Manuscript version: Author’s Accepted Manuscript
The version presented in WRAP is the author’s accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/107624

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.
The Dark Side of Information Proliferation

Thomas T. Hills
Department of Psychology, University of Warwick

Words in main text: 3088
Words in abstract: 165

Contact details:
Thomas T. Hills
University of Warwick
Gibbet Hill Road
Coventry CV47AL UK
Email: t.t.hills@warwick.ac.uk

Acknowledgements: This work was supported by the Royal Society Wolfson Research Merit Award (WM160074) and a Fellowship from the Alan Turing Institute. Thanks to Li Ying, Cynthia Siew, So Young Woo, Yonghua Cen, Annasya Masitah, Eva Jimenez Mesa, and Tomas Engelthaler for numerous helpful discussions and comments.
Abstract

There are well-understood psychological limits on our capacity to process information. As information proliferation—the consumption and sharing of information—increases through social media and other communications technology, these limits create an attentional bottleneck, favoring information that is more likely to be searched for, attended to, comprehended, encoded, and later reproduced. In information-rich environments, this bottleneck influences the evolution of information via four forces of cognitive selection, selecting for information that is belief-consistent, negative, social, and predictive. Selection for belief-consistent information leads balanced information to support increasingly polarized views. Selection for negative information amplifies information about downside risks and crowds out potential benefits. Selection for social information drives herding, impairs objective assessments, and reduces exploration for solutions to hard problems. Selection for predictive patterns drives overfitting, the replication crisis, and risk seeking. This article summarizes the negative implications of these forces of cognitive selection and presents eight warnings, which represent severe pitfalls for the naive informavore, accelerating extremism, hysteria, herding, and the proliferation of misinformation.

Keywords: misinformation, evolution, social risk amplification, social proof, attention economics
Introduction

Anti-expert speech, fake news as a political weapon, the replication crisis, relentless health warnings associating cancer with everything we eat, and the richly documented pathologies of online inattention are all symptoms of a rising awareness that information is not benign (Bawden & Robinson, 2009; Carr, 2011; Eppler, 2015; Jacoby, Speller, & Cohn, 1974; Schoenfeld & Ioannidis, 2012; Schwartz, 2004). A common contributor to each of these problems and the focus of this article is information proliferation—the capacity to access and contribute to a growing quantity of information. According to the International Telecommunications Union (2017), more than 4 billion people are now mobile-broadband users, granting each near instantaneous power to access, create, and share information. This represents a five-fold increase since 2010. Over the same period of time the number of webpages available has risen into the billions (Van den Bosch & De Kunder, 2016).¹ The result of this proliferation is that information is placed increasingly under the influence of an attention economy (e.g., Lanham, 2006) in which a growing number of people influence the evolution of information by what they choose to pay attention to (e.g., Bakshy, Messing, & Adamic, 2015; Gentzkow & Shapiro, 2010).

Herbert Simon captured the central constraint on this attention economy when he noted that “information…consumes the attention of its recipients” (p. 40, Simon, 1971). We are limited in how much information we can attend to from outside sources (sometimes called the cocktail party problem, e.g., Conway & Cowan, 2001) and how much information we can process (e.g., in working memory, Miller, 1956). Thus, too much information threatens us with information overload and other pathologies of attention that have been well-documented elsewhere (Bawden & Robinson, 2009; Carr, 2011). This article focuses on a less well-documented but perhaps more pernicious problem: how information rich environments place information under the forces
of cognitive selection, driving information evolution much like other forms of selection drive biological evolution.

**The evolution of the information ecosystem**

The cognitive life-cycle of information (see Figure 1) provides a framework for understanding how cognitive selection shapes information’s evolution as it progresses from one mind to the next. In this life-cycle, cognitive selection favors information that is more likely to be searched for, attended to, comprehended, encoded, and reproduced (e.g., Hills & Adelman, 2015; Kirby, Cornish, & Smith, 2008; Blackmore, 2000).

Information proliferation influences information evolution in two ways, by increasing competition and by reducing information’s generation time. Increasing competition for attention means that more signals compete for receivers, enhancing the role of cognitive selection. More competition is also functionally equivalent to greater amounts of background noise. Communication signals adapt to increasing noise in ways consistent with information’s life-cycle, by becoming easier for receivers to detect, recall, and reproduce (e.g., Arak & Enquist, 1995; Hills, Adelman, & Noguchi, 2017; Luther et al., 2009). This favors some kinds of information over others. For example, misinformation has an advantage in competitive environments because it is freed from the constraints of being truthful, allowing it to adapt to cognition’s biases for more distinctive and emotionally appealing information (Schomaker & Meeter, 2015; Hamman, 2001). These are both factors associated with the empirical finding that lies proliferate faster than the truth (Vosoughi, Roy, & Aral, 2018).

Information proliferation also reduces information’s generation time—the time it takes for information to move from one mind to another. This is analogous to biological evolution’s generation time. A general rule of evolution is that faster generation times accelerate adaptation
(e.g., Thomas, Welch, Lanfear, & Bromham, 2010). The more rapidly people can access, select, and reproduce preferred information, the more readily will that information reflect the cognitive biases of its users.

Though cognitive selection can produce beneficial outcomes by selecting for valuable information, the focus here is on selective forces that drive unwanted outcomes of information proliferation, such as extremism, hysteria, herding, and misinformation. In the remainder of this article we will look at biases that when combined with information proliferation produce each of these outcomes. In particular, we will focus on cognition selection for information that is 1) belief-consistent, 2) negative, 3) social, and 4) predictive.

Figure 1: Information proliferation enhances the influence of cognitive selection. A) Information undergoes cognitive selection at each stage in its life-cycle, as it passes from receiver to memory to producer and on to the next receiver. B) Information proliferation (denoted by ‘a’) increases the amount of information available to receivers and competing for attention and future production. Cognitive selection implies that information loss (denotes by ‘x’s) is not random, but is instead driven by cognition’s biases for belief-consistent, negative, social, and predictive information, which influences the evolution of information in a feedforward process.
Selection for Belief-Consistent Information

Belief-consistent selection encompasses a range of psychological phenomena associated with tendencies to seek out, encode in memory, and reproduce information consistent with prior beliefs. The processes underlying belief-consistent selection are often difficult to distinguish in practice, but roughly correspond to confirmation bias, biased assimilation, and motivated reasoning, respectively. A paradigmatic example is that giving people balanced information on ideological issues frequently leaves them more polarized on such wide-ranging topics as capital punishment, legal cases, vaccines, climate change, the effects of video games, and politics (Corner, Whitmarsh, & Xenias, 2012; Greitemeyer, 2014; Lord, Ross, & Lepper, 1979; Munro & Ditto, 1997; Munro et al., 2002; Nan & Daily, 2015; Westen, Blagov, Harenski, Kilts, & Hamann, 2006).

Though Asch (1955) is often considered a standard example of selection for social information, what this research also showed is that when an individual held an unpopular belief, the presence of one other individual who shared that belief was sufficient to reinforce public commitment to this minority opinion nearly to its levels in the absence of any opposition. Similarly, when like-minded people share views—even when they all have the same information—these views tend to be held with more confidence (Schkade, Sunstein, & Hastie, 2010; Sunstein, 2002). This happens even when performance is unchanged (Tsai, Klayman, & Hastie, 2008).

Information proliferation effectively adds fuel to this fire. Belief-consistent selection becomes more prevalent as information increases (Fischer, Schulz-Hardt, & Frey, 2008; Kardes, Cronley, Kellaris, & Posavac, 2004). Information proliferation therefore leads people to personalize information and avoid belief-inconsistent information. Algorithms do this for us in
the form of recommender systems and search engines guided by browser history. On social
media, tendencies for ingroup selection lead to filter bubbles and echo chambers, which further
reduce individual exposure to information diversity even as they increase diversity across social
media as a whole (Barberá, Jost, Nagler, Tucker, & Bonneau, 2015; Nikolov, Oliveira,
Flammini, & Menczer, 2015).

The tendency to select like-minded individuals in decision making is often associated with
groupthink, with groups defensively insulating themselves from external views. Groupthink has
been blamed for numerous political and economic fiascos (Bénabou, 2012). Information
proliferation now extends the capacity for groupthink globally, organizing and polarizing groups
from political ideologues to international terrorists (Aly, 2016).

Why does belief-consistent selection exist? There are a number of likely contributors:
Selecting belief consistent information provides a rewarding feeling of sense-making while
reducing identity-threatening information (Chater & Loewenstein, 2016; Festinger, 1962); An
unwavering argument—that sweeps inconsistencies under the rug—better influences others
(Mercier & Sperber, 2011; Trivers, 2011); and people view information and the attitudes
associated with them as cultural codes, which allow communities of like-minded individuals to
identify their ingroup and cooperate on that basis (e.g., Marshall, 2015). These are all further
reinforced by people’s tendency to understand and recall information in relation to narrative
structures for causality that they already understand (e.g., Bartlett, 1932), which represents a
cognitive blind spot for information inconsistent with prior beliefs.

**Selection for Negative Information**

Heightened sensitivity to negative information is part of our evolutionary heritage. This
sensitivity leads us to weight disadvantages over advantages in information seeking and decision
INFORMATION PROLIFERATION

making (Tversky & Kahneman, 1991) and drives the well-known news trope: “if it bleeds, it
leads.” When applied to information proliferation, a negativity bias induces us to identify and
recommunicate information about risk at the expense of more balanced information.

The preferential sharing of risk information leads to social risk amplification (Kasperson et
al., 1988). In a study by Moussaïd, Brighton, and Gaissmaier (2015), information was
communicated through a chain of individuals, similar to the children’s game ‘telephone.’ The
first individual in the chain was introduced to the information through a set of balanced articles
discussing triclosan, an antibacterial agent. As this information was proliferated from one
individual to the next, it rapidly lost key facts about triclosan’s usefulness while preserving and
adding additional but unsupported information about downside risks related to health
consequences and prevalence. As a real world example, when the first Ebola case was diagnosed
in the United States this led Twitter posts mentioning Ebola to jump from 100 posts per minute
to 6000 per minute and rapidly produced inaccurate claims that Ebola could be transmitted
through food, water, and air (Luckerson, 2014).

The Ebola example reflects another potential bias in risk communication: social risk
amplification is especially prominent for dread risks—unpredictable, catastrophic, and
indiscriminant risks to life and limb such as plane crashes, nuclear disasters, epidemics, and
terrorism (Slovic, 1987). Jagiello & Hills (2018) found that social risk amplification was
substantially larger for dread risk (nuclear power) than for non-dread risk (food additives).
Moreover, Jagiello & Hills (2018) also found that socially amplified risk was resilient to the re-
introduction of the initial balanced information. This supports one of the central threats of
information proliferation: information proliferated through individuals selects for information
that is better adapted to cognition than information designed to provide a more balanced perspective (see also Ng, Yang, & Viswanath, 2017).

Social risk amplification helps explain the recent rise in what Kahneman (2011) calls the precautionary principle, a propensity to base decisions about new technology on their potential downside risks without consideration for their potential benefits. Information proliferation rapidly makes risk the reason for taking one action over another. In the extremes, this incites hysteria and motivates Ulrich Beck’s (1992) claim that we are currently living in a risk society.

Selection for Social Information

People’s appetite for social information has led some to suggest that smartphones induce hypernatural social monitoring (Veissière & Stendel, 2018) and others to observe that information on social media crowds out other kinds of information in memory (Mickes et al., 2013). Where people do not have strong ideological convictions otherwise, social information can lead to herding and undermine collective wisdom (Asch, 1955; Raafat, Chater, & Frith, 2009; Lorenz, Rauhut, Schweitzer, & Helbing, 2011).

Salganik, Dodds, & Watts (2006) studied people’s music choices by allowing people to listen to fragments of songs and then download a song of their choice. Some participants could also see what other people had chosen. Groups of independent decision makers were more like other independent groups, sharing their diversity of views, than where groups exposed to social influence, which rapidly diverged from one another. There was also more inequality among winning songs when individuals had social information than when individuals chose in isolation. In other words, social information both added noise and amplified it. The same socially-fueled ‘preferential attachment’ is blamed on long-tailed distributions of citation counts among scientific papers (Clauset, Larremore, & Sinatra, 2017) and observed in online auctions, where
bidders follow bidders instead of other quality indicators (Huang & Chen, 2006; Simonsohn & Ariely, 2008).

When problems are easy, social information can be good: social connectivity increases the rate at which groups converge on optimal solutions. But for hard problems, high levels of social connectivity prevent groups from finding optimal solutions because they lead groups to exploit suboptimal solutions too quickly (Barkoczi & Galesic, 2016; Fang, Lee, & Schilling, 2010; Mason, Jones, & Goldstone, 2008, Lazer & Friedman, 2007).

Imitating others provides humans with what Bandura (1965) called no-trial learning. Such no-trial-learning is especially effective in uncertain environments, where the informational contingencies are complex, mistakes are costly, other people have good information, and individuals need to coordinate to produce effective outcomes (Kameda & Nakanishi, 2003; Goldstone & Gureckis, 2009). But when an individual can use information from another individual to make a decision, neuroimaging evidence indicates that this turns off the executive processing associated with a more critical evaluation of the choice (Engelmann, Capra, Noussair, & Berns, 2009). The proliferation of social information therefore threatens our better judgment.

**Selection for Predictive Information**

A bias for pattern identification can be completely devoid of other biases and still run aground in a sea of data. The problem arises because information proliferation promotes spurious correlations. Consider that the probability of a false positive (type I error) approaches 1.0 as the number of comparisons increases.\(^2\) When a single researcher performs multiple tests, they can correct for this by lowering \(\alpha\) values (Bonferonni corrections). When multiple researchers perform multiple tests this problem can be alleviated by publishing negative results, avoiding the file drawer problem. But when multiple researchers test multiple hypotheses, there is presently
no cure: even if all hypotheses were pre-registered and every individual researcher reported all tests, the absolute number of type I errors would grow with the number of researchers (Ioannidis, 2005).

If positive findings enjoy a selection bias for publication, then the proportion of false positives in print will increase with the number of unique hypotheses tested. Moreover, more information (i.e., covariates) combined with a bias for positive results invites researchers down a “garden of forking paths” during analysis, where a plethora of alternative analyses make positive findings easy to generate and deceptively intuitive given post-hoc theorizing (Munafo et al., 2017; Simmons, Nelson, & Simonsohn, 2011).

Similar biases play out in economic decision making. People who experience monetary payoffs from a set of alternatives will tend to choose riskier alternatives (seeking rare but illusory gains) as the set size of possible alternatives increases (Ert & Erev, 2007; Noguchi & Hills, 2016). This happens because high-variance alternatives are more likely to produce extreme outcomes. The probability that one of them will do so increases with the absolute number of high-variance alternatives considered (Figure 2). As a consequence, the likelihood of spurious ‘super’ performers (as well as super failures) grows with the number of contenders purely as a function of statistical noise (see Denrell, 2005; Denrell & Liu, 2012; Shermer, 2014).

Data-driven machine learning is not a solution. Without input from theory, ‘blind’ machine learning can do far worse as a result of the curse of dimensionality—too many predictors and not enough data to tease them apart (Huys, Maia, & Frank, 2016). A recent article in The Economist (2013) quotes MIT’s Sandy Pentland as reporting that three-quarters of machine learning results are nonsensical due to “overfitting.” The iconic overfitting example is Google Flu Trends which rapidly went from online oracle to victim of big data hubris when it began over-predicting
influenza-like illness by a factor of two relative to the Center for Disease Control (Lazer, Kennedy, King, & Vespignani, 2014). Similar cases have been made against data-driven approaches to election prediction (Gayo-Avello, 2012) and criminology (Chan & Bennett Moses, 2016). Given the nested black-box nature of many proprietary machine learning algorithms, overfitting is often extremely difficult to detect in real-world settings, which also happen to be the places where they can do the most harm (e.g., O’Neil, 2017).

Figure 2: The probability of choosing a risky alternative as a function of the number of alternatives considered. A decision maker chooses between \( m \) alternatives, half of which are safe and pay off \( S \) with probability 1.0 and the other half are risky, paying off \( R (>S) \) with probability \( p \) and otherwise 0. As the number of alternatives to choose from increases, a decision maker who chooses the outcome with the highest sample payoff after 1 sample from each alternative will be
increasingly likely to choose a risky option. This follows $1-(1-p)^{m/2}$ and a similar logic to type I errors.

**Conclusions**

The problems described above are not easily remedied. They are in part the negative outcomes of cognitive heuristics that have paid for themselves in our evolutionary past. The adaptive value of these heuristics represents the light side of information, including features such as ingroup identification and risk avoidance (see Table 1). The dark side is that cognitive selection’s reliance on these heuristics in an information rich environment distorts what information is available and reduces our capacity to use that information objectively.

These problems can be summarized in a list of 8 warnings for contemporary information environments:

1) Information becomes more appealing to humans as it moves through human minds.
2) Balanced information about ideological issues reinforces divergent beliefs.
3) When people share similar divergent beliefs with one another, they often become more confident in their beliefs, even when they learn nothing more than that other people know what they do.
4) Information about costs proliferates more readily than information about benefits; this can prevent cost-benefit analysis and reduce decision making to risk avoidance.
5) Social information crowds out individual quality indicators and impairs objective assessments.
6) Social information reduces exploration for hard problems.
7) More information can lead to overfitting and the detection of spurious correlations driven by rare events.

8) As we become aware of more alternatives, the most appealing alternatives become increasingly risky.

These warnings reflect advantages for misinformation in information rich environments. Like science, information is to some degree self-correcting. In a just attention economy, maladaptive information should eventually lead to costs borne by those who use it. But the rate of correction necessary to resolve misinformation increases with the rate of its proliferation. Developing methods to keep up with this tide of misinformation is a growing interdisciplinary concern, cutting across science and politics alike (e.g., Lazer et al., 2018; Lewandowsky, Ecker, & Cook, 2017; see also Simon, 1976). Understanding how cognitive selection interacts with information proliferation is a key step in this process.
Table I: The effect of cognitive selection on information proliferation

<table>
<thead>
<tr>
<th>Properties of Information Adapted to Cognitive Selection</th>
<th>Adaptive Value</th>
<th>Negative Consequences when combined with information proliferation</th>
<th>Characteristic References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belief-consistent</td>
<td>Identifying the ingroup, commitment, reduction of cognitive dissonance, sense-making</td>
<td>Polarization</td>
<td>Munro &amp; Ditto, 1997; Sunstein, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overconfidence</td>
<td>Schkade et al., 2010; Tsai et al., 2008</td>
</tr>
<tr>
<td>Social</td>
<td>No-trial learning, ingroup identity formation</td>
<td>Reduced exploration</td>
<td>Mason et al., 2008; Barkoczi &amp; Galesic, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Long-tailed preference distributions</td>
<td>Salganik et al., 2006; Huang &amp; Chen, 2006</td>
</tr>
<tr>
<td>Negative</td>
<td>Risk identification and avoidance</td>
<td>Loss of positive information</td>
<td>Kasperon et al., 1988; Jagiello &amp; Hills, 2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Truth distortion amplifying apparent risks</td>
<td>Moussaïd et al., 2015</td>
</tr>
<tr>
<td>Predictive</td>
<td>Learning, identifying contingencies associated with opportunities and threats</td>
<td>Replication crisis</td>
<td>Ioannidis, 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk-seeking</td>
<td>Ert &amp; Erev, 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Super-successes/failures driven by noise</td>
<td>Denrell &amp; Liu, 2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overfitting</td>
<td>Lazer et al., 2014; Gayo-Avello, 2012</td>
</tr>
</tbody>
</table>
References


*Science, 359*(6380), 1146-1151.
Endnotes

1 According to an international team of researchers providing real-time statistics on the internet, so far today (13:43 GMT, 16th of May, 2018) 3.1 million blog posts have been posted, 41.8 million photos have been uploaded to Instagram, 398.1 million tweets have been sent via Twitter, 3.3 billion google searches have been made, and 138.4 billion emails have been sent (Internet Live Stats, 2018).

2 If tests are independent and each have a probability of success \( p \) (i.e., rejection of the null), then in \( n \) tests the probability of seeing no successes is \( (1 - p)^n \). The probability of seeing at least one success is therefore \( 1 - (1 - p)^n \), which approaches 1 as \( n \) increases for all values of \( p \).