"This Is the Future of Advertising!" Or Is It? New Insights into the Justifiability of Deceptive Crowdwork in Cyberspace

Santtu Kauppila
University of Jyvaskyla
santtu.kauppila@gmail.com

Wael Soliman
University of Jyvaskyla
wael.soliman@jyu.fi

Abstract

Unlike classical forms of deception where the deceiver deceives their victims directly, the crowdsourcing of cyber deception provides a powerful and cost-effective mechanism for deceivers to create and spread falsehood from the shadows. But for a mass deception campaign to be effective, the crowdworkers must rationalize (and willingly accept) their role in the deceptive act. What, then, could justify participation in a mass-deception campaign? To answer this question, we adopt the qualitative vignette approach and utilize neutralization theory as our guiding lens. Our results point to several neutralization techniques that crowdworkers could invoke to convincingly rationalize involvement in a cyber deception campaign. Importantly, the findings shed new light on a growing pessimism about work ethics in cyberspace which may lead some ordinary people into joining deception campaigns, believing it to be the future of advertising. We discuss the theoretical and practical implications of these novel insights.

1. Introduction

When the news broke that "Samsung has just been slapped with a $340K fine for paying people to slam HTC in online forums," [1] many of us probably did not pay it enough attention, since the digital space is littered with anonymous opportunities to deceive and hurt others. But, who would have thought that a corporation as big as Samsung would take part in a mass deception campaign, orchestrated by a marketing agency and executed by ordinary students [1]–[4]? Cyber mass deception campaigns such as this (aka, crowdturfing, cyber-turfing, and online astroturfing) are by no means a rare occurrence. The tactic is gathering widespread use in both commerce [5]–[7] as well as in politics [8]–[10] to score quick victories. Alarmingly, research has shown that online falsehood travels "significantly farther, faster, deeper, and more broadly than the truth" [11, p. 1146], and it continues to shape people's beliefs and attitudes toward a given subject even after the false content has been debunked [12]. As such, opinions that are formed, and decisions that are made based on misleading information can have severe ramifications on the well-being of individuals, organizations, and societies at large [11]–[14].

From a cybersecurity standpoint, cyber deception is an attack on the accuracy of information circulating online [15]–[17]. What makes these disinformation attacks an exceptionally grim threat to handle is their subtle and unobtrusive nature. Various methods of detection have been utilized to detect and curb these mass-deception campaigns [18]–[20], but the people involved quickly learn how the detection algorithms work and adapt their behavior to circumvent future detection, creating a constant arms-race [21]. In addition, the most glaring problem with this countermeasure is that it typically is a late measure in an optimal security action plan [22]. Since online disinformation spreads quickly and influences deeply, by the time it is detected and removed, its goals may already have been achieved. Therefore, the security efforts have been trying to focus on pre-kinetic events, or the stages which temporally occur prior to an attack, such as the thought processes that facilitate or dissuade potential offenders from participating in cybercrime in the first place [23].

This article addresses a central question in this battle against cyber deception: what rationalizations could justify participating in the crowdsourcing of cyber deception? To answer this question, we adopt the qualitative vignette method and use neutralization theory [24]–[27] as our guiding theoretical framework. As will be revealed later, the work points our attention to five exceptionally convincing techniques deeming cyber deception justifiable. Three of these techniques are classical and widely known in IS research, namely, 'appeal to higher loyalty', 'denial of responsibility', and 'denial of injury'. The other two techniques have emerged from our analysis of the interview data. We call them the 'appeal to professionalism', and 'appeal to normative fluidity'.

2. Background: the evolution of deception

Deception is not a novel concern, and scholars have studied this behavior from various fields, including

URI: https://hdl.handle.net/10125/79684
978-0-9981331-5-7

psychology [28], criminology [29], business and organization [30], and more recently, information systems [31]–[33]. Broadly speaking, deception "implies that an agent acts or speaks so as to induce a false belief in a target or victim." [28, p. 133]. More strictly, it involves an interaction between two parties: a deceiver and their target. The deceiver's aim is to "manipulate the environment of the other party, the target, so as to intentionally foster an incorrect cognitive representation of the target's situation and instigate a desired action, one the target would be unlikely to take without the manipulation" [32, p. 95].

Two generations of deception are worth noting: classical (or traditional) deception, and Internet (or electronically mediated) deception. Classical deception is traced back to the art of conjuring, where "the successful conjurer guides the thoughts of the onlookers to the desired conclusions" [28, p. 137]. A popular example of classical deception is when a deceiver convinces their target-victim that they are buying a genuine-brand watch when in reality it is a counterfeit. Whereas this classical form of deception may involve physical interaction between the deceiver and the target (e.g., a street encounter); Internet deception, especially in its early days, utilizes the Internet as a communication medium between the deceiver and their victim. For example, instead of deceiving the victim into buying a counterfeit watch in a street exchange, the Internet made it possible for the deceiver to deceive Internet shoppers (e.g., in an email exchange or on an e-commerce site) without the need to meet them face-to-face [31], [32].

What is common to both generations (classical and Internet-mediated deception) is that there is always one degree of separation between the deceiver and their victim. In other words, the deceiver, be it a business or an individual, is always in direct interaction with the target-victim, whether in a street or in an email. Today, we are witnessing a new generation of (cyber) deception where the deceiver is utilizing recent developments in crowd-based technologies and business models, such as crowdsourcing and gig economy [34]–[39], to add two or more degrees of separation between them (the deceiver) and their targeted victims. Indeed, with the abundance of available solutions, deceivers may be asking themselves: why do the work yourself if you can hire the masses to do it for you?

Recent high-profile mass deception examples in both business [5]–[7] and politics [8]–[10] attest to the growing trend of deceivers utilizing crowdsourcing to spread misleading information and induce false beliefs among their targets/victims [28]. Following the conventional crowdsourcing platform model [34]–[37], [40], the orchestration of work requires concerted efforts among three partners: (a) requesters, those who order and pay for the gigs or tasks; (b) workers, those who execute the task for a fee; and (c) the crowdsourcing organizers, those who organize and orchestrate the exchange between requesters and workers [41]. In the same line, crowdturfing is a typical two-sided market [40], [42]–[44], where disinformation gigs are delegated to a general working populace who willingly participate in the process, receiving incentives in return [41], [45]. Unlike typical services traded on crowdsourcing platforms (such as clickwork, logo designs, image labeling, etc. [34]–[37]), the transacted service here is 'deception' primarily by means of false information. This includes fake reviews, misleading comments, and false impressions, to name a few [5], [7], [41], [46], [47]. Leveraging crowdsourcing on digital platforms enables deceivers to achieve a massive scale and rapid spread of disinformation, which sets it apart from traditional, less sophisticated forms of deception.

The disinformation acts done by the crowdworkers are generally anything a normal user is capable of, as the worker's intent is to pretend to be a typical user. These include, for example, posting, commenting, liking/upvoting, writing reviews and sharing links. Of course, the crowdsourcing of deception is only one of various forms of deception that exist in cyberspace. For instance, there is considerable research on deceptive practices that utilize automated bots, or accounts that appear to belong to humans when in reality they are operated fully or partially by a program [48]–[50]. Notwithstanding the importance of these various forms of cyber deception, the main focus of this article is on the justifiability of deceptive practices by the human worker.

In sum, most current work on cyber deception has characterized the network of actors who participate into the business, as well as developed technical methods of detecting the attacks after they have been executed (i.e., after the fact). However, the pre-kinetic events of the phenomenon have received less attention, and thus approaches that focus on the motivational and behavioral aspects of the problem have been called for [51]. In addition, prior work has examined the types of services offered by black-hat marketplaces [41], [52] but thus far, little attention has been given to how the crowd workers justify their involvement in taking part in deceptive work on behalf of the deceiver.

3. Neutralization theory

Neutralization theory [24]–[27] is one of the most popular and influential explanations of deviant behavior. The original work on neutralization theory dates back to the 1950s: Sykes and Matza challenged the dominant perspective at that time, which assumed that delinquents adopt a completely inverted moral code that runs contrary to conventional morality [53]. Sykes and

Matza [24] suggested that this perspective (often called the sub-culture view) was rather simplistic, especially in that it rested upon the strong assumption that there is a clear demarcation between the two worlds: that of the law-abiding citizen, and that of the delinquent [27]. If that was indeed the case, Sykes and Matza argued, then there would be no reason to believe that crime brings shame or guilt to those who commit it. However, "many delinquents *do* experience a sense of guilt or shame, and … [these feelings should not] be dismissed as a purely manipulative gesture to appease those in authority" [24], pp. 664-5, *emphasis* original].

Sykes and Matza [24] proposed that many delinquents may have devised various techniques that enabled them to *drift* between the two worlds: that of crime and that of social order [27]. Their work, later to be known as neutralization theory, provides an explanation for how early-stage criminals cope with the dilemma of committing crime yet respecting social order. Specifically, neutralization theory assumes that delinquents are "partially committed to the dominant social order" [24, p. 666], and that this partial commitment is what drives them to utilize different neutralization techniques that dull the feeling of shame that comes with engagement in anti-social, deviant behavior. As Maruna and Copes [25] point out, neutralizations' main purpose is to free "the offenders from the moral binds that prevent offending" [p. 298].

Sykes and Matza's [24] work resulted in the identification of five popular guilt-neutralizing mechanisms, which they called techniques of neutralization. These are 'denial of responsibility', 'denial of injury', 'denial of victim', 'condemnation of condemners', and 'appeal to higher loyalties'. More details on these techniques will be provided in the findings section. Each of these techniques provides a counterargument to justify the temporary deviant behavior from the offender's point of view. Later research in criminology has unearthed further techniques that offenders could deploy to neutralize their anti-social behavior in different contexts. These include the metaphor of the ledger [54] defense of the necessity [55], and claim of entitlement [56]. For a more complete list of techniques see Maruna and Cope's [25] work in criminology and Willison et al.'s [23] work in information systems.

In the IS field, neutralization theory has been used to explain deviant behavior in two main research themes. The first theme covers workplace deviance committed by employees. Two distinct topics have dominated this area of research: one explores the justifiability of crimes committed by employees, such as hacking and stealing [57], [58]; and the other explores non-criminal violations of organizational policy, such as cyberloafing [59], [60] and shadow IT use [61], [62]. The second research theme explores the justifiability of software and music piracy, [63]–[65] covering deviant behavior outside organizational context.

To our best knowledge, the neutralization theory lens has never been used to explore the justifiability of mass deception crowdwork, the focus of this article.

4. Research approach

The crowdsourcing of disinformation is a secretive mode of operation in cyberspace, and therefore it is a challenging feat to identify, let alone invite, participants who are willing to share their experiences. Things get more complicated when the research question delves into self-reflection on a sensitive subject such as the justifiability of an unethical behavior. When first-hand experiences are not readily available, the vignette research approach provides an optimal alternative for empirical data collection [66]–[69].

In principle, the vignette is a "technique used in structured and depth interviews as well as focus groups, providing sketches of fictional (or fictionalized) scenarios. The respondent is then invited to imagine, drawing on his or her own experience, how the central character in the scenario will behave" [68, pp. 175-176]. This approach allows for a less confrontational discussion on sensitive topics, such as crowdturfing, without pressuring the interviewees to answer on unethical matters on a personal level. Vignette research designs are suitable to both qualitative and quantitative approaches [66]–[69]. In the IS field, however, vignettes have been mostly used in quantitative studies [26], [70], [71], where research participants are presented with one or more scenarios and then asked to respond to a range of measurements using predetermined scales. By contrast, in the qualitative vignette method both the interviewer and the interviewee engage in a conversation about a topic of interest, typically reflected by the fictionalized scenario, in an attempt to access the participants' own 'stock of knowledge' [68] and co-construct an explanation that makes sense of the scenario at hand.

In the current study, the vignette is formulated to resemble as much as possible the experience of an individual crowdworker taking part in the *Samsung vs. HTC* mass deception campaign [1]–[4]. In this campaign, Samsung's appointed agent company hired ordinary university students to do their bidding, hence we believe that conducting the research among of students suits the purpose of this study. The vignette itself is divided into four main parts. The first part gives a brief background on the narrative's protagonist, a student called Paul, and some of his interests. The second part introduces a digital marketing campaign associated with the release of a new mobile phone, for

which Paul is invited to participate for a decent amount of money. The third part describes some of the gigs that Paul is required to complete before he can receive his paycheck. These gigs include, among other things, fabricating positive comments and reviews about the new phone and its manufacturer (similar to Samsung), as well as spreading negative comments about a leading competitor (similar to HTC). The fourth part depicts a confrontation between Paul and one of his dear friends who did not approve of his actions, which concludes with Paul's admission of guilt and attempting to justify his participation in the campaign.

This vignette was presented to the interviewees as reading material right before the interview, where they were asked to imagine themselves as this dear friend and reflect on Paul's actions and gauge the plausibility of the different possible ways in which his behavior could be justified. We made sure that the vignette reflected Paul's admission of guilt as a result of his deceptive behavior, since, according to neutralization theory, without feelings of guilt or shame, there would be nothing to neutralize [24], [53], [72].

Twenty interviewees agreed to participate in this study, all of whom are university students (similar to the real campaign). The interviewees were first asked for their initial impressions and the excuses Paul would give to justify his participation in the campaign. After this unprompted discussion, the interviewees would be asked to reflect on the plausibility of various justifications based on the five classical neutralization techniques. Analysis of the data utilized both theory-driven and data-driven coding. Whereas neutralization theory sensitized us to the five classical techniques, data-driven codes allowed us to capture the potentially emerging ones. Based on their answers, we were able to gain rare insights into the justifications that could enable ordinary people to engage in unethical (and potentially illegal) behavior.

5. Findings

In this section, we report our findings on the most prevalent rationalizations deeming participation in cyber deception justifiable. We begin by presenting our findings regarding the acceptability of the five classical techniques. Whereas the appeal to higher loyalty, denial of responsibility, and denial of injury were found to be acceptable justifications, the denial of the victim and condemnation of the condemners techniques were generally unacceptable. Then, we present our findings regarding the emergence of two novel techniques that emerged from the in-depth interview discussions: namely, appeal to professionalism and appeal to normative fluidity.

5.1. Appeal to higher loyalties

The analysis reveals that rationalizations alluding to higher loyalties are the most acceptable form of justification among our interviewees. According to neutralization theory, the appeal to higher loyalties technique allows the offender to commit norm-breaking by claiming loyalty to a **person**, such as family and friends [24], or to a **thing** [73], such as a value or a belief [74]. In the vignette, this rationalization is reflected by Paul claiming loyalty to *"a friend"*, to *"money"*, and to *"having fun"*.

Interestingly, the interviews suggest that the appeal to 'a thing' is more acceptable as an excuse for online deception vis-à-vis the appeal to 'a person'. Appeal to money and making profit, for example, appears to be a widely accepted justification for participating in the campaign. For instance, one interviewee said, *"I think he's after the money and this seemed like a smart way to do it."* [Interviewee_M10]; while another rationalized: *"If Paul was starving, or whatever, and needed the money, I'd be okay with it."* [Interviewee_M2]. By contrast, justifying Paul's role in the campaign as a favor to his friend, although was generally accepted, received some dispute from several interviewees. For instance, some disputed by saying *"I don't think a friend's request warrants such involvement"* [Interviewee_M4], and *"If a friend asks you to do a crime, you should stop that friend"* [Interviewee_F4].

In IS research, the appeal to higher loyalties is among the most applied neutralization technique to justify norm-breaking in both organizational [26], [75], [76] and non-organizational contexts [63], [64], [77]. For instance, in the context of software piracy, this technique is often used to argue that creating unauthorized copies of software is justifiable when it is done to help a friend who cannot afford to purchase the software [65].

5.2. Denial of responsibility

The findings suggest that rationalizations alluding to denial of responsibility are highly acceptable among the study participants. The denial of responsibility technique is used to liberate oneself from the sense of accountability with respect to the situation at hand, arguing for example, that the crime is "due to forces outside of the individual and beyond his control" [24, p. 667]. In the vignette, this rationalization is reflected by Paul dodging accountability for participating in the campaign by making arguments like *"I didn't really know this wasn't allowed"*.

Interestingly, most interviewees considered this technique to be a convincing justification, since Paul could plausibly argue that he was not aware of any rules that prohibit posting false content on his own social

media account. Some of the justifications supporting Paul's denial of responsibility diverted the attention from Paul's actions to the less-than-optimal regulations governing the operations of social media platforms. For instance, some implied that it is not Paul's fault when "*there's probably a loophole somewhere that allows these campaigns to exist."* [Interviewee_M4]. This is not the full picture, however, since some interviewees were not satisfied with this being a good rationalization to join online deceptive campaigns. To them, Paul was only making up an excuse to justify being caught, and hence they did not accept this as justification. In their view, Paul should know better.

In IS research, the denial of responsibility is considered one of the most effective neutralization techniques used to justify violations in both workplace [26], [78] and non-workplace contexts [64], [77]. For instance, in the context of organizational Internet abuse, employees have used this technique to argue that using Internet access provided by the organization for personal purposes is justifiable if they are not sure whether there is Internet use policy in the organization, or if they do not understand it properly [78].

5.3. Denial of injury

Our findings suggest that rationalizations alluding to the denial of injury are highly acceptable. The denial of injury technique provides an argument that diminishes the impact of the act on the victim, arguing for example that act "does not really cause any great harm" [24, p. 668]. In the vignette, this rationalization is reflected with Paul attempting to minimize the impact of his deceptive actions by making statements like *"it's not a big deal"*, *"It's digital environment, so it doesn't really matter"* and *"nobody that could be hurt was hurt"*.

Surprisingly, several interviewees considered Paul's actions as minor if not close to victimless. One interviewee questioned the impact of Paul's role in the campaign and exclaimed: *"what's one guy anyway?"* [Interviewee_M5]. Others justified Paul's actions by claiming that content social media should not be a source of information anyway: "*nobody should be believing social media anyway at this point."* [Interviewee_F1].

In contrast, some interviews challenged Paul's attempt at denying injury by highlighting that an adult person, such as Paul, should recognize that his virtual behavior has physical, non-virtual ramifications. They argued that Paul should be cognizant that people in his social network who trusted him would likely make real-life decisions based on his views. Contesting Paul's justification, one interviewee argued: "*Yeah, but he's trying to sell a physical thing."* [Interviewee_M4].

In IS research, denial of injury is among the most popular justifications in various research contexts, such as insider crime [79], information security policy violations [76], [80] and digital piracy outside the workplace [64], [81]. For instance, in a study on digital music piracy, denial of injury was found to be the most commonly encountered neutralization technique among the study participants [81].

5.4. Denial of the victim

Our analysis suggests that rationalizations reflecting the denial of the victim technique are barely acceptable, considering the rich discussions and disputes this technique spurred in the interviews. According to neutralization theory [24], the denial of the victim technique is transformative in that the offender transforms the victim into a wrongdoer and the offender into a rightful avenger, arguing for example that the victim deserved what happened and that "they had it coming" [24, p. 669]. The vignette reflected this rationalization with Paul blaming his social media followers for not doing proper research, for example, that *"nobody should believe the Internet anyway"* and that *"it's the consumer's fault for being fooled"*.

Most interviewees disputed this rationalizing, and implied that such justification is rather unlikely to be accepted if their friend Paul used it as an excuse. The interviewees generally saw that somebody striving for Internet influence, such as Paul, should not downplay the effectiveness of that influence as doing so would be self-defeating and hypocritical. In other words, one who is doing influence professionally, even as part of a deceptive campaign, cannot downplay the impact of influence!

This finding was surprising considering that in IS research, denial of the victim has been most visible in justifying non-organizational violations [64], [77], [81]–[83]. For instance, in the context of digital piracy, Bhal and Leekha [82] found that denial of victim to be a central cognitive logic underlying not considering software piracy unethical. On the other hand, most workplace IS research has dismissed this technique from investigation on the basis of it not being a plausible rationalization for violating organizational policy [26], [76].

5.5. Condemnation of the condemners

The analysis suggests that rationalizations attempting to condemn the condemners are extremely unacceptable. To Sykes and Matza [24], condemnation of the condemners is a deflective technique whereby the offender turns the spotlight on those who disapprove of the deviant act, arguing for example that those who disapprove are "corrupt, stupid, and brutal" [24, p. 668].

Our vignette reflected this rationalization with Paul's attempts to shift the blame from his actions to the flaws in *"the rules that regulate social media campaigns", "the morals that have you feeling wronged"* and *"the monopolistic practices of social media platforms"*.

Most interviewees dismissed this line of argumentation as being unreasonable. For instance, in response to Paul's attack on the rules, an interviewee vehemently rejected the justification and noted: *"No reasonable person is going to say that about deception!"* [Interviewee_F3].

Interestingly, despite this finding, in IS research, condemning the condemner is often reported as a common justification for violating organizational policy. This contrast is understandable considering that in organizational contexts, employees violate the organizational password policy (e.g., writing it down on a paper) could reasonably argue that the security requirements are complex, frustrating or even annoying [26], [73]. However, in the context of crowdturfing, there is hardly any maneuver that justifies the condemnation of rules that prohibit deception, much less morals and how individuals feel about the practice.

5.6. Appeal to professionalism

We call this emergent technique 'appeal to professionalism'. Discussions with the interviewees pointed our attention to a rationalization rooted in the professional outlook of the campaign. After all, this is a marketing campaign, organized by a legitimate marketing agency and aiming to help a large corporation increase its market share. If all these professionals are involved, the rationalization goes, then the gig itself must be legitimate; *"…this is essentially just a business thing he [Paul] has done."* [Interviewee_F5] "*It was his job and he did his job and that was about it."* [Interviewee_F2].

In addition, several interviewees talked about the grayness of the situation and legality surrounding it, or as Interviewee_M2 puts it: *"Legislation always lags behind technology."* Thus, appeal to (professional) legitimacy seems to be a highly acceptable justification for Paul's participation in the campaign. Our findings point to a very surprising revelation: wherein many interviewees would categorically reject the abstract notion of deception, the same interviewees could accept participating in a deceptive social media campaign if it was offered to them by a legitimate business in the form of a professional job. This became quickly clear from the interviews since several participants struggled to see any wrongdoing from Paul's part. At the end of the interview session with Interviewee_M11, he concluded: *"I'm not sure I can form an opinion on how allowed or not-allowed this [practice] is yet."*

All this combined, it can be easily seen why a lay-user like Paul could be misled into thinking they had joined a professional and legitimate advertising campaign, only after which they came to realize that they were conducting unethical or illegal activities, which they now had to justify somehow. *"Of course, you're going to make excuses now that you've been caught with your pants down."* [Interviewee_M12]. One interviewee explained it succinctly: "*It probably dawned on Paul much later that there was something shady going on but, in the beginning, probably, he didn't know this."* [Interviewee_F6].

The interviewees saw that this justification was good for Paul's participation since the authority at hand, the professional firm which now pays him salary, had deemed the actions and tasks they are giving to Paul as acceptable. This compounded by the perception that these campaigns are prevalent and legal (or at least gray area) can easily lead one to think they can gain legitimate working experience from participating in such a campaign. *"The marketing aspect is important for his future job prospects."* [Interviewee_M10]

What is unique about this technique is that it enables an actor to pass on the blame for their unethical behavior to a superior (e.g., the hiring agency) who has deemed these actions as legitimate, regardless of the actor's own moral assessment of the situation. As such, the actor may see their actions as immoral, but since they were approved by a legitimate authority, the responsibility would eventually lie with the authority and not the actor.

5.7. Appeal to normative fluidity

We call this technique 'appeal to normative fluidity' to reflect a novel justification adopted by several interviewees alluding to the fact that normative rules (e.g., laws, regulations, social norms, etc.) are dynamic and ever changing. As such, what is considered forbidden (or frowned upon) today could be the new norm in the future. Although hiring workers to spread false information on social media might not be accepted today; we cannot dismiss the possibility that *"this is the future of advertising."* [Interviewee_M6]. Of all possible rationalizations, appeal to normative fluidity (e.g., this is the future of advertising) is considered the second most plausible. Interviewee_M9, for example, considers it *"the most realistic"* justification.

This finding points our attention to one of the oldest (and often unrecognized) neutralization theory assumptions: namely, the flexibility of the normative system. Specifically, Sykes and Matza [24] highlighted the complexity stemming from the interpretive flexibility of the normative system where delinquency occurs. This is particularly the case when "the quality of

the values is obscured by their context" [84, p. 715]. Recognizing this sensitizes us to the conflict that may arise from co-existence of formal and informal regulative structures, and that in the era of social media, novel informal rules might be in the making, and when the time is right they would replace today's norms. Benson [85] clarifies this process by pointing the reader's attention to the informal structure. He notes: "… an informal structure exists below the articulated legal structure, one which frequently supersedes the legal structure. The informal structure may define as moral and 'legal' certain actions that the formal legal structure defines as immoral and 'illegal.'" [p. 593]. The flexibility of the normative system makes it plausible to accept that the same act may be seen as 'good' or 'bad', as 'right' or 'wrong', and as 'justified' or 'unjustified' depending on the timeframe and context of assessment.

Could it be that we are experiencing a moral shift where what we consider today immoral/unethical will become the new norm of the near future? Or are these simply the dejected resignations of individuals tired and pessimistic about the prospects of social media manifesting as opportunistic self-interest? This is something for future research to ponder.

6. Discussion

Our work has explored the possible justifications of cyber deception. The central question guiding this endeavor was: what rationalizations could justify participating in the crowdsourcing of deception? The work presented thus far offers some important implications to both theory and practice, as well as directions for future research, which we briefly discuss next.

From a theoretical standpoint, our work demonstrates how neutralization theory could help us understand a core pre-kinetic [23] cognitive process which enables individuals to justify a behavior they themselves generally do not approve of. Neutralization techniques, however, are not universally applicable to all types of deviance. Rather, as Sykes and Matza [24] explain, neutralization techniques are learned in interaction, and that certain techniques are "better adapted to particular deviant acts than to others" [p. 670]. What this means is that we should not expect justification techniques that are suitable to shoplifting [72] or car theft [53] to be readily applicable to other forms of deviance, such as cyber deception. In this regard, our findings provide the first empirical evidence in favor of five exceptionally salient techniques. Of the classical techniques, the appeal to higher loyalty, denial of responsibility, and denial of injury were the most dominant ones. In addition, two newly discovered techniques emerged from our data: appeal to professionalism and appeal to normative fluidity.

As we suggested earlier, deception may have moved on to a new age where the separation between the deceiver and their victims increases as more and more players are involved in the commercialization of deception. For example, in the Samsung-vs-HTC smear campaign [1]–[4], a company was hired by Samsung which eventually had the deception done by students. Most probably, these students, like our study participants, believe themselves to have good moral judgement and to be generally decent human beings. While the interviews were based around deception scenarios committed by a fictional character, many interviewees in our study agreed, by their own admission, that students would be the prime candidate for this work, since they would be tempted to take part in similar campaigns for marginal amounts of money. Effort should be dedicated to making sure this slip does not happen. This leads to some practical implications which we discuss next.

Our work has very important practical implications in combating cyber deception. The key points we can tackle with our findings is to prevent the utilization of neutralization techniques by the people participating in these campaigns. One of the critical steps to successfully combating cyber deviance, such as software piracy, has been to understand how software piracy may be justified [63], [64], [82]. Such understanding has served as a foundation for both policy makers and anti-piracy organizations in their development of counterarguments and campaigns. With this insight we can place two campaigns of similar purpose, one for policy makers and one for social media platforms themselves; these being, respectively, de-neutralization programs and awareness campaigns.

First, our findings can help devise actionable countermeasures against cyber deception through de-neutralization programs. Being attentive to the techniques deployed to justify deceptive crowdwork can inform the design of de-neutralization intervention programs targeting specific techniques. De-neutralization campaigns aim to target common justifications of digital deviancy, hindering their usage to begin with. For example, an anti-neutralization campaign targeting the 'appeal to professionalism', can inform its target audience (e.g., social media users, etc.) that the professional outlook and the business glamour do not make deception any less insidious; that acts (including in cyberspace) should be appraised based on their own virtue, not based on how they are packaged and sold. These arguments could be put in schools together with Internet etiquette and other digital usage studies, in order to have the maximum outreach in preventing digital deception.

Second, awareness campaigns should focus on shedding light on the emerging practice of servitizing cyber deception (i.e., offering it as a service) under the guise of professional crowdwork and gig economy. The intent with awareness campaigns is to familiarize users with more proper channels for utilizing their social media presence, preventing them from slipping into cyber deception campaigns, believing this to be legitimate Internet advertising practice. In relation to this, social media platforms need to make more transparent what is allowed and what is not and make examples out of these faults. Even if one campaign is dealt with internally by a social media platform, to the users this action remains largely invisible.

The crowdsourcing of deception is a largely under-researched area, which is understandable considering its novelty and the challenges associated with its secretive operations. These challenges, however, should inspire future researchers to adopt (and develop) creative research approaches that allow us to gain deeper understanding of how and why these operations succeed. Our work paves the road to several future research directions. Another area of interest for future research is exploring the role of the crowdworkers' cultural norms on rationalizing and neutralizing. It is also important in the future to consider shedding more light on the persuasive role of marketing agencies (i.e., the middleperson) in making deception a viable business model. Finally, it could also be interesting to shed more light on the perspective of the (potential) target-victims and how they perceive, consume, and communicate disinformation in today's cyberspace, and what mitigation measures they should have in place.

7. Conclusion

In this article we explored the techniques by which ordinary users might rationalize their participation in crowdsourced deception. We set out to study these justifications, utilizing neutralization theory as our framework and qualitative vignette as our methodology. Our findings pointed to five exceptionally convincing techniques deeming cyber deception justifiable, namely, appeal to higher loyalty, denial of responsibility, and denial of injury appeal to professionalism, and appeal to normative fluidity. Our work offers groundbreaking contributions to both theory and practice. From the theoretical standpoint, the work extends neutralization theory by demonstrating its applicability in the domain of cyber deception. Furthermore, the work extends neutralization theory by discovering two techniques from the inductive part of the research. Our work highlights the practical importance of awareness campaigns and de-neutralizing programs in the battle against cyber deception.

8. References

[1] C. Moss, "Samsung has just been slapped with a $340K fine for paying people to slam HTC in online forums," Business Insider, 2013. [Online]. Available: https://www.businessinsider.com/samsung-fined-for-posting-negative-htc-reviews-online-2013-10?r=US&IR=T.

[2] A. Souppouris, "Samsung fined $340,000 for faking online comments," The Verge, 2013. [Online]. Available: https://www.theverge.com/2013/10/24/5023658/samsung-fined-340000-for-posting-negative-htc-reviews.

[3] V. Yalburgi, "Samsung Taiwan Apologises for Fake Negative Reviews of HTC Smartphones," International Business Times, 2013. [Online]. Available: https://www.ibtimes.co.uk/samsung-taiwan-fake-negative-reviews-htc-smartphones-459577.

[4] M. Kan, "Taiwan's FTC investigating Samsung for defaming HTC on forums," Tech Advisor, 2013. [Online].

[5] T. Lappas, G. Sabnis, and G. Valkanas, "The impact of fake reviews on online visibility: A vulnerability assessment of the hotel industry," Inf. Syst. Res., vol. Articles i, pp. 1–22, 2016.

[6] H. Luo, "A review of research on identification of false reviews in e-commerce," J. Manag. Humanit. Res., vol. 3, pp. 9–15, 2021.

[7] Y. Wu, E. W. T. Ngai, P. Wu, and C. Wu, "Fake online reviews: Literature review, synthesis, and directions for future research," Decis. Support Syst., vol. 132, no. March, p. 113280, 2020.

[8] W. L. Bennett and S. Livingston, "The disinformation order: Disruptive communication and the decline of democratic institutions," Eur. J. Commun., vol. 33, no. 2, pp. 122–139, 2018.

[9] G. King, J. Pan, and M. E. Roberts, "How the Chinese government fabricates social media posts for strategic distraction, not engaged argument," Am. Polit. Sci. Rev., vol. 111, no. 3, pp. 484–501, 2017.

[10] N. Pham, "Vietnam admits deploying bloggers to support government," BBC, 2013. [Online].

[11] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," Science (80-.)., vol. 1151, no. March, pp. 1146–1151, 2018.

[12] E. Thorson, "Belief echoes: The persistent effects of corrected misinformation," Polit. Commun., vol. 33, no. 3, pp. 460–480, 2016.

[13] R. Han, "Manufacturing consent in cyberspace: China's 'Fifty-Cent Army,'" J. Curr. Chinese Aff., vol. 44, no. 2, pp. 105–134, 2015.

[14] E. Kapantai, A. Christopoulou, C. Berberidis, and V. Peristeras, "A systematic literature review on disinformation: Toward a unified taxonomical framework," New Media Soc., pp. 1–26, 2020.

[15] R. O. Mason, "Four ethical issues of the information age," MIS Q., vol. 10, no. 1, pp. 5–12, 1986.

[16] A. R. Peslak, "PAPA revisited: A current empirical study of the Mason framework," J. Comput. Inf. Syst., vol. 46, no. 3, pp. 117–124, 2006.

[17] R. Von Solms and J. Van Niekerk, "From information

security to cyber security," Comput. Secur., vol. 38, pp. 97–102, 2013.
[18] J. Song, S. Lee, and J. Kim, "CrowdTarget: target-based detection of crowdturfing in online social networks," in Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, 2015, pp. 793–804.
[19] K. Lee, S. Webb, and H. Ge, "Characterizing and automatically detecting crowdturfing in Fiverr and Twitter," Soc. Netw. Anal. Min., vol. 5, no. 2, pp. 1–16, 2015.
[20] K. Lee, P. Tamilarasan, and J. Caverlee, "Crowdturfers, campaigns, and social media: Tracking and revealing crowdsourced manipulation of social media," Icwsm 2013, pp. 331–340, 2013.
[21] D. Zhang, L. Zhou, J. L. Kehoe, and I. Y. Kilic, "What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews," J. Manag. Inf. Syst., vol. 33, no. 2, pp. 456–481, 2016.
[22] D. Straub and R. J. Welke, "Coping with systems risk: Security planning models for management decision making," MIS Q., vol. December, no. 4, pp. 441–469, 1998.
[23] R. Willison and M. Warkentin, "Beyond deterrence: An expanded view of employee computer abuse," MIS Q., vol. 37, no. 1, pp. 1–20, 2013.
[24] G. M. Sykes and D. Matza, "Techniques of neutralization: A theory of delinquency," Am. Sociol. Rev., vol. 22, no. 6, pp. 664–670, 1957.
[25] S. Maruna and H. Copes, "What have we learned from five decades of neutralization research?," Crime and Justice, vol. 32, no. 2005, pp. 221–320, 2005.
[26] M. Siponen and A. Vance, "Neutralization: New insights into the problem of employee information systems security policy violations," MIS Q., vol. 34, no. 3, pp. 487–502, 2010.
[27] D. Matza, Delinquency and Drift. New York: Wiley, 1964.
[28] R. Hyman, "The psychology of deception," Annu. Rev. Psychol., vol. 40, pp. 133–154, 1989.
[29] L. Humphreys and M. Peelo, "Understanding deception: Disentangling skills from conviction," Howard J. Crim. Justice, vol. 52, no. 1, pp. 55–64, 2013.
[30] P. E. Johnson, S. Grazioli, and K. Jamal, "Fraud detection: Intentionality and deception in cognition," Acc., Organ. Soc., vol. 18, no. 5, pp. 467–488, 1993.
[31] S. Grazioli and S. L. Jarvenpaa, "Perils of Internet fraud: An empirical investigation of deception and trust with experienced Internet consumers," IEEE Trans. Syst. Man, Cybern. Part A Syst. Humans, vol. 30, no. 4, pp. 395–410, 2000.
[32] S. Grazioli and S. L. Jarvenpaa, "Consumer and business deception on the Internet: Content analysis of documentary evidence," Int. J. Electron. Commer., vol. 7, no. 4, pp. 93–118, 2003.
[33] D. P. Biros, J. F. George, and R. W. Zmud, "Inducing sensitivity to deception in order to improve decision making performance: A field study," MIS Q., vol. 26, no. 2, pp. 119–144, 2002.
[34] X. N. Deng, K. D. Joshi, and R. D. Galliers, "The duality of empowerment and marginalization in microtask crowdsourcing: Giving voice to the less powerful through value sensitive design," MIS Q., vol. 40, no. X, pp. 1–24, 2016.
[35] D. C. Brabham, "Moving the crowd at Threadless," Information, Commun. Soc., vol. 13, no. 8, pp. 1122–1145, Dec. 2010.
[36] W. Soliman and V. K. Tuunainen, "Understanding continued use of crowdsourcing systems: An interpretive study," J. Theor. Appl. Electron. Commer. Res., vol. 10, no. 1, pp. 1–18, 2015.
[37] B. Bergvall-Kåreborn and D. Howcroft, "Amazon Mechanical Turk and the commodification of labour," New Technol. Work Employ., vol. 29, no. 3, pp. 213–223, 2014.
[38] D. Schlagwein and N. Bjørn-Andersen, "Organizational learning with crowdsourcing: The revelatory case of LEGO," J. Assoc. Inf. Syst., vol. 15, no. Special Issue, pp. 754–778, 2014.
[39] I. Blohm, J. M. Leimeister, and H. Krcmar, "Crowdsourcing: How to benefit from (too) many great ideas," MIS Q. Exec., vol. 12, no. 4, pp. 199–211, 2013.
[40] A. J. Wood, M. Graham, V. Lehdonvirta, and I. Hjorth, "Networked but commodified: The (dis)embeddedness of digital labour in the gig economy," Sociology, vol. 53, no. 5, pp. 931–950, 2019.
[41] G. Wang et al., "Serf and turf: Crowdturfing for fun and profit," in WWW, 2012, p. 10.
[42] T. Eisenmann, G. Parker, and M. Van Alstyne, "Strategies for two-sided markets," Harv. Bus. Rev., vol. 84, no. 10, p. 12, 2006.
[43] D. S. Evans, "Some empirical aspects of multi-sided platform industries," Rev. Netw. Econ., vol. 2, no. 3, pp. 191–209, 2003.
[44] G. Parker and M. Van Alstyne, "Two-sided network effects: A theory of information product design," Manage. Sci., vol. 51, no. 10, pp. 1494–1504, 2005.
[45] M. Leiser, "AstroTurfing, 'CyberTurfing' and other online persuasion campaigns," Eur. J. Law Technol., vol. 7, no. 1, pp. 1–27, 2016.
[46] K. Lee, S. Webb, and H. Ge, "The dark side of micro-task marketplaces: Characterizing fiverr and automatically detecting crowdturfing," Assoc. Adv. Artif. Intell., pp. 275–284, 2014.
[47] M. Luca and G. Zervas, "Fake it till you make it: Reputation, competition, and Yelp review fraud," Manage. Sci., vol. 62, no. 12, pp. 3412–3427, 2016.
[48] Y. Yao, B. Viswanath, J. Cryan, H. Zheng, and B. Y. Zhao, "Automated Crowdturfing Attacks and Defenses in Online Review Systems," arXiv Prepr. arXiv1708.08151, 2017.
[49] N. Abokhodair, D. Yoo, and D. W. McDonald, "Dissecting a social botnet: Growth, content, and influence in Twitter," in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15, 2015, pp. 839–851.
[50] S. Rossi, M. Rossi, B. Upreti, and Y. Liu, "Detecting political bots on Twitter during the 2019 Finnish parliamentary election," Proc. 53rd Hawaii Int. Conf. Syst. Sci., vol. 3, pp. 2430–2439, 2020.
[51] G. Li, W. Niu, L. Batten, and J. Liu, "New advances in

securing cyberspace and curbing crowdturfing," Concurr. Comput. Pract. Exp., vol. 29, no. 20, pp. 1–3, 2017.
[52] S. Farooqi et al., "Characterizing Key Stakeholders in an Online Black-Hat Marketplace," arXiv:1505.01637v2, 2017.
[53] V. Topalli, "When being good is bad: An expansion of neutralization theory," Criminology, vol. 43, no. 3, pp. 797–835, 2005.
[54] C. B. Klockars, The professional fence. New York: Free Press, 1974.
[55] W. W. Minor, "Techniques of neutralization: A reconceptualization and empirical examination," J. Res. Crime Delinq., vol. 18, no. 2, pp. 295–318, 1981.
[56] S. McGregor, "Conceptualizing immoral and unethical consumption using neutralization theory," Fam. Consum. Sci. Res. J., vol. 36, no. 3, pp. 261–276, 2008.
[57] S. J. Harrington, "The effect of codes of ethics and personal denial of responsibility on computer abuse judgments and intentions," MIS Q., vol. 20, no. 3, pp. 257–278, 1996.
[58] M. Nicho and F. Kamoun, "Multiple case study approach to identify aggravating variables of insider threats in information systems," Commun. Assoc. Inf. Syst., vol. 35, no. 18, pp. 333–356, 2014.
[59] V. K. G. Lim, "The IT way of loafing on the job: cyberloafing, neutralizing and organizational justice," J. Organ. Behav., vol. 23, no. l, pp. 675–694, 2002.
[60] L. Khansa, J. Kuem, M. Siponen, and S. S. Kim, "To cyberloaf or not to cyberloaf: The impact of the announcement of formal organizational controls," J. Manag. Inf. Syst., vol. 34, no. 1, pp. 141–176, 2017.
[61] M. Silic, J. B. Barlow, and A. Back, "A new perspective on neutralization and deterrence: Predicting shadow IT usage," Inf. Manag., vol. 54, no. 8, pp. 1023–1037, 2017.
[62] S. Haag, A. Eckhardt, and A. Schwarz, "The acceptance of justifications among shadow IT users and nonusers – an empirical analysis," Inf. Manag., vol. 56, no. 5, pp. 731–741, 2019.
[63] S. Hinduja, "Neutralization theory and online software piracy: An empirical analysis," Ethics Inf. Technol., vol. 9, no. 3, pp. 187–204, 2007.
[64] J. R. Ingram and S. Hinduja, "Neutralizing music piracy: An empirical examination," Deviant Behav., vol. 29, no. 4, pp. 334–366, 2008.
[65] M. Siponen, A. Vance, and R. Willison, "New insights into the problem of software piracy: The effects of neutralization, shame, and moral beliefs," Inf. Manag., vol. 49, no. 7–8, pp. 334–341, 2012.
[66] C. Barter and E. Renold, "The use of vignette in qualitative research," Soc. Res. Updat., vol. 25, no. 9, pp. 1–6, 1999.
[67] C. Barter and E. Renold, "'I wanna tell you a story': Exploring the application of vignettes in qualitative research with children and young people," Int. J. Soc. Res. Methodol., vol. 3, no. 4, pp. 307–323, 2000.
[68] N. Jenkins, M. Bloor, J. Fischer, L. Berney, and J. Neale, "Putting it in context: The use of vignettes in qualitative interviewing," Qual. Res., vol. 10, no. 2, pp. 175–198, 2010.
[69] N. Rahman, "Caregivers' sensitivity to conflict: The use of the vignette methodology," J. Elder Abus. Negl., vol. 8, no. 1, pp. 35–47, 1996.
[70] J. B. Barlow, M. Warkentin, D. Ormond, and A. R. Dennis, "Don't make excuses! Discouraging neutralization to reduce IT policy violation," Comput. Secur., vol. 39, no. PART B, pp. 145–159, 2013.
[71] J. B. Barlow, M. Warkentin, D. Ormond, and A. R. Dennis, "Don't even think about it! The effects of antineutralization, informational, and normative communication on information security compliance," J. Assoc. Inf. Syst., vol. 19, no. 8, pp. 689–715, 2018.
[72] P. Cromwell and Q. Thurman, "The devil made me do it: Use of neutralizations by shoplifters," Deviant Behav., vol. 24, no. 6, pp. 535–550, 2003.
[73] M. Siponen, P. Puhakainen, and A. Vance, "Can individuals' neutralization techniques be overcome? A field experiment on password policy," Comput. Secur., vol. 88, 2020.
[74] B. Byers, B. W. Crider, and G. K. Biggers, "Bias crime motivation: A study of hate crime and offender neutralization techniques used against the Amish," J. Contemp. Crim. Justice, vol. 15, no. 1, pp. 78–96, 1999.
[75] P. L. Teh, P. K. Ahmed, and J. D'Arcy, "What drives information security policy violations among banking employees? Insights from neutralization and social exchange theory," J. Glob. Inf. Manag., vol. 23, no. 1, pp. 44–64, 2015.
[76] G. D. Moody, M. T. Siponen, and S. Pahnila, "Toward a unified model of information security policy compliance," MIS Q., vol. 42, no. 1, pp. 285–311, 2018.
[77] L. C. Harris and A. Dumas, "Online consumer misbehaviour: An application of neutralization theory," Mark. Theory, vol. 9, no. 4, pp. 379–402, 2009.
[78] L. Cheng, W. Li, Q. Zhai, and R. Smyth, "Understanding personal use of the internet at work: An integrated model of neutralization techniques and general deterrence theory," Comput. Human Behav., vol. 38, pp. 220–228, 2014.
[79] T.-C. Lin, J. S.-C. Hsu, Y.-C. Wang, and S. Wu, "Examining the antecedents of employee unauthorized computer access," J. Stat. Manag. Syst., vol. 21, no. 3, pp. 493–517, 2018.
[80] A. Vance, M. T. Siponen, and D. W. Straub, "Effects of sanctions, moral beliefs, and neutralization on information security policy violations across cultures," Inf. Manag., p. 103212, 2019.
[81] R. Moore and E. C. Mcmullan, "Neutralizations and rationalizations of digital piracy: A qualitative analysis of university students," Int. J. Cyber Criminol., vol. 3, no. 1, p. 441, 2009.
[82] K. T. Bhal and N. D. Leekha, "Exploring cognitive moral logics using grounded theory: The case of software piracy," J. Bus. Ethics, vol. 81, no. 3, pp. 635–646, 2008.
[83] S. Yu, "Digital Piracy Justification: Asian Students Versus American Students," Int. Crim. Justice Rev., vol. 23, no. 2, pp. 185–196, 2013.
[84] D. Matza and G. M. Sykes, "Juvenile delinquency and subterranean values," Am. Sociol. Rev., vol. 26, no. 5, pp. 712–719, 1961.
[85] M. L. Benson, "Denying the guilty mind: Accounting for involvement in a white-collar crime," Criminology, vol. 23, no. 4, pp. 583–607, 1985.