

Research Insight

Understanding the Internet Water Army: Domain, Characteristics, Purposes, Messages and Affordances, and Consequences

Laurent H. Wang¹ and Ronald E. Rice¹

¹ Department of
Communication, University of
California, Santa Barbara

**Corresponding to
Laurent Wang**

Department of
Communication, University of
California, Santa Barbara, CA
93106-4020 USA
Email: hwang02@umail.ucsb.edu

Disclosure Statement

No potential conflict of interest
was reported by the author.

Received

28 Feb 2024

Revised

22 Apr 2024

Accepted

17 May 2024

ABSTRACT

The Internet and social media bring many benefits to individuals, groups, organizations, and society. However, some users, groups, agents, businesses, institutions, and governments also manipulate and deceive the online environment, fostering negative consequences. This article describes one specific manifestation of this phenomenon: the Internet Water Army, a growing Internet/social media phenomenon, especially in China. There has been no robust framework for referring to the Internet Water Army, nor an integrated social science review of this type of Internet/social media user, thus creating gaps and contradictions in discussions and research about it. First, the article defines, with examples, first the general domain of organized multiple online postings or contributions, second the more specific domain of paid posters or influencers promoting misinformation, and third the particular case of the Internet Water Army. Then, the explanation identifies four central dimensions of the Internet Water Army—characteristics (user demographics and activities, organization and process), purposes (at individual, organizational, and political levels), messages (content and valence) and an affordance (visibility), and consequences (desirable/undesirable, direct/indirect, and anticipated/unanticipated)—with a wide variety of examples. The article ends with a discussion of several theoretical perspectives that could increase our understanding of phenomena such as the Internet Water Army.

KEYWORDS

China, explication, Internet Water Army, social influencers, social media

Digital and online media constantly evolve and improve based on user-generated content and user interaction, especially social media. Social media facilitate user interactivity (boyd & Ellison, 2007), generate masspersonal communication (Carr & Hayes, 2015), and support collaborative activities (Rice et al., 2017). At the same time, social media also allow, foster, and disseminate misinformation and

harm, due to its low threshold of gatekeeping (Metzger & Flanagin, 2013) and easy access by people, companies, bots, and algorithms around the world. Users with malicious purposes may take advantage of the internet to distort the ordinary online environment.

One example is the *Internet Water Army* (IWA), a prevailing phenomenon in Chinese social media platforms such as Sina Weibo (Chinese version of X). The IWA has been referred to in several ways: Internet Navy, Chinese Astroturfers, Internet Zombies, Chinese army of sock puppet, or its Chinese name, Wang Luo Shui Jun (网络水军) (Wikipedia, 2020; Zeng et al., 2014). This paper uses the term *Internet Water Army* (IWA), as it corresponds the best with its original Chinese name and has been used the most frequently in past literature. Liao et al. (2021) state that the “water” portion of the Internet Water Army refers to “flooding” the online space with such posts. We define the Internet Water Army (IWA) as organized networks of paid workers (primarily low-income individuals), who are hired by individuals, organizations, or political actors, and who generate massive amounts of social media content in forms of posts and comments with varied quality, for promotional or propaganda purposes. For example, social influencers and companies may hire IWA workers to write predominantly positive reviews to promote their new product.

EXPLICATING THE IWA

In order to gain more knowledge about the IWA, and develop ways to identify and suppress IWA posting, researchers (primarily computer scientists) have been concentrating on detecting and characterizing the IWA by applying computational methods (e.g., Guo & Jiang, 2023). These analyses are primarily based on the needs of specific studies as well as the researcher’s personal understandings of the structure and

characteristics of the IWA. Yet there lacks a thorough review of the IWA phenomenon from a social science perspective. Further, few studies have addressed the consequences of the IWA. Establishing a comprehensive framework of the IWA may not only inform us about such an online phenomenon in the Chinese cultural context, but also improve our understanding of its social and political implications globally, and help researchers attend to salient distinctions.

Explication sheds light on central aspects of a concept while drawing boundaries among various terms, and thus provides a useful framework for identifying relevant aspects of the IWA phenomenon (Chaffee, 1991). Adapting Chaffee’s approach to clarifying and distinguishing concepts, the following sections describe general and specific domains of interest, and four dimensions (with subdimensions) of the IWA.

Explication Domain

“Domain” in fields such as psychology and education refers to broader (general) or more focused (specific) areas of expertise or behavior. For example, Plucker and Beghetto (2004), in their explication of creativity, assessed whether it is primarily domain general or domain specific; Kalén et al. (2021) showed that sports performance primarily depends on domain specific (sports-related), rather than domain general, cognitive functions and skills. Developing and validating measures requires conceptually defining the construct’s domain, clarifying what are the boundaries of a construct as well as its related but distinct terms, and the involved entity and properties (MacKenzie et al., 2011).

General

The general domain consists of large numbers of users, whether identified or anonymous, contributing to an online site. These refer to processes (general domain), rather than named

entities (specific domain), and focus on generally acceptable behavior, such as contributing meta-information or promoting products. A very general domain is *crowdsourcing*, or “a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task” (Estellés-Arolas & González-Ladrón-De-Guevara, 2012, p. 197). We would generalize this to remove the requirement of a specific call by and organization to engage, or in some cases even the conscious awareness of the users. Examples include Wikipedia, Flickr photo tags, Amazon recommendations, YouTube likes, Q&A sites, Linux, and GoFundMe (Doan et al., 2011). Most crowdsourcing is unpaid, but some involves compensation, such as prizes (e.g., for innovation challenges) (Doan et al., 2011). *Human flesh search* is a form of online crowdsourcing, whereby many users are invoked via a Q&A format to identify, track down, and even punish a perpetrator – a form of cyber-vigilantism (Xu, 2015, p. 264). Similarly, voluntary online *sedition hunters* have identified a number of the U.S. Capital insurrectionists on January 6, 2021 (Mak, 2021).

Another general concept is *social media influencers*, “independent, third-party endorsers who shape attitudes through... social media” (Freberg et al., 2011, p. 90). A form of online social influencing is a *brand ambassador*, who focuses on particular brands (Smith et al., 2018, p. 8). Social influencers are typically independent from the referred organization, while the brand ambassador is typically associated through an explicit affiliation, helping both to legitimize the organization and increase the online influence and status of the promoter. They typically do not receive financial compensation, but may receive free products they then discuss (Dhanesh & Duthler, 2019). Other general related user types include *crowdsourced promotion* (Kim et al.,

2018), *Internet public relations* (Wu et al., 2013), *smart and flash mobs* (Rheingold, 2002), and *self-branders* (Khamis et al., 2017). Currently receiving vast public, legal, political, and research attention is the general multi-dimensional concept of *fake or fabricated news* (Molina et al., 2021), the dissemination of false content.

Another instance of this general domain is the use of *machine-generated bots* and *digital propaganda* – in particular, their effects on interfering with the political agenda and public opinion, such as Russian bots in the 2016 U.S. election (Singer & Brooking, 2018; see also Stukal et al., 2017), or social bots in manipulating public opinion about India-Pakistan conflicts and Indian elections (Neyazi, 2020). *Computational propaganda*, applying a mixture of algorithms, anonymity, automation, and human management, is an organized attempt to generate false amplification and misinformation in order to create a bandwagon effect or false consensus. A particular form is *automated fake engagements* (Kwon et al., 2022).

Specific

Click farms and *follower factories* use (X) and increasingly Instagram (esp. its Stories feature) to game the use of likes, views, numbers of followers, and other indicators of attention and popularity (Lindquist, 2018) – essentially, an informal and often illegal process of purchasing followers. Lindquist describes the extensive accessibility to follower/click sellers, and how they are self-taught, began working while teenagers, and come from lower-middle-class families. These entrepreneurs use software and sites that offer a wide array of services, targeting different categories of followers (e.g., by gender, country), and provide followers and clicks to many resellers. The “followers” may be actual users, or generated by automated scripts; the followers may be automatically purchased and resold through website interfaces; and they may actually originate in other countries, such as India

or Russia.

Crowdturfing is a negative and more specific side of crowdsourcing, disseminating malicious URLs, developing fake grassroots campaigns, and distorting search engine results (Lee et al., 2013). Sites with from 70-95% of tasks involving crowdturfing include MinuteWorkers, MyEasyTask, Microworkers, and ShortTask (Wang et al., 2012). PR firms, agents, or businesses generally use *crowdturfing platforms or systems* developed for specific crowdsourcing applications, to initiate and coordinate mass action, for positive and negative purposes. These systems themselves have to deal with evaluating and filtering both the content (e.g., poor or false quality) and users (e.g., malicious, fraudulent) involved in these activities. Crowdturfing systems are rapidly growing, both in revenue and users (Wang et al., 2012). Similar services can be found at the US website Subvert and Profit (www.subvertandprofit.com), which claims they can access 25,000 users to perform tasks, and the UK website socioniks.com (Fielding & Cobain, 2011). Search optimization, account creation, and spam campaigns are also offered on Freelancer and eBay (Wang et al., 2012).

A *sock puppet* is defined as malicious and/or deceptive individuals holding multiple identities to vandalize online discussion and/or control public opinions (Bu et al., 2013; Kumar et al., 2017). However, some define sock puppets as individuals who, through alternative accounts, provide fake reviews of, or praise or defend, their *own* works that were earlier posted under their identified account (Seymour, 2011).

A large group of services is specifically devoted to governmental or political purposes. One is the Chinese *50 Cent* (50c) or *US Cent* party (King et al., 2017; Wu et al., 2013). “50c party” (五毛党) users provide “social media comments posted at the direction or behest of the regime, as if they were the opinions of ordinary people” (King et al., 2017, p. 484). King et al.’s (2017) analysis of the estimated nearly half billion Chinese 50c

Party posts in 2013 show that while posters have typically been thought of as similar to the IWA – i.e., as atomized, largely independent users – they are actually primarily governmental employees (though others voluntarily make similar posts). Their posts are designed primarily to divert attention away from any oppositional or controversial content (esp. if that might lead to collection action) rather than to engage in confrontation or argument, and to present positive conversations about salient topics and cheerleading for government actions or policies.

Bradshaw and Howard (2017) review what they call *cyber troops*, or “government, military or political party teams committed to manipulating public opinion over social media” (p. 3) in 28 countries (see their Table 1, p. 13). Cyber troops include contracted companies, paid individuals, political supporters, computational propaganda (bots), and fake or stolen accounts, all to spew fake news and spam, manipulate likes and retweets, shape or distract attention, and misrepresent posts as grassroots opinion. For example, China implemented a worldwide campaign aimed at Americans to promote disinformation and physical protests in a wide array of languages, social media platforms, websites, and forums (Serabian & Foster, 2021). For example, Mozur et al. (2024) report on leaked documents showing how China’s government-sponsored campaign targeting foreign governments, telecommunications firms, Chinese citizens domestically and abroad, and ethnic minorities buys data and cyberattack tools, and hires hackers, from private enterprises and contractors. Bradshaw and Howard (2017) show that all authoritarian regimes (in the countries studied) use these cyber troops to manipulate their publics. Democratic countries also implement cyber troops to influence other countries’ populations, as well as to counter terrorist propaganda, in addition to political parties using them for domestic campaigns.

Jothi and Me (2020) also note that this

business model exists around the world, such as the U.K. SocioNiks or the Australian uSocial. In The Philippines, *keyboard trolls* were hired to promote and continue to support candidate and President Duterte (Bradshaw & Howard, 2017). In Nigeria, social media entrepreneurs (using cheap smartphones and a variety of inexpensive social media and private messaging apps, esp. WhatsApp groups) who provide political support, discredit opponents, deliberately spread false information, and manipulate photos are called *propaganda secretaries* (Hassan & Hitchen, 2019). Other related specific governmental or political user types include the Russian *online army*, *troll armies*, or web brigades (Karpan, 2018), seminar users (Darwish et al., 2017), (Internet) *pushing hands*, and *vote spammers* (Wu et al., 2013).

The Internet Water Army

The specific phenomenon of the *Internet Water Army* in Chinese social media primarily involves real people (i.e., not bots). Taobao and other Chinese online platforms enable different forms of discussions such as posts, articles, videos, and Q&A sessions, in which IWA users may create various types and forms of misinformation (Emerging Technology, 2011), or create the false perception of supportive networking links and retweets. IWA workers may engage in “slander, entrapment or defamation to attack their competitors, fabricate pseudo-events, and confuse the public” (Kuang, 2018, p. 116). Silverman et al. (2020) also describe examples of marketing and PR organizations in over 15 countries creating and spreading disinformation, deception, and harassment for hire, as similar to the activities of the IWA; this growing development has been referred to as *the professionalization of deception*.

IWA workers differ from sock puppet representatives in several ways: (1) IWA users are usually a low-income population, and thus they are less likely to create fake accounts to promote

their own books or Wikipedia articles, while sock puppets are likely to be higher status content producers (such as authors), or professionals hired by a government or corporation; (2) IWA users are usually part of large, transient groups of people hired by PR companies to work for someone else rather than themselves, while sock puppets typically operate for their own purposes; (3) the IWA industry has a fairly common organizational structure (see below), while sock puppets are mostly independent users; (4) the IWA does not emphasize an individual identity to tell the story, and a frequent goal is to create an impression of general consensus among the public, while sock puppets heavily rely on fake (not anonymous) identities to foster a fake reality. Separately, King et al. (2017) distinguish the 50c Party from the IWA and other similar terms (e.g., *volunteer 50c members*, *little red flowers*, *American Cent Party*, etc.) in that the 50c Party exclusively serves the interests of Chinese government, while the IWA mainly works for companies and non-political individuals.

The IWA is an example of what Grohmann and Corpus Ong (2024) call “disinformation for hire as everyday digital labor.” They emphasize the wide range of industries, platforms, and workers involved in this digital ecosystem, and the tensions and challenges of developing, implementing, and enforcing policies and regulations concerning such activities. Lindquist and Weltevrede (2024), focusing on these governance issues, refer to this general phenomenon as the “engagement as a service” market.

To further explain the IWA, the following sections describe four dimensions of the IWA: characteristics, purposes, messages and affordances, and consequences (adapted from Chaffee, 1991; see Table 1 for a summary).

Table 1. *Dimensions and Subdimensions of the Internet Water Army*

Dimension	Subdimension	Summary
Characteristics	User demographics & activities	People with lower socio-economic status, plenty of spare time, and need for money; non-identifiable; multiple accounts & IDs at multiple locations; creating new posts rather than responding to other posts
	Organization and process	Top-down approach; includes project managers, trainers, posters/workers, PR
Purposes/ Level	Individual	Promote social influencers
	Organizational	Promote companies and products; attack adversaries or competitors
	Political	Promote government policies and ideology; frame a discussion or set an agenda; exert influence on public opinion; suppress discussion of a controversial topic or by oppositional forces
Messages and Affordances	Message content	Quantity-based messages: mass production of similar content to increase or decrease popularity of a topic
		Quality-based messages: logical expression, reasoning, credible, and persuasive to influence evaluation
	Message valence	Primarily positive
	Communication visibility	Message transparency; network translucence; user anonymity
Consequences	Desirable/undesirable	Functional/beneficial (e.g., attract public attention) or dysfunctional/harmful (e.g., diminished online trust) for different stakeholders
	Direct/ indirect	First-order response to IWA activities (e.g., gathered public opinion) or second-order result of those initial consequences (e.g., undermined participation in user-generated content)
	Anticipated/unanticipated	Expected (e.g., attack competitors) or unexpected (e.g., generate public panic)

Dimension One: Characteristics

User Demographics and Activities

The Internet Water Army consists of individual posters (workers), organized for a short time, often posting an overwhelming amount of potentially meaningless messages, operating through multiple online identities, to simulate a large number of other (regular) users, and who are paid after each task (Wu et al., 2013). The individual poster is often referred to as a “sailor” and the professional team leader the “navy head.” There is no face-to-face contact, and the identity of the individual “pusher” or sailor is typically not known (either anonymous or pseudonymous), even to the hiring PR firm or IWA agent (Kuang,

2018). IWA workers are typically activated for a given task, and dismissed afterwards, though typically return for multiple missions. The main force consists of people with lower socio-economic status, plenty of spare time, and need for money, such as migrants, housewives, college students, and even prisoners (Chen et al., 2013, 2016; Elsner, 2013). A general requirement is knowledge about and potential user experience with the target social media platform(s).

Because the initial purposes of the IWA are to attract (or distract) attention and influence opinion, IWA users attempt to create an impression on other internet users that the topics they are discussing are popular, and that their opinions strongly resonate with other users. To

accomplish this goal, they may initiate a large number of discussions by using multiple IDs to spread the same message (Nanjing Marketing Group, 2011; Xu et al., 2014), indicating that a given topic, post, or user is receiving extensive attention.

Researchers have noted some common characteristics of IWA accounts. Chen et al. (2013) collected user data from an IWA campaign and manually identified 70 potential IWA users out of 552 in total. They concluded that IWA users tend to stay active for a shorter time compared to legitimate users, often discard their IDs, and do not keep track of the conversations once they finish their tasks. The same IWA user ID may be active at different geographical locations within a short time interval, as the IWA account may be assigned by the resource team to different IWA users across their diverse geographical locations. IWA users do not build their personal networks through their IWA accounts, so there is an extremely low possibility that they have as many bilateral relationships as regular users (Wang et al., 2014). Moreover, Chen et al. (2013) reported that IWA users tend to create posts or comments rather than reply to other people's posts and comments, largely because most IWA users do not have the intention of holding conversations or exchanging opinions with others. Chen et al. (2013) estimated that 50% of IWA users post every 2.5 minutes. Moving from one topic to another and posting to different types of discussions are also more likely to be higher than for legitimate users.

Organization and Process

IWA businesses are distinctively organized and operated based on their own systems and rules. Chen et al. (2013) explain that the structure typically consists of the requester; the mission (set of tasks); and the IWA "agent", which consists of a team or project manager (there may be multiple project leaders) who coordinates a trainer, posters/workers, online

resources, and relationships with the public (see also Kim et al., 2018). The resources team is in charge of initiating and registering new IDs for the posters, and serves as human resource staff responsible for recruiting adept writers. A specific IWA mission or campaign (a set of goals and associated tasks) is initiated when a client requests a mission/campaign. The general process involves: receive request, input task and funds into system; distribute tasks (may be multiple); finish tasks; submit results (such as screenshot or links); evaluate task completion; make (typically very low) payments (Chen et al., 2013; Wang et al., 2012). More specifically, the mission is usually transferred and forwarded by a project manager, who is responsible for carrying out the whole project and leading the team. The project (or "mission") may include several teams with different responsibilities. The trainer team coordinates with the poster team (IWA workers), gives them the necessary training, and schedules tasks for specific posters. The poster team is grouped and assigned with tasks by the training team after receiving instructions on the specific tasks. There is also a quality control process conducted by the project manager in order to validate if the posts meet the numeric and quality threshold required by the clients (Chen et al., 2013; Emerging Technology, 2011). Such a top-down organizational structure facilitates IWA task completion efficiency and maximizes worker utility, as each individual worker's role is distinctly defined.

The IWA provides several different services depending on the requirements of their clients. In order to boost the digital popularity of celebrities, a PR company may pay the IWA to be fake followers and to extensively comment on, repost, or like the celebrity's (or their PR agent's) posts. The posts are designed to be seen by and propagated through online networks of people interested in some aspect of either the original or the subsequent artificial post (Wu et al., 2013). The mission may include a range of posting tasks

such as articles, Q&A sessions, and even video clips. An analysis of 2869 Zhubajie campaigns for 1280 customers on the Chinese microblog Sina Weibo included tasks to extend a sponsored message via tweets, retweets, or buying followers (Wang et al., 2012). IWA users are hired to create *fake reviews* and posts on Taobao, the Chinese version of eBay, in an attempt to promote their clients' products (Nanjing Marketing Group, 2011).

Wang et al. (2012) noted a variety of associated obstacles to the desired IWA activities: bots or search engines generating clicks on the crowdurfing site, workers using multiple accounts, and spam detector systems deleting worker posts. Once the targeted website notices the IWA activities and attempts to delete their posts, the IWA public relationship team is expected to contact the manager of the targeted website and persuade them to reverse the deletion. Establishing and maintaining a good relationship with the web manager benefits and even multiplies the exposure of the IWA posts. For example, the website may prioritize IWA comments among others under a certain post (Wang et al., 2012).

Dimension Two: Purposes

The general purposes of the IWA are to generate attention by, and influence the opinions or attitudes of, a targeted population, typically for hidden purposes, such as influencing internet users' attitudes towards a celebrity or a social or political event, consumption of a product or service, or supporting government policy and ideological frames (Chen et al., 2013; Wang et al., 2014). Other purposes include dominating online discussion forums, accomplished through large numbers of posts (Xu et al., 2014), or to influence the attitudes of internet users through high quality persuasive content (Guo et al., 2017). Purposes may occur on one or more levels: individual, organizational, and political.

Understanding IWA purposes at these levels not only contributes to more tailored detection methods, but also facilitates digital literacy training for online users to critically evaluate user-generated content online.

Individual Level

The IWA may be employed by social influencers to promote the influencers themselves or their works, which may consequently make them stand out and lead to various commercial opportunities (Pan, 2019). Celebrities, movie stars, and singers (via their agents) may hire the IWA to publicize their new movies, albums, or TV shows (Chen et al., 2013). For example, on August 2019, a Chinese idol quickly gained extensive coverage by China's national media CCTV. When he publicized his new music video on Weibo, there were roughly 100 million reposts, which constituted at that time one-third of all Weibo users (Pan, 2019). The extraordinarily large number of reposts drew the attention of several governmental officials, and he was later accused by the Communist Youth League of China on their official social media account of faking his online popularity (Sohu, 2018). Regardless, these inauthentic numbers brought him numerous commercial opportunities (Pan, 2019).

Organizational Level

Product companies may also employ the IWA to promote their products and compete with other companies. Paid online reviewers may receive professional training in writing such reviews and developing relationships with website webmasters (Liao et al., 2021). Some IWA users may pretend to be real product users and comment or write positive reviews on the product. This strategy is often applied to attract potential adopters of the products at the beginning stage of product diffusion because the early majority category of adopters relies heavily on the information from opinion leaders to evaluate whether they wish to adopt it (Rogers,

2003). Also, some companies may employ the same strategy as celebrities promoting their movies and TV shows, by initiating extensive discussion to attract attention. For example, a 2009 online article titled “Junpeng Jia, you mother asked you to go back home for dinner” (translated) posted in the online community of the computer game *World of Warcraft* attracted over 300,000 replies and more than 7 million views in only two days after it was posted (Chen et al., 2013).

Based upon a client’s request, the IWA may attempt to crowd out or delete negative reviews or comments about a certain person or company (Duan, 2010). The hiring PR company may first bribe the online media platforms such as Weibo to delete the unfavorable content. If the bribing fails, it may hire the IWA to post a massive amount of comments under a certain entry to dominate the comment section so that other online users will not be able to notice the negative comments.

Companies may also employ the IWA to attack their adversaries or competitors (Nanjing Marketing Group, 2011). By comparing similar products with different brands, companies often attempt to demonstrate the relative advantages (Rogers, 2003) of their products, services, personalities, ideology, etc. over those of their competitors to persuade customers to purchase their brand. This strategy can be convincing when companies are able to present statistical or other types of valid evidence. However, it may also be misleading, unethical, or illegal when companies intentionally post misinformation, sometimes even causing social upheaval. For example, a 2010 post on Sina Weibo that claimed that the milk powder of Yili (a company that sells milk-related products) could cause premature puberty was created and disseminated by a PR company employed by competitor Mengniu (Herold, 2015).

Political Level

A government may utilize the IWA as a tool to promote their policies and ideology, frame a discussion or set an agenda, exert influence on public opinion, or suppress discussion of a controversial topic or by oppositional forces (Wikipedia, 2020). For example, in China, 50c Party is recognized as a form of the IWA that is used only for political purposes (King et al., 2017). The main purposes are to influence public opinion, provide positive support for the government, and promote unity and stability (especially during emergency events), as well as to distract public attention from activities associated with collective action, and avoid controversy, while not explicitly censoring other content (King et al., 2017), via tens of thousands of online posters on microblogs and chatrooms (Barr, 2012). In the early 2010s, over 2 million IWA users, including prisoners (in return for reduced sentences) were reported being hired by websites, municipalities, provinces, and various government offices (Elsner, 2013).

Dimension Three: Messages and Affordances

The majority of the IWA literature emerged from computer science with an aim to create computational methods to detect IWA activities and posts, in order to be able to suppress, filter, or delete them. Sun et al. (2014) concluded that IWA detection algorithms can be categorized into two groups: *behavior-based* (user activities, such as bilateral friending ratio; Wang et al., 2014) and *content-based* (the posts) recognition (see also Chen et al., 2013). Relevant content-based work identifies two primary characteristics of IWA messages: content and valence. Both behavior- and content-based research could be grounded in an affordance approach; in particular, communication visibility (message transparency, network translucence, and user anonymity).

Message Content

Two types of posting strategies emerge from the literature: *quantity-based* messages and *quality-based* messages. *Quantity-based* messages are generated when there is a need for public attention toward a specific post or topic within a short time period (Xu et al., 2014). To achieve this goal, IWA workers may copy existing posts or comments, making slight changes, and repost them (Chen et al., 2013, 2016). Their comments or posts can sometimes be meaningful or related to the topic, while sometimes not, as their main objective is to attract attention and increase, or decrease, the popularity of the topic. As a result, similarity and relatedness are often the main criteria used to distinguish IWA messages from others (Chen et al., 2013; Sun et al., 2014). *Quality-based* messages are designed with logical expression, reasoning, credible, and persuasive (Xu et al., 2014) in order to influence the audience's evaluation and even decision making. IWA users may initiate both quantity- and quality-based content to create an impression that their reviews are popular *and* reliable (Fayazi et al., 2015).

Message Valence

The way that the IWA message conveys valence (positivity or negativity) may affect how the target audience views a specific person, product, or social event (Binder et al., 2015). Message valence is often overlooked in *quantity-based* message design. Within *quality-based* messages, the use of phrases and words that express emotions becomes important as the goal is to persuade the audience. For example, King et al. (2017) identified six types of 50c Party "cheerleading" content: "expressions of patriotism, encouragement and motivation, inspirational slogans or quotes, gratefulness, discussions of aspirational figures, cultural references, or celebrations" (p. 486). Guo et al. (2017) proposed that IWA product reviews may include as many positive emotional words as

possible. A sentiment analysis of a large number of IWA workers' and non-workers' tweets found that workers' content included less swearing, less anger, and less first person singular; overall, they were less personal (Lee, 2013).

Communication Visibility

Media affordances provide additional insight to such messages. "Media affordances are relationships among action possibilities to which agents perceive they could apply a medium, within its potential features/capabilities/constraints, relative to the agent's needs or purposes, within a given context" (Rice et al., 2017, p. 109), and include pervasiveness, editability, self-presentation, searchability, visibility, and awareness. We specifically discuss communication visibility, as it is a central affordance (Treem et al., 2020). Communication visibility includes message (content) transparency and user network (relationship) translucence, as well as user anonymity (low identifiability or low visibility). We argue that though media affordances do not constitute an IWA message per se, they provide critical boundary conditions based on which an IWA campaign may be more or less effective in achieving its goals.

Message Transparency. The very nature of most online and social media (except perhaps for those featuring only short-term, disappearing posts), is to share and make visible content, both within and across groups (boyd & Ellison, 2007). The pervasive transparency of social media messages, due to rapid dissemination, enduring storage, low cost, easy access by unintended or unknown users, searchability, links, retweets, hashtags, and various forms of notifications, facilitate the fundamental aspect and purpose of IWA activities. To the extent that different poster strategies (e.g., quantity, quality) and platform features (e.g., popular hashtags, trend monitoring) afford different levels or kinds of message transparency, IWA posters and their

clients may be more or less successful. Indeed, many content trends in Sina Weibo are due to ongoing retweets by a low percent of fraudulent accounts designed to inflate specific posts, making the content more transparent (Yu et al., 2012).

Network Translucence. Network patterns associated with crowdsourcing sites, IWA workers, followers, and threads are a central focus of IWA detection studies (Jothi & Me, 2020). In many IWA campaigns, the message content and its quality (transparency) are not the primary goal form of visibility; rather network translucence data such as message quantity and attention (the number of followers, likes/dislikes, ratings, posts, comments, retweets) are.

Crowdsourced promotion workers tend to be interconnected through a small-world network (Kim et al., 2018). Kim et al. (2018) posted two tasks promoting two YouTube videos before they were formally released, hired 28 workers for one week, and used the resulting data to determine activity, network patterns, and ratio of legitimate users to workers' followers, with effectiveness measured by number of hits to the video links. They noted that workers can falsely promote their own effectiveness by following themselves through additional self-accounts or by having co-workers follow them. Indirect measures of effectiveness include the number of tasks performed over time, the number of followers, or Klout scores (since shut down), though in their study, none of these was associated with actual effectiveness.

Lee et al. (2013) distinguished three types of crowd-turfing workers: professional, casual, and middlemen. They specifically analyzed twitter-related campaigns, comparing nearly 350,000 worker tweets to nearly 1.9 million non-worker tweets. Professionals had more followings and followers, included more links, but far fewer tweets, with well-connected networks fostering dissemination; however, they rarely communicated directly with other

users. 50c Party posts are neither random nor equally distributed, but organized into bursts of high-volume activity synchronized with major government policy statements or national events (King et al., 2017, p. 488), indicating substantial strategic governmental coordination. Liu et al. (2015) refer to the quite sparse links among IWA posters as "an unnatural social network." They analyzed 50 source nodes, a total of 9,000 nodes and 15,000 edges, over 10 days, from the massively popular Sina microblog. They distinguished professional workers, who take on tasks very frequently, from gray workers, who do so only occasionally. However, based on their overlapping membership in multiple worker cliques, the gray workers generated greater propagation than did the professional workers.

User Anonymity. Anonymity refers to the extent to which the source of message is unidentified and unacknowledged by communicators (Scott, 1998). Anonymity is a central feature or option for much computer-mediated communication, online sites, and social media (Szulc, 2019), and it lowers communication visibility. Many Chinese social media platforms such as Weibo still allow anonymity, in spite of the 2011 "Beijing Municipal Provisions on Microblog Development and Management," which outlawed anonymity and even pseudonyms (Jia & Han, 2020). In IWA posts, typically the goal is to make the worker's identity anonymous (including to the IWA agent), but to make the message and network links, both from the worker and from and among the users, as visible (transparent and translucent) as possible. Further, the invisibility of the anonymized IWA worker (including using multiple anonymous accounts) is an attempt to increase the credibility of the post(s) by portraying the content or links as generated by unique third parties rather than by the source personality or company or by a persistent pseudonym.

Dimension Four: Consequences

There is less research on the consequences or effects of IWA activities. Using Rogers' (2003) diffusion of innovations framework, we define the consequences of IWA activities as effects on an individual, group, organization, or social system as a result of those activities, with three polar characteristics: desirable/undesirable, direct/indirect, and anticipated/unanticipated. These may be salient individually or in combination. Further, consequences may be culturally contextual.

Desirable/undesirable Consequences

This type of consequence depends on whether the effects of IWA activities (e.g., quantity and quality of online posts) are functional/beneficial or dysfunctional/harmful, for different stakeholders. For example, companies may hire IWA users to promote their movie and tag it as “worth watching”, and thus people may be attracted to buy tickets (Chen et al., 2016). This consequence may be perceived as desirable by the company and its media client, but, depending on their experience, undesirable by the customers. Another undesirable consequence is diminished online trust and information credibility (Chen et al., 2013, 2016). Vote-spamming distorts shared information and discourse outcomes (Zhang & Labiod, 2019), contributing to cyber-pessimist views of the civic consequences of new media (Zhao, 2014). Massive criticism of a particular position can generate a *spiral of silence*, crowding out discussion, minimizing minority opinions, and maximizing a false perception of the majority view (Noelle-Neumann, 1993).

Kwon et al. (2022) identified two complementary individual and social consequences of a computational propaganda operation (the 2018 South Korean “Druking” case). Their analysis of 1389 fake accounts generating over 56,000 engagements designed to significantly alter top-ranked posting visibility, and comments by over

45,000 users, on an online news site within a three-hour period, showed that bot-assisted flooding of anti-government opinions counter to those in more general online media, represented two related effects. First was the false *amplification* of opinions by the fake posters; second was the resulting *diminution* of users' political commenting, representing an artificial spiral of silence. These both lead to what the authors refer to as “a distorted opinion environment.”

Direct/indirect Consequences

This type of consequence depends on whether the effects occur as a first-order response to IWA activities or as a second-order result of those initial consequences. For example, a direct consequence can be that public attention is gathered and public opinion is affected (Xu et al., 2014), while an indirect consequence can be lowered authenticity and general perceived value of information in online discussions (Guo et al., 2017). Manipulating public opinion may undermine participation and trust in cyber collective actions such as crowdsourcing (Estellés-Arolas & González-Ladrón-De-Guevara, 2012; Zeng et al., 2014). The low level of relevance of some IWA content to the referred earlier posts may confuse general users and lower their level of perceived interactivity and their impetus to create their own user-generated content (Carr & Hayes, 2015). Official associations such as the Public Relations and Communication Association, and the International Communications Consultancy, have attempted to reinforce principles against such unethical practices, which harm the reputation of legitimate PR firms (Silverman et al., 2020; Jiaye, 2021), yet PR companies may feel they have to engage in those activities in order to stay competitive (Wu et al., 2013).

Anticipated/unanticipated Consequences

This type of consequence depends on whether the effects are intended or expected by the

innovation promoters or members of a social system. As an anticipated marketing strategy, companies may hire the IWA to attack their competitors or spread rumors (Sun et al., 2014), yet generate unanticipated consequences. As an example, the fake news that Yili's milk powder is detrimental to health that Mengniu disseminated caused extreme panic and antipathy at the national level (Nanjing Marketing Group, 2011). Although the anticipated outcome of Mengniu was that potential customers would switch their purchases from Yili to Mengniu, the unanticipated outcomes were more critical, social, and legal. The contradictions and tensions associated with paid online posting in China are to some extent a result of ambiguous Chinese regulation stemming from emphasizing both political control and economic growth, and the facilitation of digital communicative labor (Han, 2018). For example, in China, *Internet mercenaries* are not legally recognized, but they are also not legally prohibited (Wu et al., 2013).

Combinations

As with innovations in general (Rogers, 2003), the different types of IWA consequences typically occur in combination (such as undesirable, indirect, and unanticipated). Low (2018) reported that IWA companies may post inappropriate content such as pornography in an attempt to trigger deletions of message threads or discussions, but that may diminish the level of civility and undermine the online communication environment. The negative impacts of the IWA on the quality of online discussion have received attention from the Chinese government. As a legislative reaction, a Cybersecurity Law was introduced by the Chinese government in June 2017, which put 200 suspects under arrest and deleted more than 5000 accounts (Low, 2018). While one perspective on this transnational economy is to emphasize the unregulated and harmful nature of the system as well as the ethics of the sweatshop aspect of digital

click workers, Lindquist (2018) instead focuses on the reality, economics, and motivations of click farmers (specifically, in Indonesia). Thus, Lindquist positions this digital and personal capitalist ecosystem as multi-consequential: dynamic, evolving, multiple, and interdependent patron-client agreements requiring trust but also frequent turnover and repair. Similarly, Ong and Cabañes (2019) situate the generation of digital disinformation by anonymous paid click workers as part of production culture, akin to highly competitive piece work, short-term projects, and freelance gig work.

THEORETICAL APPROACHES AND FUTURE RESEARCH

The IWA and the affiliated actors, content, networks, and consequences offer a wide array of research foci, complementing the already extensive work by computer and information scientists on categorizing and detecting IWA users and content. We propose four communication approaches that may further the study of this phenomenon.

Framing and Agenda-Setting

Message framing is the way that communicators (e.g., IWA agents or users) design the message to emphasize or deemphasize aspects of an issue or an event, in order to shape audience attention to and interpretations of that message (Binder et al., 2015). At the individual level, future research may continue to study what frames influence how social media users evaluate the credibility and trustworthiness of an IWA message, especially among quality-based messages (Xu et al., 2014).

At the macro-level, Kuang (2018) applied agenda-setting theory and spiral of silence theory to understand the IWA process. For a given topic, the IWA attempts to move the topic to, or higher up on, the agenda of a specific public, focusing attention on praise or condemnation through

massively repeated content. It is therefore important to gain a better understanding of the uses and effects of IWA campaigns at the organizational and political levels. Novel approaches to measuring such effects, such as computational tools, may be especially helpful.

Affordances

As noted above, digital media provide numerous features or affordances to leverage IWA efforts, such as copy-and-pasting, reposting, retweeting, hashtags. In turn, this process may generate the spiral of silence effect, whereby ordinary people quickly come to feel their opinions are in the minority or deviant, and thus refrain from counter-posting. Therefore, future research may test how media affordances such as visibility and anonymity facilitate or constrain the intended effects of IWA messages, while at the same time, enable other undesirable, indirect, or unanticipated consequences.

Diffusion Process

Chen et al. (2003) developed the aging theory of online bursting topics to study the life cycle of online discussion topics, involving five stages: latency, emergence, evolution, recession, and death. Xu et al. (2014) argued that most IWA messages are sourced during the latency and early evolution stages, in which public attention starts to be aroused. The persuasive effects of the IWA might be strongest during this time because the audience is receiving information and forming their personal attitudes. They also suggested that IWA users may initiate their activities with quantity-based messages to gather attention and then engage in quality-based messages to increase their persuasiveness. Yet, few studies specify the stages during which a specific IWA message type is utilized.

The sequenced development of online discussion topics and public attitudes through IWA activities can also be studied from the perspective of diffusion of innovations theory

(Rogers, 2003). This theory identified five stages through which an innovation is diffused and leads to a decision of adoption, reinvention, discontinuance, or rejection: knowledge, persuasion, decision, implementation, and confirmation. During the knowledge stage, the audience is exposed to the basic information about the innovation. Then, during the persuasion stage, they start to form a personal attitude toward the innovation by receiving other people's evaluation information. From this perspective, IWA users may start posting the discussion topics and increasing their digital popularity with quantity-based messages during the knowledge stage. Then, they may apply quality-based valenced messages during the knowledge stage, in which the target audience starts actively seeking other people's opinions, retweeting posts, and shaping their own attitudes. The diffusion of innovations concept of critical mass is useful to focus on when online links (followers, reposts) become viral. Studying IWA strategies during each diffusion stage can help facilitate development of interventions that detect and label potential IWA content.

Cultural Contexts

In the case of IWA, a few Chinese cultural and political issues are particularly relevant. Official Chinese discourse about Internet governance emphasizes personal security, social stability, and moral goodness. So, a large number of messages from IWA workers reinforcing this discourse, attempting to reduce the possible "destabilizing field of contentions" (Cui & Wu, 2016, p. 265), resonates with many online users in China as well as the Communist Party. Kuang (2018) also notes that most Chinese users expect a stable social context, that governmental (especially local) units deal with problems, and to feel that their opinions are valued. The Chinese government seeks to avoid the incursion of external ideologies, while maintaining control, through shaping Internet technology use and

features, as well as online dialog (Wang, 2013). Thus, practices such as 50c Party and the IWA may serve the purpose of focusing, congealing, and maintaining attitudes. Future research, especially in the political arena, may study how the phenomenon of IWA-type behavior intertwines with the Chinese practice of *weiguan* (围观), or mass gatherings for public spectacles, applied to online political participation with virtual crowds expressing public opinions and complaints about offline controversies, both to experience expression but also to generate pressure for resolution: that is, online *weiguan* (网络围观) (Xu, 2015).

CONCLUSION

Transformative online and digital media are being used in constantly expanding ways, for better and for worse. Researchers need to similarly continuously identify and define new types of uses, users, and consequences. This review articulated the domain and four dimensions (adapted from the explication categories in Chaffee, 1991) of the “Internet Water Army”, and distinguished it from other similar general and specific phenomena, drawing upon literature across disciplines. In particular, it applied a typology of consequences (Rogers, 2003) to illustrate some implications of the IWA. This framework can serve as a guide for scholars advancing research on the Internet Water Army and similar practices in the broad domain of digital disinformation and engagement as service.

REFERENCES

- Barr, M. (2012). Nation branding as nation building: China’s image campaign. *East Asia*, 29(1), 81–94.
- Binder, M., Childers, M., & Johnson, N. (2015). Campaigns and the mitigation of framing effects on voting behavior: A natural and field experiment. *Political Behavior*, 37, 703–722. <https://doi.org/10.1007/s11109-014-9292-2>
- boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Bradshaw, M., & Howard, P. N. (2017). *Troops, trolls and troublemakers: A global inventory of organized social media manipulation*. Working Paper 2017.12. University of Oxford. <https://fpmag.net/wp-content/uploads/2017/11/Troops-Trolls-and-Troublemakers.pdf>
- Bu, Z., Xia, Z., & Wang, J. (2013). A sock puppet detection algorithm on virtual spaces. *Knowledge-Based Systems*, 37, 366–377. <https://doi.org/10.1016/j.knosys.2012.08.016>
- Carr, C., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. *Atlantic Journal of Communication*, 23, 46–65. <https://doi.org/10.1080/15456870.2015.972282>
- Chaffee, S. H. (1991). *Explication*. Sage.
- Chen, C., Wu, K., Srinivasan, V., & Zhang, X. (2013, August). Battling the internet water army: Detection of hidden paid posters. *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)* (pp. 116–120). IEEE.
- Chen, G., Cai, W., Huang, J., & Jiao, X. (2016, June). Uncovering and characterizing Internet Water Army in online forums. *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)* (pp. 169–178). IEEE.
- Cui, D., & Wu, F. (2016). Moral goodness and social orderliness: An analysis of the official media discourse about Internet governance in China. *Telecommunications Policy*, 40(2-3), 265–276. <https://doi.org/10.1016/j.telpol.2015.11.010>
- Darwish, K., Alexandrov, D., Nakov, P., &

- Mejova, Y. (2017, September). Seminar users in the Arabic Twitter sphere. In T. Yasseri, A. Mashhadi, & G. L. Ciampaglia (Eds.), *Social Informatics* (pp. 91–108). Springer.
- Dhanesh, G. S., & Duthler, G. (2019). Relationship management through social media influencers: Effects of followers' awareness of paid endorsement. *Public Relations Review*, 45(3), 101765.
<https://doi.org/10.1016/j.pubrev.2019.03.002>
- Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the World-Wide Web. *Communications of the ACM*, 54(4), 86–96.
<https://doi.org/10.1145/1924421.1924442>
- Duan, Y. (2010, June 17). *The invisible hands behind Web postings*. China Daily.
http://www.chinadaily.com.cn/china/2010-06/17/content_9981056.htm
- Elsner, K. (2013, November 27). *China uses an army of sock puppets to control public opinion - and the US will too*. Guardian Liberty Voice.
<https://guardianlv.com/2013/11/china-uses-an-army-of-sockpuppets-to-control-public-opinion-and-the-us-will-too/>
- Emerging Technology from the arXiv. (2011, November 22). *Undercover researchers expose Chinese Internet water army*. MIT Technology Review.
<https://www.technologyreview.com/s/426174/undercover-researchers-expose-chinese-internet-water-army/>
- Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200.
<https://doi.org/10.1177/0165551512437638>
- Fayazi, A., Lee, K., Caverlee, J., & Squicciarini, A. (2015, August). Uncovering crowdsourced manipulation of online reviews. *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 233–242). ACM.
- Fielding, N., & Cobain, I. (2011, March 17). *Revealed: US spy operation that manipulates social media*. The Guardian.
<https://www.theguardian.com/technology/2011/mar/17/us-spy-operation-social-networks>
- Freberg, K., Graham, K., McGaughey, K., & Freberg, L. A. (2011). Who are the social media influencers? A study of public perceptions of personality. *Public Relations Review*, 37, 90–92.
<https://doi.org/10.1016/j.pubrev.2010.11.001>
- Grohmann, R., & Corpus Ong, J. (2024). Disinformation-for-hire as everyday digital labor: Introduction to the special issue. *Social Media + Society*, 10(1).
<https://doi.org/10.1177/20563051231224723>
- Guo, B., & Jiang, Z. B. (2023). What is the Internet water army? A practical feature-based detection of large-scale fake reviews. *Mobile Information Systems*.
<https://doi.org/10.1155/2023/2565020>
- Guo, B., Wang, H., Yu, Z., & Sun, Y. (2017, July). Detecting the Internet water army via comprehensive behavioral features using large-scale e-commerce reviews. *2017 International conference on computer, information and telecommunication systems (CITS)* (pp. 88–92). IEEE.
- Han, D. (2018). Paid posting in Chinese cyberspace: Commodification and regulation. *Television & New Media*, 19(2), 95–111.
<https://doi.org/10.1177/1527476417701994>
- Hassan, I., & Hitchen, J. (2019, April 18). *Nigeria's propaganda secretaries*. Mail & Guardian.
<https://mg.co.za/article/2019-04-18-00-nigerias-propaganda-secretaries/>
- Herold, D. K. (2015). Whisper campaigns: Market risks through online rumours on the Chinese Internet. *China Journal of Social Work*, 8(3), 269–283.
<https://doi.org/10.1080/17525098.2015.1082493>
- Jia, L., & Han, X. (2020). Tracing Weibo (2009–2019): The commercial dissolution of public

- communication and changing politics. *Internet Histories*, 4(3), 304–332.
<https://doi.org/10.1080/24701475.2020.1769894>
- Jiaze, X. (2021, November). Analysis on the influence of water army on public opinion on the Internet. *2021 3rd International Conference on Literature, Art and Human Development (ICLAHD 2021)* (pp. 718-722). Atlantis Press.
- Jothi, D. H., & Me, T. R. (2020, January). Enhanced detection of Internet water army based on supernetwork theory. *2020 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1–6). IEEE.
- Kalén, A., Bisagno, E., Musculus, L., Raab, M., Pérez-Ferreirós, A., Williams, A. M., Williams, A. M., Araújo, D., Lindwall, M., & Ivarsson, A. (2021). The role of domain-specific and domain-general cognitive functions and skills in sports performance: A meta-analysis. *Psychological Bulletin*, 147(12), 1290–1308.
<https://doi.org/10.1037/bul0000355>
- Karpan, A. (Ed.) (2018). *Troll factories: Russia's web brigades*. Greenhaven Publishing LLC.
- Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of social media influencers. *Celebrity Studies*, 8(2), 191–208.
<https://doi.org/10.1080/19392397.2016.1218292>
- Kim, H. J., Lee, J., Chae, D. K., & Kim, S. W. (2018). Crowdsourced promotions in doubt: Analyzing effective crowdsourced promotions. *Information Sciences*, 432, 185–198.
<https://doi.org/10.1016/j.ins.2017.12.004>
- King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *American Political Science Review*, 111(3), 484–501.
<https://doi.org/10.1017/S0003055417000144>
- Kuang, W. (2018). Network forums. In W. Kuang (Ed.), *Social media in China* (pp. 101-122). Palgrave Macmillan.
- Kumar, S., Cheng, J., Leskovec, J., & Subrahmanian, V. S. (2017, April). An army of me: Sockpuppets in online discussion communities. *Proceedings of the 26th International Conference on World Wide Web* (pp. 857–866).
- Kwon, K. H., Lee, M. H., Han, S. P., & Park, S. (2022). Fake thumbs in play: A large-scale exploration of false amplification and false diminution in online news comment spaces. *New Media & Society*, 26(6), 3252–3272.
<https://doi.org/10.1177/14614448221099170>
- Lawton, T. (2011, May 10). *The online water army: How businesses deceive Chinese Internet users*.
<https://nanjingmarketinggroup.com/blog/online-water-army-how-businesses-deceive-chinese-internet-users>
- Lee, K., Tamilarasan, P., & Caverlee, J. (2013, June). Crowdturfers, campaigns, and social media: Tracking and revealing crowdsourced manipulation of social media. *Seventh International AAAI Conference on Weblogs and Social Media* (pp. 331–340).
- Liao, H. L., Huang, Z. Y., & Liu, S. H. (2021). The effects of negative online reviews on consumer perception, attitude and purchase intention: Experimental investigation of the amount, quality, and presentation order of eWOM. *Transactions on Asian and Low-Resource Language Information Processing*, 20(3), 1–21.
- Lindquist, J. (2018). Illicit economies of the internet: Click farming in Indonesia and beyond. *Made in China*, 4(3), 88–92.
- Lindquist, J., & Weltevrede, E. (2024). Authenticity governance and the market for social media engagements: The shaping of disinformation at the peripheries of platform ecosystems. *Social Media + Society*, 10(1).
<https://doi.org/10.1177/20563051231224721>
- Liu, W., Cao, Y., Li, D., Niu, W., Tan, J., Hu, Y., & Guo, L. (2015, November). Structural analysis of IWA social network. *International Conference on Applications and Techniques in Information Security* (pp. 141–152). Springer.

- Low, Z. (2018, December 13). *Chinese police shut down 'water army' of Internet trolls paid US\$4.3 million to blitz websites and social media*. South China Morning Post.
<https://www.scmp.com/news/china/society/article/2177800/chinese-police-shut-down-water-army-internet-trolls-paid-43>
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293–334.
<https://doi.org/10.2307/23044045>
- Mak, T. (2021, August 18). *The FBI keeps using clues from volunteer sleuths to find the Jan. 6 Capitol rioters*. National Public Radio.
<https://www.npr.org/2021/08/18/1028527768/the-fbi-keeps-using-clues-from-volunteer-sleuths-to-find-the-jan-6-capitol-riote>
- Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220.
<https://doi.org/10.1016/j.pragma.2013.07.012>
- Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021). Fake news is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, 65(2), 180–212.
<https://doi.org/10.1177/0002764219878224>
- Mozur, P., Bradsher, K., Liu, J., & Krolik, A. (2024, February 22). *Leaked files show the secret world of China's hackers for hire*. The New York Times.
<https://www.nytimes.com/2024/02/22/business/china-leaked-files.html>
- Neyazi, T. A. (2020). Digital propaganda, political bots and polarized politics in India. *Asian Journal of Communication*, 30(2), 39–57.
<https://doi.org/10.1080/01292986.2019.1699938>
- Noelle-Neumann, E. (1993). *The spiral of silence: Public opinion, our social skin*. University of Chicago Press.
- Ong, J. C., & Cabañes, J. V. A. (2019). When disinformation studies meets production studies: Social identities and moral justifications in the political trolling industry. *International Journal of Communication*, 13, 20, 5771–5790.
<https://ijoc.org/index.php/ijoc/article/view/11417>
- Pan, Y. (2019, February 26). *China's state media investigates fake followers of celebrities*. Jing Daily.
<https://jingdaily.com/fake-followers-celebrities/>
- Plucker, J. A., & Beghetto, R. A. (2004). Why creativity is domain general, why it looks domain specific, and why the distinction does not matter. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 153–167). American Psychological Association.
<https://doi.org/10.1037/10692-009>
- Rheingold, H. (2002). *Smart mobs: The next social revolution*. Perseus Publishing.
- Rice, R. E., Evans, S. K., Pearce, K. E., Sivunen, A., Vitak, J., & Treem, J. W. (2017). Organizational media affordances: Operationalization and associations with media use. *Journal of Communication*, 67, 106–130.
<https://doi.org/10.1111/jcom.12273>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Scott, C. R. (1998). To reveal or not to reveal: A theoretical model of anonymous communication. *Communication Theory*, 8, 381–407.
<https://doi.org/10.1111/j.1468-2885.1998.tb00226.x>
- Serabian, R., & Foster, L. (2021, September 8). *Pro-PRC influence campaign expands to dozens of social media platforms, websites, and forums in at least seven languages attempted to physically mobilize protesters in the U.S.* Threat Research Blog.
<https://www.fireeye.com/blog/>

- threat-research/2021/09/pro-prc-influence-campaign-social-media-websites-forums.html [no longer accessible]
- Seymour, R. (2011, September 16). *The Johann Hari debacle*. The Guardian.
<https://www.theguardian.com/commentisfree/2011/sep/16/johann-hari-debacle>
- Silverman, C., Lytyvenko, J., & Kung, W. (2020, January 6). Disinformation for hire: How a new breed of PR firms is selling lies online. *Buzzfeed News*.
<https://www.buzzfeednews.com/article/craigsilverman/disinformation-for-hire-black-pr-firms>
- Singer, P. W., & Brooking, E. T. (2018). *LikeWar: The weaponization of social media*. Houghton Mifflin.
- Smith, B. G., Kendall, M. C., Knighton, D., & Wright, T. (2018). Rise of the brand ambassador: Social stake, corporate social responsibility and influence among the social media influencers. *Communication Management Review*, 3(1), 6–29.
<https://doi.org/10.22522/cm20180127>
- Stukal, D., Sanovich, S., Bonneau, R., & Tucker, J. A. (2017). Detecting bots on Russian political Twitter. *Big Data*, 5(4), 310–324.
<https://doi.org/10.1089/big.2017.0038>
- Sun, W., Zhao, W., Niu, W., & Chang, L. (2014, October). A DBN-based classifying approach to discover the Internet water army. *Proceedings of the 2014 3rd International Conference on Intelligent Information Processing* (pp. 78–89). Springer.
- Zsulc, L. (2019). Profiles, identities, data: Making abundant and anchored selves in a platform society. *Communication Theory*, 29, 257–276.
<https://doi.org/10.1093/ct/qtz031>
- Treem, J. W., Leonardi, P. M., & van den Hooff, B. (2020). Computer-mediated communication in the age of communication visibility. *Journal of Computer-Mediated Communication*, 25(1), 44–59.
<https://doi.org/10.1093/jcmc/zmz024>
- Wang, A. (2013). Analysis of the enhancement of the attraction of China's mainstream ideology in Internet media. *Journal of Dalian University of Technology (Social Sciences)*, 34(4), 103–107.
- Wang, G., Wilson, C., Zhao, X., Zhu, Y., Mohanlal, M., Zheng, H., & Zhao, B. Y. (2012, April). Serf and turf: Crowdturfing for fun and profit. *Proceedings of the 21st International Conference on World Wide Web* (pp. 679–688).
- Wang, K., Xiao, Y., & Xiao, Z. (2014, January). Detection of internet water army in social network. Paper presented at the *International Conference on Computer, Communications and Information Technology*, Beijing, China.
- Wikipedia. (2020, March 10). *Internet water army*.
https://en.wikipedia.org/wiki/Internet_Water_Army
- Wu, M., Jakubowicz, P., & Cao, C. (Eds.) (2013). *Internet mercenaries and viral marketing: The case of Chinese social media*. IGI Global.
- Xu, H., Cai, W., & Chen, G. (2014, August). Forum-oriented research on water army detection for bursty topics. *2014 9th IEEE International Conference on Networking, Architecture, and Storage* (pp. 78–82). IEEE.
- Xu, J. (2015). Online Weiguan in Web 2.0 China: Historical origins, characteristics, platforms and consequences. In G. Yang (Ed.), *China's contested Internet* (pp. 257–282). NIAS Press.
- Yu, L. L., Asur, S., & Huberman, B. A. (2012, September). Artificial inflation: The real story of trends and trend-setters in Sina Weibo. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (pp. 514–519). IEEE.
- Zeng, K., Wang, X., Zhang, Q., Zhang, X., & Wang, F. Y. (2014). Behavior modeling of Internet water army in online forums. *IFAC Proceedings Volumes*, 47, 9858–9863.
- Zhang, J., & Labiod, H. (2019, July). Retrieve the hidden leaves in the forest: Prevent voting spamming in Zhihu. *International symposium on security and privacy in social networks and big data*



(pp. 165–180). Springer.

Zhao, Y. (2014). New media and democracy: 3 competing visions from cyber-optimism and cyber-pessimism. *Journal of Political Sciences & Public Affairs*, 2(1), 1–4.
<https://www.omicsonline.org/open-access/new-media-and-democracy-competing-visions-from-cyberoptimism-and-cyberpessimism-2332-0761.1000114.pdf>