To be overloaded with something is, necessarily, undesirable relative to some evaluative standard. Thus, the concept of “information overload” militates against the admittedly natural inclination to regard information as an unconditional good, either because information necessarily helps us achieve our practical/moral goals, or because information is intrinsically valuable. However, there are many interconnected ways in which we can be negatively impacted by virtue of having too much information (Himma 2007, pp. 267-72). Too much information can i) deplete our attention, a limited resource needed for other important tasks, ii) make it more difficult to sort the useless from the useful information, iii) engender the intrinsically unpleasant psychological state sometimes called “technostress” (Brod 1984), and iv) provoke the feeling of being overwhelmed, which can undermine our ability to make choices (Schwartz 2004).

 To the list of problems with information overload that Himma identifies, we would add “increased vulnerability to phantom patterns.” Not only can falling prey to phantom pattern negatively impact well-being in general (by, e.g., fostering reckless actions with respect to the COVID-19 pandemic), our greater susceptibility to phantom patterns also negatively impacts what we might call our “epistemic well being” (Hookway 2003, p. 79). That is, the growth of phantom patterns makes it more difficult for agents to successfully gain meaningful explanatory and predictive knowledge. Being overloaded with information makes us more liable to mistakenly regard some coincidental correlation as a good predictor of future events, or to mistakenly accept some causal explanation for a set of events, which considered jointly, has no causal explanation. Consequently, the phantom pattern problem, along with the paradox of Big Data, is an especially insidious manifestation of the problem of information overload, as it poses a threat to *both* our epistemic and our practical agency.

**4. The Wrongful Exploitation of the Pattern-Finding Instinct**

*4.1. What Is “Exploitation”?*

In the Big Data era, the myriad ways in which targeted advertising leads people to consume misinformation and makes them more susceptible to believing phantom patterns can be regarded as a kind of *exploitation*, in the moral sense of the term. Sometimes we use the term “to exploit” in a way that is synonymous with morally neutral expressions, such as “to make use of” or “to avail oneself of”. For example, one might *exploit* their own natural artistic talent to complete a beautiful painting, where this sort of “exploitation” does not necessarily imply a wrongdoing. In this non-moral sense then, it is non-controversial that predictive algorithms *exploit* the pattern-recognition instinct in order to foster online interactions, typically in the form of clicks, with dubious, sensationalistic, or outright false content. However, what we argue is that in virtue of exploiting the pattern-recognition instinct in this way, the *individuals* that interact with this content are exploitedin the *moral sense.[[6]](#footnote-6)*

Here, we adopt relatively standard definition of exploitation put forward by Zwolinksi and Wertheimer (2016), according to which: “To exploit someone is to take unfair advantage of them. It is to use another person’s vulnerability for one’s own benefit.” In a discussion of the ethics of the living wage, Brennan (2019, p. 19) provides a helpful “archetypical case” of exploitation, which is adapted from Nozick (1969, p. 447):

*Drowning Man*: Peter’s boat capsizes in the ocean. He will soon drown. Ed comes along in a boat. He says to Peter, “I’ll save you from drowning, but only if you provide me with 50% of your future earnings.” Peter angrily agrees.

This case is naturally described as an instance exploitation: Ed wrongs Peter by exploiting him. Peter happens to find himself in a vulnerable situation, and Ed takes advantage of Peter’s vulnerability in order to derive some (presumably unnecessary) monetary benefit.

 What is most important about the Drowning Man case is that it illustrates a number of key points about what *is* and *is not* required for *A* to exploit *B*. First, it is clear that even though Ed exploits Peter in offering to save Peter for half of his income, Ed does not harm Peter in the sense of making Peter *worse off*. If Ed does not rescue Peter, then Peter will drown, which is presumably a worse state of affairs for Peter than having to give up half of his income. In fact, rather than being made worse off, it seems that Peter is made *better off* by the deal. Peter can very easily reject Ed’s offer, but the fact that Peter accepts the deal is strong, though defeasible, evidence that the deal is in his self-interest, at least relative to the alternative scenario in which Peter drowns. Indeed, one of the challenging aspects about explaining what is morally wrong with exploitation derives from the fact that “exploitation can also be mutually beneficial, where both parties walk away better off than they were *ex ante*” (Zwolinksi and Wertheimer 2016).

Second, even though Ed exploits Peter in offering him a bad deal, Ed does not *coerce* Peter into accepting the offer. While one might be inclined to describe an “offer that one can’t refuse” as coercive (e.g., Feinberg 1989), as Giublini and Savulescu (2021, p. 377) point out, “not everything that makes an option difficult or unreasonable to refuse is coercive.” Coercion necessarily involves the removal of options (Wertheimer 2014), typically accompanied by threats of violence for refusal—the paradigmatic example being robbery at gunpoint. However, in the Drowning Man case, Ed *expands* Peter’s range of options without removing any previously available options; moreover, Ed does not threaten Peter in any way. Ed does not, for instance, promise to expedite Peter’s imminent drowning if Peter refuses to give up 50% of his future earnings. Third, it is clear that even though Ed exploits Peter, Ed does not *deceive* Peter. It is clear what the terms of the deal for being rescued are, and Peter, we can suppose, is of sound mental capacity at the time. So, if Peter accepts the deal, then Ed has acquired Peter’s *informed* consent.

Fourth, Ed exploits Peter in the Drowning Man case even if Ed did not causally contribute to Peter’s boat capsizing. It is because of some unjust or simply unfortunate set of background conditions that Peter finds himself in danger. Ed arrives just in time to exploit this vulnerability in Peter’s situation, but Ed need not have contributed to these background conditions in order to have exploited Peter. It is enough for Ed to simply offer Peter an unfair deal in light of Peter’s vulnerability for this to count as an instance of exploitation. If Ed significantly causally contributed to some defect in Peter’s boat, which led to its capsizing, then Ed would probably have a moral obligation to save Peter without receiving any compensation. To offer Peter the same unfair deal in the alternate scenario where Ed is morally responsible for the boat capsizing would count as a more serious moral wrongdoing. In the alternate scenario Ed *exploits* Peter and wrongs him in some additional, more serious way, by violating his rights, or by treating him unjustly.[[7]](#footnote-7)

In sum then, what the Drowning Man case shows is that it is possible for *A* to exploit *B*, even if i) *A* does not harm *B*, ii) *A* benefits *B*, iii) *A* does not coerce *B*, iv) *A* does not deceive *B*, and v) *A* does not causally contribute to *B*’s vulnerable position. Generally, *A* exploits *B* whenever *A* takes unfair advantage of *B*’s vulnerability, whether or not *A* contributed to *B*’s vulnerable state.

*4.2. Can We Dispense with the Pattern-Recognition Instinct?*

 As we suggested above, we believe that social media websites that host online misinformation that Big Data algorithms can predict that users will click on—which ultimately leads to advertising revenue—are exploiting the epistemic vulnerabilities of Internet users. We have up to this point attempted to show how online misinformation takes advantage of the weakened epistemic state of consumers, by reference to our discussion of the pattern-finding instinct. It is clear that the general public is made more vulnerable by the explosion of information that characterizes the Big Data movement along with manipulative market forces, and that this vulnerability is primarily epistemic in nature. Now, it is true that users of social media websites derive a benefit from being targeted in this way, namely the free use of these apps and websites. However, just like in the Drowning Man case, we argue that this mutually beneficial transaction falls under standard accounts of exploitation.

Unlike in the case of the drowning Peter, one might wonder whether the vulnerability individuals have because of the pattern-processing instinct doesn’t have an easy solution: attempting to override this instinct when it manifests itself and become wary of patterns we are inclined to see online. Wouldn’t this eliminate the vulnerability? As we discussed early on, the pattern-finding instinct was crucial for the survival of early humans, but it probably serves has somewhat less survival value in the life of the average United States citizen today. After all, the average person no longer needs to use patterns to determine which plants should be avoided or consumed; they no longer need to follow herds of animals to locate water; they no longer need to piece together a forecast from previous weather patterns. For the most part, we have technology that predicts the weather, grocery stores that carry only edible food, and virtually never find ourselves without access to potable water. To be sure, we owe this technology partly to the pattern-recognition abilities successfully employed by the scientists and engineers responsible for the creation of this technology. So, of course, it would be unfeasible to recommend a wholesale condemnation of patterns. But for those of us who are not domain-specific experts, perhaps the rational response to the phantom pattern problem is, following David Hume, to simply refuse to go beyond our data. Given the current epistemic situation, if we can survive without assenting to any patterns *whatsoever* then, perhaps we should do so.

Ultimately, we think this line of thinking is far too extreme and is thus misguided. While overriding as best as we could our pattern-finding abilities would serve to reduce our vulnerabilities, this proposal is far from feasible. Often, patterns demand our assent without our having much choice in the matter. Furthermore, we want to argue that the pattern-recognition instinct is an intrinsic good that is crucial, not only for human survival, but also for human flourishing. The pattern-recognition instinct should be valued, we believe, as a particularly humane instinct, and recognized as a crucial tool for the participation in activities that lie at the heart of what is unique about humanity itself. It is because the pattern-finding instinct conduces to human flourishing in this way that it seriously wrong to exploit individuals by means of taking advantaging of their propensity to see patterns in the world.

*4.3. The Pattern-Finding Instinct as a Humane Instinct*

In 1910, Pablo Picasso put paint to a canvas in various shades of brown, grey, and white, composing a piece representative of the Cubist school of art. If one were to take the painting and divide it into several parts, it would be easy to see it as a work of abstract art. Near the upper left-hand corner of the painting, one can discern an elongated cube that vaguely resembles a fuzzy shoebox. Other segments of the canvas look similar—shoebox shaped objects standing out from or sinking into the canvas, but no discernable object being depicted. Four areas of the canvas boast the painting’s only non-straight lines, loose sweeping rounded objects reminiscent of the “golden spiral,” a geometric figure whose growth factor is the golden ratio. Each part of the canvas is lovely and haunting, if not identifiable, on its own. But taken together, what becomes immediately apparent to the observer of the piece is that one is looking at, as the title of the painting suggests, a *Girl with a Mandolin*.

 Picasso’s ability to create this piece relies on the assumption of functioning pattern-recognition abilities in his observer. When observing the painting as a whole, one no longer sees abstract sea-shell shaped curved figures dotting the middle of the canvas, but rather the sweep of the mandolin player’s hair, the curve of her breasts, the body of her musical instrument. The image we see changes once we recognize it *as a pattern*, and it is often difficult or impossible to return to seeing it as a random collection of abstract shapes. This is the common technique of Cubism, as well as other styles of art: Pointillism and Symbolism, for instance, just to name two modern schools. The artist takes a chance on her audience—she bets that they will immediately perceive the pattern she is indirectly representing, and that, from the audience’s recognition, her creation in its full form will emerge.

 The visual arts are, of course, not the only domain of art in which the pattern-finding instinct is relied upon. In an interview with the *Paris Review*, Vladimir Nabokov describes the method of his writing as a process of creating and filling in patterns: “The pattern of the thing precedes the thing. I fill in the gaps of the crossword at any spot I happen to choose. These bits I write on index cards until the novel is done”.[[8]](#footnote-8) Patterns are also frequently used as a literary device for conveying a theme or something about a character’s motivations or fate. For example, in Virginia Woolf’s *Mrs. Dalloway*, the intermittent chiming of Big Ben, over time, draws the reader’s attention to the novel’s emphasis on the passage of time and the briefness of each hour of a human life. And in Vonnegut’s *Slaughterhouse-Five*, the pattern of the utterances of the phrase “so it goes” comes to signify death.

In addition to using patterns of our own creation for communication with others, the recognition of natural—or perhaps even *supernatural*—patterns has structured human life since the beginning of the known history of humanity. For religious believers, the religious life is animated by the faith that there is a higher reality beyond the world we know. In traditional Christianity, Judaism, and Islam, as well as many other non-Western religions, signs and wonders sometimes offer glimpses behind the curtain. Patterns of ceremonial practice are often the conduit for (what practitioners recognize as) patterns of divine experience. A Catholic crosses himself to receive the Eucharist and experiences the real presence of Christ. In preparing the Seder, a Jewish family becomes connected with the historical liberation as depicted in *Exodus*. In rituals of *puja*, a Hindu marks the significance of an anniversary or holiday by offering fire, foliage, and sustenance to the Divine, who in turn recognizes the worshiper. The parts of our lives that are at the core of our identity, both as humans and as individual persons, often rely on this innate ability to use, recognize, and interpret patterns.

 Why does the importance of the pattern-processing instinct for art, literature, religion, etc. matter? Something that makes pattern-recognition abilities stand out among other survival instincts is that it is not *merely* a survival instinct. Far beyond its utility as a tool for avoiding predators and poisons or locating food and water, the pattern-finding instinct has a distinctly *humane* function. The ability to recognize, interpret, and create patterns is part of what is unique and complex about human communication. This much was already suggested in section 2.2 in our discussion of the evolution of superior pattern-processing in human beings. Our discussion of the role of pattern-finding in connection with *humanity* offers further support for this view. Given the centrality of the pattern-recognition instinct to our cognitive, social, and emotional lives, anything that takes advantage of this instinct in a way that either destroys it or renders it useless is *prima facie* seriously morally wrong. This is true, even if we consent to and benefit from the transactions in which this exploitation takes place.

*4.4 An Analogy with Advertising and the Question of Fairness*

 As we’ve suggested, one way in which the pattern-recognition instinct may be destroyed or rendered useless is if the agent were in an environment that was uncooperative; an environment that made relying on a previously truth-conducive cognitive faculty a genuine liability. The problem that we have discussed can be thought of as another instance of a “mismatch” between our evolved tendencies and the radically different, and increasingly digital, environment in which we find ourselves in the 21st century (Li et al. 2018).[[9]](#footnote-9) To further illustrate this point, it would be helpful to look at another nearby example of an instinct that is taken advantage of in a way that looks to be exploitative. The human instinct to consume, for instance, or the instinct to gain the admiration of our peers, or the instinct to attract romantic partners, is the basis for nearly all advertising and marketing. Philosophers have raised questions regarding false advertising, also known as “puffery” (Arrington 1982), and in recent years there has been a flurry of documentaries devoted to showing the undesirability of living in society in which one is constantly bombarded with encouragements to consume more and more.[[10]](#footnote-10) The concerns raised about the overabundance of sensational, targeted advertising take many forms, one of which is a worry about whether advertising allows consumers to learn about how to fulfil the desires they already have, or whether advertising implants *new* desires in consumers that (it is promised) only the advertised product can fulfill (Arrington 1981: p. 285; Braybrooke 1967). In either case, it is certainly true that advertisements give consumers desires for particular products, whether or not the consumer had previously desired, for instance, a more modern-looking wristwatch or more voluminous hair. The current state of incessant online advertisement in the age of Big Data has produced a rather unsurprising result: the increase of global clothing consumption has increased by 400% since 1998,[[11]](#footnote-11) while new trends like skin care have a projected market growth of nearly 70% by 2026.[[12]](#footnote-12) Such skyrocketing consumption, not at all in step with world population growth, leads one to conclude that people are increasingly making purchases they would have not made just a few decades ago. Following the instinct to consume is no longer merely aiding our survival, but in many ways contributing to our demise.[[13]](#footnote-13)

What we can see in the case of advertising is the intentional exploitation and manipulation of a human instinct which, in a more cooperative environment, should be fairly reliable: when we desired things, these desires could be taken as some evidence that the thing desired would improve our lives in some way. This evidence is defeasible, of course, but in a more cooperative environment (that is, in an environment where there is not a prolific intention to manipulate the desire to consume) we could likely be less skeptical, more assured that these desires were not the result of a campaign to grab our attention. And while many have asked the requisite moral questions about advertising, the same cannot be said of the ways that the pattern-recognition instinct is thrown into a similar kind of environment. Rather than inundation with advertisements meant to exacerbate the consumption instinct, people in technologically advanced societies today find themselves adrift in a sea of data. Our pattern-recognition instincts are alerted by so many pretenders—trapped in such an uncooperative environment—that the instinct no longer reliably guides us to see things as they are.

Given the analysis of exploitation in section 4.1, in order for these instances of exploitation of the pattern-finding instinct to count as exploitation in the *moral sense*, those who benefit from this online ecosystem must do so *unfairly*. One might wonder whether the morally objectionable features of the exploitation of the pattern-finding instinct really counts as *unfair*, even if morally problematic. Discussions of fairness can be divided into two broad categories: questions about procedural fairness, and questions about substantive fairness (Zwolinksi and Wertheimer 2016). Accounts of procedural fairness are irrelevant to our discussion here—insofar as there are fair or unfair results of the market for digital data, there is no agreed-upon procedure that reaches this outcome. Theories of substantive fairness are varied and multifarious, but many accounts find a point of agreement that a fair exchange must *respect* persons and the *basic needs* of those persons (Zwolinksi and Wertheimer 2016). Some who argue for this as a criterion for a substantively fair exchange hold this view on Kantian grounds. Put simply, when we are made aware that an outcome or condition of exchange fails to respect persons, or that it deprives one of more of the parties of a basic need, we have a duty to help (Sample 2003, p. 57; Snyder 2008, p. 309). Such a criterion for substantive fairness, of course, need not be grounded in a Kantian normative theory. Many other deontic normative frameworks, as well as Rule-Utilitarianism and (with a little work) virtue theories can make sense of this criterion for a fair exchange. So, if the market for digital data creates conditions which fail to respect persons or deprive them of a basic need, as we have argued it does, then on this view the market exchange is unfair.

**5. Epistemic Rights Violations**

In the previous section, we have primarily tried to show why the exploitation of the pattern-finding instinct is morally wrong. However, as Wertheimer (1996, p. 6) points out, there is a distinction between the *moral weight* of exploitation, i.e., its “degree of wrongness”, and its *moral force*, i.e., which actions we have reason to take in light of the presence of exploitation. It is conceivable that the exploitation of individuals in virtue of the unfair taking advantage of the pattern finding-instinct is morally wrong, but nevertheless it would be morally *im*permissible to take steps to interfere with or prohibit these transactions, especially in light of the fact they are consensual (i.e., we choose to use social media) and beneficial (i.e., we get to use it for free). One way of showing the moral weight of exploitation has moral force is to argue that the exploitation of the pattern-finding instinct constitutes a kind of rights-violation. Typically, rights violations give third parties some reason for action.

 Recently, Watson (2018) has argued for a category of human rights that are specifically *epistemic* in nature, where our “epistemic rights” can be violated if we are deprived of the epistemic goods to which we have rights. Watson defines an “epistemic right” as follows (2018, pp. 91-92):

A moral right is an epistemic right just in case it pertains to epistemic goods. My right to information about my blood-sugar levels, after being tested for diabetes, is a moral right that pertains to the epistemic good of information. It is, therefore, accurately characterized as an epistemic right. Recognizing that epistemic rights are often (perhaps always) derivative from moral rights helps to underscore the normative force of epistemic rights.

So, epistemic rights, as Watson conceives of them, are things I am owed (moral rights) which pertain to specifically epistemic goods. What counts as an “epistemic good”? While Watson does not go into detail about exactly how to understand an epistemic good, the bulk of her examples involve information that someone has a right to, e.g., the results of a blood test, or (the primary focus of her paper), political information regarding the Brexit proposals, etc. But one can imagine other sorts of epistemic goods to which one may have a right. For example, I may have the right to retain the integrity a particular epistemic capacity. Admittedly, it is easier to see how one can have an epistemic right to information, and easier to see how such rights could be violated—namely, the information is simply withheld. But epistemic rights to epistemic capacities can be violated as well. To illustrate our contention, consider a detective who possesses a photographic memory that she uses to help her solve murder cases. This is an epistemic capacity, and one whose integrity the detective has a right to retain. Her epistemic right could be violated in a few different ways. It could be violated if she were kidnapped by a murderer whom she was pursuing who lobotomized her in a way that destroyed her photographic memory. It could also be violated if, instead of *destroying* her photographic memory, the murderer rendered the capacity *useless* by implanting in her thousands of false photographic memories.

 Of course, there are a multitude of conceptions of human rights, and they are likely not all compatible with a category of epistemic rights as Watson conceives of them. For Nozick (1974), human rights are purely negative—we have rights to not be interfered with, so long as we avoid interfering with others. Nozick calls these “side constraints” on the actions of others. I have the right to not have my property forcibly taken from me, the right to not be touched or struck without my consent, etc. I do not, however, have rights *to* things except insofar as types of non-interference can be things to which I have rights. For a Nozickian, I have no automatic right to any (to use Watson’s example) personal information about myself that I do not already possess. I do, however, have property rights over things for which I have paid or entered into some binding agreement (Nozick 1974, p. 171; Mack 2018). Even so, a Nozickian conception of rights may have room for the sort of rights Watson is talking about in her examples—one’s medical information, for instance, but this would likely be subsumed under a property right. But what about epistemic capacities? A Nozickian may have greater room for a category of specifically epistemic rights if those rights concern capacities that one possesses naturally. Actions from others that destroy or render useless such capacities may fall under the umbrella of “interference”. So even given a thin conception of rights as purely negative, one may accept that there is a human right to not have their epistemic capacities rendered useless.

 Additionally, one need not even accept any ontologically robust understanding of “rights” to accept our general argument that the moral weight of exploitation has moral force. For example, someone who thinks that talk of “rights” is merely (misleading) shorthand for a more detailed analysis of moral desert/obligation/wellbeing, etc. (for instance, MacIntyre 1983) may be more than happy to agree that, in Watson’s example, the doctor has the obligation to give the patient her blood test results, or that the patient deserves these results, or that the doctor would be acting badly to withhold the patient’s information, and so on. There are many ways the foundational normative landscape may look given the analysis that, we believe, most of us will agree on: that there are some epistemic goods (be they information, abilities, capacities, etc.) that it is *prima facie* wrong to destroy or render useless

 As we have argued, the pattern-finding instinct is very plausibly an epistemic good. Beyond its initial adaptation as a tool for survival, we find that we rely on our ability to identify, interpret, and create patterns as we navigate aspects of a meaningful human life: understanding and communicating with others linguistically, creating and appreciating the visual, musical, and literary arts, reading faces and emotions, and participating in communal and religious rituals. We could say that, beyond being an innate *human* instinct, the pattern-finding ability is a distinctly *humane* instinct. It is plausible that each of us possesses an epistemic right to retain this deeply good epistemic capacity. To be deprived of this capacity, therefore, would fit the definition of an epistemic rights violation. For this reason, the exploitation of the pattern-finding instinct *does* give us some normative grounds for action.

**6. Some Proposed Solutions and Some Difficulties**

To our knowledge, the problem of the exploitation of the pattern-recognition instinct has not yet been addressed in the philosophical literature. However, a number of solutions have been proposed, by both academics and politicians, for tackling problems of fake news, online misinformation, etc., i.e., conditions that gives rise to this exploitation. While our focus in this paper is not to suggest a solution to this kind of epistemic exploitation, here we will briefly discuss some of the more prominent proposals, as well as some of the criticisms that have been raised against them.

*6.1 Critical Thinking*

 A natural response to this more general problem of online misinformation is that we should increase our support for “critical thinking” education. One might argue that the epistemic failures in socio-political discourse that have manifested in the last few years are the result of insufficient attention to the established forms of inductive or deductive reasoning. On this view, the problem is that too many participants in public discourse are suffering from “epistemic vices”, which lead them to hold beliefs that should, if they were reasoning correctly, would appear to be demonstrably false (e.g., Cassam 2016). Perhaps critical thinking education could help, then, with the phantom pattern problem. If we were better critical thinkers, then we would be made less vulnerable, and thus exploitation of this instinct would be less of a problem.

 The critical thinking response depends, of course, on whether an epistemic vice explanation is the best explanation of the general problem of belief in online misinformation. Admittedly, this seems like an antecedently plausible explanation, and what’s more recent empirical work on the correlation between epistemic vice and belief in COVID-19 conspiracy theories seems to confirm it (Meyer et al. 2020). However, one should be cautious about extrapolating too hastily here. For one thing, it has recently been shown that different conspiracy theories about the origins of the COVID-19 virus have “different psychological and social determinants” (Hartman et al. 2021). Furthermore, if one thinks the current situation with regards to belief in online misinformation is uniquely worrisome, as we do, then the proponent of the epistemic vice explanation owes us some further explanation for the apparent precipitous decline in intellectual virtue. In our view, it is much more plausible that the radically new digital environment, rather than a massive decline in intellectual character *per* *se*, explains the current the situation with regards to belief in online misinformation.

However, even if epistemic vice explains why people fall prey to online misinformation in the form of phantom patterns or otherwise, as Fritts and Cabrera (forthcoming) argue at length, there is good reason, both empirical and theoretical, to be skeptical of proposals to increase critical thinking education as a means of combatting belief in misinformation. Often, courses in critical thinking education focus on *formal*, abstract elements of reasoning, when the problems in reasoning are often *material*, having more to do with domain-specific considerations that are difficult to address in a general critical thinking course. And, unfortunately, there is little empirical evidence to suggest that interventions that focus on *media literacy* diminish belief in fake news headlines (Bulger & Davison 2018). Furthermore, in an uncooperative epistemic environment, perhaps an increase in critical thinking education will only lead to more entrenched conspiracy theorizing because of the “sophistication effect” (Taber & Lodge 2006), “whereby being more knowledgeable provides more ammunition with which to counter unpalatable claims” (Levy 2017, p. 33). It may be that a citizenry possessed of intellectually virtuous character traits such as conscientiousness, thoroughness, carefulness, humility, etc. would be less credulous with respect to online misinformation. However, it is difficult to inculcate these traits in those that lack them (Tanesini 2016; Kotzee et al. 2021). Ultimately then, we think that any intervention to deal with the problems that we have highlighted should focus more on modifying the environment, rather than attempting to mold individual’s intellectual character in the right ways.

*6.2 Flagging Misinformation*

Another attempted solution to reducing the *spread* of online misinformation, which may help address the phantom pattern problem, has been to alert social media users of false or misleading media being shared on their sites. In 2017, Facebook began to “flag” article that contained false or misleading information. Users themselves could report the article as fake news by clicking a button located next to the headline of an article. The article was then evaluated by third-party “fact-checkers”. Alfano et al. (2020, p. 19) suggest that false or misleading videos, on platforms such as YouTube, also ought to be flagged and reviewed by fact-checkers. Regina Rini (2017) advances a similar proposal, arguing that social media sites should let users report other site users who engage in the spread of online misinformation. So, perhaps, Facebook should have flagged the post in which Marjorie Taylor Greene discussed the phantom pattern involving the Rothschilds and the California wildfires.

Unfortunately, Facebook’s flagging campaigned resulted in more deeply entrenched beliefs in the claims of the disputed articles by users who saw the flagging of articles as a partisan move to suppress dissenting voices and counter-narratives. Facebook product manager Tessa Lyons wrote: “Academic research on correcting misinformation has shown that putting a strong image, like a red flag, next to an article may actually entrench deeply held beliefs – the opposite effect to what we intended.”[[14]](#footnote-14) It is not clear, then, that flagging articles or prominent posts that discuss phantom patterns would have the desired effect. Perhaps this solution would make things worse.

*6.3 Amend Section 230 of the Communications Decency Act*

Finally, in another effort to reduce the spread of online misinformation, some politicians, legal theorists, and public policy centers have proposed amending Section 230 of the Communications Decency Act (CDA 230).[[15]](#footnote-15) Passed in 1996 when the Internet was in its infancy, CDA 230 has proved to be one of the most important laws in the US legal code dealing with the regulation of the Internet. In broad outline, CDA 230 provides a broad and sweeping liability shield for internet companies, e.g., social media outlets such as Facebook or Twitter, for any unlawful content created or developed by any third party using their services, subject to a few exceptions. It is because of CDA 230, for example, that Alice cannot sue Facebook for defamatory content posted by Bob on Facebook. With CDA 230 protections in place, Facebook is free to permit a wide latitude of speech on its platform without worrying about being “sued out of oblivion” (Kosseff 2019, p. 276). It is for this reason that CDA 230 has been called the “most important law protecting Internet speech” by the *Electronic Frontier Foundation*, a prominent digital civil rights and liberties non-profit.[[16]](#footnote-16) Indeed, the entire business model of many of the largest and most successful tech companies, e.g., Yelp!, is made possible by CDA 230 and its subsequent broad interpretation in the courts (Kosseff 2019, pp. 120-2).

For a recent example of a proposal to modify CDA 230 in order to curtail the spread of online misinformation, consider the “Health Misinformation Act of 2021”, introduced by Senators Amy Klobuchar (D-Minnesota) and Ben Ray Luján (D-New Mexico) in July of 2021.[[17]](#footnote-17) This bill, which is obviously a response to online misinformation during the COVID-19 pandemic, would revoke CDA 230 protections for any interactive computer service, e.g., social media websites, that “promotes…health misinformation through an algorithm.” The purpose of this bill is to incentivize internet companies to take greater measures to combat the spread of misinformation by ensuring that their algorithmic recommender systems don’t amplify health misinformation; otherwise, these companies would risk exposing themselves to a flurry of potentially devastating lawsuits by being denied CDA 230 liability protections. To be sure, any proposal to amend CDA 230 along the lines of the Health Misinformation Act raises several serious concerns, which we’ve already hinted at, regarding freedom of expression, state censorship, the future of a free and open Internet, etc. We cannot say with certainty what the Internet would look like without CDA 230, but plausibly a world without CDA 230 would be one where the Internet is more akin to a “closed, one-way street that looks more like a broadcaster or newspaper and less like the Internet we know today” (Kosseff 2019, p. 276). So, while we won’t go so far as to endorse repealing CDA 230 as some have done,[[18]](#footnote-18) we find the general idea of amending CDA 230 in some way to nudge social media companies to do a better job moderating their platforms to be a proposal worth pursuing.[[19]](#footnote-19) Perhaps a change like this in the *legal* environment may help us to better contend with the problem of online misinformation.

**7. Conclusion**

 In this paper we have sought to draw attention to an under-appreciated aspect of the appeal of some forms of digital misinformation: the way they exploit the pattern-finding instinct by drawing our attention to phantom patterns that emerge from the flood of data available online. We have argued that, in virtue of relying on our pattern-recognition instinct to appear plausible, the market for digital misinformation often epistemically exploits consumers of this media. This exploitation occurs in virtue of two facts about the market: first, that the market takes advantage of the pattern-recognition instinct, allowing social media websites to benefit from our vulnerability and second, that this is an unfair advantage because the pattern-recognition instinct is a basic epistemic good.

 It has not been our goal here to offer a complete moral analysis of the market for digital. However, as we have argued, the exploitation of individuals in virtue of taking advantage of the pattern-finding instinct very plausibly constitutes at least a *prima facie* moral wrong, and it is difficult to see what kinds of factors could override this wrong in such a way that the creation and dissemination of such media is rendered permissible. While we have argued that the moral weight of this exploitation has moral force, we have not attempted to provide a solution to this problem here, though many proposals have been raised by others which may help to mitigate the initial conditions that give rise to this exploitation. It is our hope that, in thinking more deeply about how humans recognize and respond to patterns, we can continue to explore ways to retain this basic epistemic good—this deeply human capacity—without falling prey to those who would wield it against us.

References

Alfano, M. & Skorburg, J.A. (2018). “Extended knowledge, the recognition heuristic, and epistemic injustice,” In D. Pritchard, J. Kallestrup, O. Palermos, and J. A. Carter (eds.), *Extended Knowledge*. Oxford University Press.

Alfano, M. and Klein, C. (2019). “Trust in a Social and Digital World,” *Social Epistemology Review and Reply* Collective, 8(10): 1-8.

Alfano, M., Carter, J.A., & Cheong, M. (2018). “Technological Seduction and Self-Radicalization,” *Journal of the American Philosophical Association* (3):298-322.

Alfano, M., Fard, A.E, Carter, J.A., Clutton, P. & Klein, C. (2020). “Technologically scaffolded atypical cognition: The case of YouTube’s recommender system,” *Synthese*: 1-24.

Aliapoulios, M., Bevensee, E., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., & Zannettou, S. (2021). “An early look at the parler online social network,” *arXiv:2101.03820.*

Amarasingam, A., & Argentino, M. A. (2020). “The QAnon conspiracy theory: A security threat in the making,” *CTC Sentinel* 13(7): 37-44.

Arrington, R. L. (1982). “Advertising and Behavior Control”. Journal of Business Ethics 1 (1):3 - 12.

Barrett, P. (2020). “Regulating Social Media: The Fight Over Section 230—and Beyond,” *NYU STERN – Center for Business and Human Rights.*

Benkler, Y., Faris, R. & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics.* Oxford University Press.

Berenstain, N. (2016). "Epistemic Exploitation," *Ergo*, 3(22): 569–90.

Braybrooke, D. (1967). “Skepticism of Wants, and Certain Subversive Effects of Corporations on American Values”. in Hook, S. (ed.), *Human Values and Economic Policy*. New York University Press.

Brennan, J. (2019). “Should employers pay a living wage?” *Journal of Business Ethics,* 157(1): 15-26.

Brod, C. (1984). *Technostress: The Human Cost of the Computer Revolution*. Reading, Mass.: Addison-Wesley.

Bulger, M., & Davison, P. (2018). “The Promises, Challenges, and Futures of Media Literacy*,” Journal of Media Literacy Education*, 10(1): 1-21.

Calude, C.S. & Longo, G. (2017). "The Deluge of Spurious Correlations in Big Data,” *Foundations of Science* 22(3): 595-612.

Cassam, Q. (2019). Conspiracy Theories. Cambridge: Polity Press.

Dennett, D.C. (1996). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Schuster.

Dentith, M. R. (2019). “Conspiracy theories on the basis of the evidence,” *Synthese*, 196(6): 2243-2261.

Dentith, M.R. (2014). The Philosophy of Conspiracy Theories. Palgrave-Macmillan.

Feinberg, J. (1989). *The moral limits of the criminal law*. Oxford University Press.

Fitelson, B. and Sober, E. (1998). “Plantinga's probability arguments against evolutionary naturalism,” *Pacific Philosophical Quarterly*, 79(2): 115-129.

Flick, D. (2017). “Combatting Fake News: Alternatives to Limiting Social Media Misinformation and Rehabilitating Quality Journalism”, 20 SMU Sci. & Tech. L. Rev.

Fritts, M. & Cabrera, F. (forthcoming). “Fake News and Epistemic Vice: Combating a Uniquely Noxious Market,” *Journal of the American Philosophical Association.*

Gelfert, A. (2018). “Fake News: A Definition,” *Informal Logic* 38(1): 84-117.

Gelfert, A. (2021). “What is Fake News?” In M. Hannon & J. de Ridder (eds.), *The Routledge Handbook of Political Epistemology*. Routledge, pp. 171-80.

Giubilini, A., & Savulescu, J. (2021). “Stopping exploitation: Properly remunerating healthcare workers for risk in the COVID‐19 pandemic,” *Bioethics*, 35(4): 372-379.

Hartman, T.K., Marshall, M., Stocks, T.V., McKay, R., Bennett, K.M., Butter, S., et al. (2020). “Different conspiracy theories have different psychological and social determinants: Comparison of three theories about the origins of the COVID-19 virus in a representative sample of the UK population,” *Frontiers in Political Science*, 3: 1-17.

Hey, T., Tansley, S., and Tolle, K. (2009). “Jim Grey on eScience: A transformed scientific method.” In: Hey T, Tansley S and Tolle, K. (eds.) *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond: Microsoft Research, pp. xvii–xxxi.

Himma, K. E. (2007). “The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition,” *Ethics and Information Technology*, 9(4): 259-272.

Hookway, C. (2003). “Affective states and epistemic immediacy,” *Metaphilosophy*, 34(1‐2): 78-96.

Kitchin, R. (2014). “Big data, new epistemologies and paradigm shifts,” *Big Data & Society*, 1(1): 1–12.

Kofta, M., Soral, W., & Bilewicz, M. (2020). “What breeds conspiracy antisemitism? The role of political uncontrollability and uncertainty in the belief in Jewish conspiracy,” *Journal of Personality and Social Psychology*, 118(5): 900-918.

Kosseff, J. (2019). *The twenty-six words that created the Internet*. Cornell University Press.

Kotzee, B., Carter, J.A., & Siegel, H. (2021). "Educating for Intellectual Virtue: A Critique from Action Guidance," *Episteme*, 18(2): 177-99.

Levy, N. (2017). “The Bad News About Fake News.” *Social Epistemology Review and Reply Collective*, 6(8): 20-36.

Li, N. P., van Vugt, M., & Colarelli, S. M. (2018). “The evolutionary mismatch hypothesis: Implications for psychological science*,” Current Directions in Psychological Science*, 27: 38–44.

MacIntyre, A. C. (1983). *After Virtue: A Study in Moral Theory*. University of Notre Dame Press.

Mattson, M. (2014). “Superior pattern processing is the essence of the evolved human brain,” *Frontiers in Neuroscience*, 8(265).

Mayer-Schonberger, V. and Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray Publisher.

Mazzocchi, F. (2015). “Could big data be the end of theory in science? A few remarks on the epistemology of data-driven science,” *EMBO Reports* 16(10): 1250–1255.

McIntyre, L. (2018). *Post-Truth.* Cambridge, MA: MIT Press.

Meyer, M., Alfano, M., & De Bruin, B. (2021). “Epistemic vice predicts acceptance of Covid-19 misinformation,” *Episteme*, 1–22. doi: 10.1017/epi.2021.18.

Nozick, R. (1969). “Coercion. In S. Morgenbesser, P. Suppes, & M. White (eds.), *Philosophy, science, and method*. New York: St. Martin’s Press.

Nozick, R. (1974). *Anarchy, State, and Utopia*. Basic Books.

O’Connor, C & Weatherall, J.O. (2019). *The Misinformation Age*. New Haven, CT: Yale University Press.

Pigliucci, M. (2010). “What Darwin Got Wrong,” *Philosophy Now*, 81: 38-39.

Resnik, M. (1997). *Mathematics as a Science of Patterns*. Oxford University Press.

Rini, R. (2017). “Fake News and Partisan Epistemology,” *Kennedy Institute of Ethics Journal*, 27 (S2):43-64.

Sample, R. (2003). *Exploitation, What It Is and Why it is Wrong*. Lanham, MD: Rowman and Littlefield.

Sauer, H. (2018). *Debunking Arguments in Ethics*. Cambridge University Press.

Schwartz, B. (2004). “The Tyranny of Choice,” *Scientific American*, 290(4): 70-75.

Simon, H. A. (1979). *Models of Thought*. Yale University Press.

Smith, G. and Cordes, J. (2020). *The Phantom Pattern Problem: The Mirage of Big Data*. Oxford University Press.

Snyder, J. (2008). “Needs Exploitation”. *Ethical Theory and Moral Practice*, 11(4): 389–405.

Taber, C. S. & Lodge, M. (2006). “Motivated Skepticism in the Evaluation of Political Beliefs,” *American Journal of Political* Science, 50: 755–769.

Tanesini, A. (2016). “Teaching Virtue: Changing Attitudes,” *Logos and Episteme*, 7(4): 503-527.

Wertheimer, A. (1996). *Exploitation*. Princeton: Princeton University Press.

Wertheimer, A. (2014). *Coercion*. Princeton University Press.

Wicker, A. (2016). “Fast fashion is creating an environmental crisis.” *Newsweek.* https://www.news week.com/2016/09/09/old-clothes-fashion-waste-crisis494824.html?fbclid=IwAR3 \_PmY0w1kIGurhqGkE2ExJ6wCHUtdrb-G6vRuRkqOreRfeWrZ-zdwmzQw

Williams, B. (1985). *Ethics and the Limits of Philosophy*. Harvard University Press.

Wood, A. (2016). “Unjust Exploitation,” *The Southern Journal of Philosophy*, 54: 92-108.

Zwolinski, M. & Wertheimer, A. (2016). "Exploitation", *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition)*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/ exploitation/>.

1. Berenstain (2016) introduces the concept of “epistemic exploitation” to refer to cases in which “privileged persons compel marginalized persons to produce an education or explanation about the nature of the oppression they face” (p. 570). While our focus differs from that of Berenstain in important ways, our analysis of epistemic exploitation may be of sufficient generality such that the harms to marginalized persons that Berenstain investigates may indeed count as “epistemic exploitation” in our sense. [↑](#footnote-ref-1)
2. https://www.idc.com/getdoc.jsp?containerId=prUS47560321 [↑](#footnote-ref-2)
3. https://www.vox.com/22256258/marjorie-taylor-greene-jewish-space-laser-anti-semitism-conspiracy-theories [↑](#footnote-ref-3)
4. https://www.independent.co.uk/news/world/americas/naomi-wolf-twitter-ban-vaccine-b1860320.html [↑](#footnote-ref-4)
5. As Himma (2007, p. 266) points out, awareness of this problem goes back to at least the 80s, when the psychologist Warren Thorngate remarked upon some “ominous” trends in rates at which information was being produced. [↑](#footnote-ref-5)
6. For ease of expression, we will often speak of the “exploitation of the pattern-finding instinct”, but this should be regarded as shorthand for “the exploitation of individuals whose pattern-finding instinct is taken advantage of.” [↑](#footnote-ref-6)
7. See Wood (2016) for a helpful discussion of the distinction between unjust exploitation and cases of exploitation that are morally objectionable, but which do not constitute injustices. [↑](#footnote-ref-7)
8. https://www.theparisreview.org/interviews/4310/the-art-of-fiction-no-40-vladimir-nabokov [↑](#footnote-ref-8)
9. See Sauer (2018, pp. 38-40) for a discussion of this mismatch problem in the context of disgustsensitivity and moral judgments, as well as Alfano and Skorburg (2018) for a discussion of the ways in which the “representation heuristic” has become unreliable when placed in a malfunctioning media environment. [↑](#footnote-ref-9)
10. See: *The True Cost* (2015), *The Great Hack* (2019), and *The Social Dilemma* (2020) for three prominent ones. [↑](#footnote-ref-10)
11. https://sustainability.uq.edu.au/projects/recycling-and-waste-minimisation/fast-fashion-quick-cause-environmental-havoc?fbclid=IwAR3zctssU8eBiz20JjjnDrhOzb1DNKBY7AVQVgBIHf8RwogQpIFVPY0inb0 [↑](#footnote-ref-11)
12. https://www.globenewswire.com/news-release/2021/04/15/2210827/0/en/Skincare-Products-Market-Growth-Trends-COVID-19-Impact-and-Forecasts-2021-2026.html#:~:text=The%20Global%20Skincare%20Products%20Market,the%20market%20growth%20by%201. [↑](#footnote-ref-12)
13. For example, consider the “fast fashion industry”, which, in the United States alone, results in 14 million tons of clothing going to landfills each year (Wicker 2016). [↑](#footnote-ref-13)
14. https://www.bbc.com/news/technology-42438750 [↑](#footnote-ref-14)
15. See Barrett (2020) for an overview of the debate over CDA 230, including some recent proposals to amend it. Some proposals to amend CDA 230 are concerned with other online ills aside from misinformation, e.g. hate speech and harassment. See Kosseff (2019) for a book-length treatment of the origins and history of CDA 230. [↑](#footnote-ref-15)
16. https://www.eff.org/issues/cda230 [↑](#footnote-ref-16)
17. https://www.govtrack.us/congress/bills/117/s2448/text/is [↑](#footnote-ref-17)
18. Both former U.S. president Donald Trump and current U.S. president Joe Biden have expressed support for repealing CDA 230 (Barrett 2020, p. 2). [↑](#footnote-ref-18)
19. See Hwang (2020) for further discussion of the proposal to modify CDA 230 specifically to deal with the problem of online misinformation. [↑](#footnote-ref-19)