A (Nudge) Psychological Reading of the "Nigerian Scam"

Till Neuhaus

Bielefeld University,
Faculty of Education,
Germany

till.neuhaus@uni-bielefeld.de

Abstract. The "Nigerian Scam" (so named because of its prevalence in the country, especially during the 1990s, then continued by many organizations in other different regions around the world) is a scheme in which the sender requests help in facilitating a transfer of money. In return, he offers a large commission, sometimes up to several million dollars. The scammers request that money be sent to pay for some of the costs associated with the transfer. These attempts (also known as "advance fee fraud "or "419 fraud" - there is a section in the Nigerian Criminal Code, *i.e.* Section 419, that point at this type of fraud as illegal), are widely regarded as a joke among digital natives. However, forms and variations of the Nigerian Scam have been successful since the 16th century and continue to do so, even in the 21st century. The longevity of the scam hints at the exploitation of very basic human processes. Therefore, this article tries to analyse these processes from a psychological standpoint, trying to derive the mechanisms that these texts exploit. The different phases of the scam (from the creation of the target group, until the final contact) are analysed from the psychology of persuasion as well as behavioural economics standpoints – both being subsumed under the label of "Nudging" – trying to identify the settings, scenarios, framings, and signals which make the scam one of the most successful scams in human history.

Keywords: scam mail, nudge, nudging, behavioural economics, Nigerian scam

1. Introduction and the Nudge Theory

Almost all texts are written with an intention. The text's intention can be – just to name a few – to share information, to publicly position oneself towards a topic, or to manipulate others. Every day,

we consume countless (fractions of) texts (*i.e.* news, adverts, text messages, books etc.) and have to select the corresponding mode of perception. Almost instantaneously we figure out the author's intentions, their agenda, and our relationship with him. While our systems to analyse the aforementioned aspects developed under different evolutionary conditions, the world and, thereby, also the production, positioning, and form of texts has changed tremendously. Often, authors of particular texts try to exploit our evolutionarily grown systems to perpetuate their agenda; an (in)famous example may be the advertisement's struggle for our attention. However, some other authors have more malicious intentions and whose fraud generates roughly $3 billion in damages annually (cf. Park et al. 2014: 1) – scam emails. May it be a Nigerian Prince, a clergyman in need, or a tragedy-struck high-ranking military official, most scam emails are degraded to the status of a running-joke for online natives. However, the scam proves to be relatively stable, having a history of more than 400 years. Therefore, it can be assumed that the authors of scam emails (and the genres prior to online communication) have perfected the art of exploiting evolutionary grown mechanisms. Therefore, this contribution will analyse scam mails from a psychological standpoint trying to highlight the mechanisms which famous scam mail narratives and their presentation attempt to exploit. The 21st century is certainly a century of mass media but it is also the century of persuasion, may it be in politics, public debate, or personal matters. The study of scam emails will prove to be valuable to understand the basic mechanisms of human perception, recognition, and thinking, highlighting the mechanisms which a plethora of actors – from advertisement professionals to politicians – attempts to exploit.

In this contribution, the communication between scammers and victims will be conceptualized by highlighting the four crucial stages of the scam: (1) the pre-conversation phase (referred to as the *bait* phase), (2) the *hook* phase in which the scammers try to convince the victim to transfer money, (3) the *line* phase in which it is attempted to extract more money from the victim, and (4) the *sinker*

phase. In each of the paragraphs, the underlying psychological processes will be highlighted, contextualized, and explained. In the final section of this paper, the prior analysis' insights will be summarized and it will be attempted to outline the future potential of reading texts from a psychological/Nudge theory perspective. First, however, the theoretical framework of this paper, the so-called Nudge psychology approach, will be presented.

The *Nudge theory* assumes that human beings categorize decisions by either letting system 1 or system 2 decide (cf. Kahneman 2012). Instead of system 1 and 2, Thaler/Sunstein (2017, 34) use the terms *Humans* (for system 1) and *Econs* (for system 2) as they associate the traits of system 1 with rather irrational and thereby human behaviour while system 2 represents a more analytical, rational-choice-like approach – a way of thinking they associate with the discipline of economics. However, Kahneman as well as Thaler and Sunstein, both describe the same differentiation. System 1/*Humans* is intuitive, uncontrolled, decides fast, is highly associative, and uses props and cues it has learned earlier. On the contrary, system 2/*Econs* is rule-governed, highly reflective, slow, conscious, and controlled (cf. Thaler/Sunstein 2017, 34). At the very core, this systemization of decision-making is the psychological explanation of and differentiation between the *blink vs. think* dichotomy (cf. Gladwell 2005).

Nudging further assumes that there are patterns in human decision-making, especially when decisions deviate from rationality and/or the best interest of the decision-maker. The deviations of actual behaviour from the rational choice paradigm cannot be predicted by von Neumann's and Morgenstern's (1944) (normative) rational-choice theory as "deviations of actual behaviour from normative models are too widespread to be ignored, too systematic to be dismissed as a random error, and too fundamental to be accommodated by relaxing the normative system" (Tversky/Kahneman 1986, 3). This led Tversky and Kahneman to the conclusion that rational choice theory "is grossly inadequate as a descriptive model of individual choice behaviour "(Tversky 1975,

163). The first systematization of irrationalities in decision-making processes under uncertainty has been presented by Amos Tversky and Daniel Kahneman as they focused on the cognitive-perceptual aspects of this phenomenon – a line of thinking they coined Prospect Theory and which can be understood as a descriptive psychological model of decision making and judgement under uncertainty (cf. Tversky/Kahneman 1974). Prospect Theory is the umbrella term for a plethora of effects and biases in which the presentation or availability of information changes the decision outcome. One such example is a study in the field of medicine in which participants of all level of expertise (expert/physician, amateur/patient) have been confronted with a treatment option. In the first case, it was presented with a 10% mortality rate, in the other experiment with a 90% chance of survival. While both presentations describe the same treatment, the expert as well as the amateur test subjects "were influenced by several variations in the nature of the data and the form in which they were presented" (McNeil et al. 1982, 1262) – a clear violation of rational-choice principles which should later find its way into the literature as the *framing effect* (cf. Tversky/Kahneman 1979: 3/4; Tversky/Kahneman 1981). Generally speaking, Prospect Theory identifies cases, scenarios, and modes of presentation which allow a predictable modification of decision outcomes (cf. Kahneman/Tversky 1977). Nudge theory widens this approach as not just cognitive-perceptual aspects of the environment are considered but all aspects with the help of which a decision is made – the sum of these aspects is often referred to as the "decision architecture" (Thaler/Sunstein 2017). The decision architecture incorporates all relevant aspects, such as provided information, contextualization/framing of information, reference groups, social aspects, wording, and a plethora more; nudging thereby enriches the insights from decision-making psychology and behavioural economics with knowledge from social psychology, socio-linguistics, and sociology (cf. Neuhaus 2020, 74). The conscious modification of decision architectures has proven to be a potent tool as it has the potential to influence the decision-maker

and ultimately change the outcome/decision; in the following, institutions, corporations, and governments used nudging to improve settings in their interest (cf. Straßheim/Jung/Korinek 2015).

The power of nudging lies in its possibility to create a decision architecture – providing the *right* information and present it accordingly – which leads the decision-maker to a fast, intuitive, and unconscious System 1 decision while the problem the decision-maker is actually tackling would rather require the analytical, slow, and reflective System 2. This mismatch of the presented problem and the responding system can be used to nudge people into a certain direction. However, as this paper will show, the same mechanisms can also be exploited to modify decision architectures aiming at the victimization of people as decision architectures play a tremendous role in online communication as the owner of a website, the host of a blog, or the writer of an email fully controls the decision architecture. S/he solely decides which information is provided and how the information is presented. Decision architectures are a powerful tool as people do not act based on objective reality or facts but what they assume reality to be (cf. Haltermann 2012, 64) – Daniel Kahneman coined this availability bias the *What You See Is All There Is* (WYSIATI) rule (cf. Kahneman 2012, 115). As this paper will show, the scammers conducting the Nigerian Scam have perfected the modification of decision architectures by applying a plethora of nudges – an, up to this point, neglected perspective on spam emails as one sub-category of online communication. This paper will analyze the different steps of the scam, highlighting, contextualizing, and explaining the applied nudges and later outline their potential for legitimate online communication.

2. How Nigerian Scammers Nudge their Victims

The Bait Phase: Spam emails are, compared to other media and modes of communication, relatively cheap while being quick, direct,

reliable, well-scalable, and globally accessible (cf. Saini 2012, 1). Therefore, they are sent to either stolen, collected, or otherwise made available email accounts; the scammers do not discriminate based on any observable trait. It could be assumed that scammers aim at a high response rate as it enhances their chances to victimize responders. However, the way the spam mails are written directly contrasts this assumption as scammers emphasize their Nigerian nationality – whether this is actually the case or not cannot be confirmed – and also construct highly unlikely and dubious scenarios, a red flag for every well-informed internet user (cf. Schaffer 2012, 162). The disparity between the scammer's approach (the mails) and their obvious aim (victimize people for monetary benefit) can be explained by looking deeper into the economics of the scam itself. Initiating the contact by sending a spam mail is comparatively cheap. Costs for the scammers are caused by on-going and continuing exchange with the potential victim as writing more personal messages requires time and human capital. As a result, it is in the scammers' interest that, if they invest time, money, and effort, the endeavour results in the victimization of their contact and pays off monetarily. The worst-case scenario for the scammers would be the investment of time, manpower, and money by having a long-lasting conversation which results in no payment (cf. Herley 2012, 11). The scammers rightfully assume that naïve internet users make better targets as they are less likely to research the presented narratives online, seek advice (*i.e.* in forums), or discuss such matters with informed people offline (cf. Herley 2012, 11/12). Based on these observations, it could be concluded that the core traits of a potential victim are social isolation, advanced age, and naivety. All of these items are unobservable based on the information source the scammers use, *ergo* the email addresses (cf. Levitt/Dubner 2016, 160). For the scammers, this poses the question of how they can make a usually unobservable trait visible. The solution the scammers came up with is a mechanism called self-selection, sometimes also referred to as the self-selection bias, which could be described as a selection by proxy; in legitimate businesses, pricing is often used to

trigger this self-selection (cf. Harford 2007, 52/53). The Nigerian Scam, on the contrary, uses the absurdness of the initial story to select for unobservable traits and thereby exploits the core insight from the game theory which assumes a change in behaviour from all involves actors in case of a change in rules and/or signals (cf. Binmore 2013, 155/156).

By making their narratives highly dubious and using fraud-connoted buzz words, the scammers target a specific sub-population of the people they approach. Only the ones who respond to these outrageous narratives have proven to have a tendency for naivety – the key trait of a potential victim. By responding to the initial mail, the – prior to the self-selection process – uncategorizable crowd of people has organized itself into potential victims and non-responders. Simultaneously, this selection process makes the scam financially more profitable as the non-responders – which could have been lured into a conversation as they would have responded to a less fraud-laden mail – would have had a higher probability of opting-out at a later stage of the scam; the worst-case scenario for the scammers as it produces the highest cost while generating zero revenue. By including certain connoted buzz words, the scammers reduce the amount of responses dramatically as only one out of 12.5 million emails is responded to (cf. Wainwright 2017, 150; Kanich et al. 2008, 11). However, by letting the tremendous amount of approached people categorize themselves among the line of desirable traits for the scammers, the initiators of the scam improve their likeliness for monetary reward and minimize the probability for high costs, low revenue scenarios. By intelligently using the right wording and narratives, the scammers do not only reduce their target group, they optimize their target group. This *target group optimization* has a long history with the Nigerian Scam. In the 1980s, online communication was far away from being a mainstream phenomenon and scams of this kind took place by letter mail. This changed the setting dramatically as initial costs were much higher compared to the digital version of the scam. As the scammers could not address all people indiscriminately and hope for self-selection

processes, the scammers showed a prevalence for addressing people with partly visible traits. In the times of letter-based communication, the scammers primarily targeted people registered with drug abuse and/or addiction backgrounds. Furthermore, they addressed people related to charities, churches, or registered debtors and/or alcoholics (cf. Kich 2005, 129). The reason for this kind of pre-selection can be explained by the hidden, yet often related, traits of the aforementioned groups. Alcoholics, debtors, and/or addicted people constitute a sub-population which is, compared to the rest of society, relatively short on money. As the scam promises financial rewards for providing advanced-fees, people with a lack of monetary resources are expected to respond more positively to a *make money fast* scheme (cf. Stajano/Wilson 2011, 73). The selection of charity and/or religion-related populations can be explained by the assumed trait of empathy. People being engaged in such kind of organizations are assumed to be more empathetic and therefore respond more positively to people in need – the second catch of the scam. These reflections of the scam's past, as well as its present, illustrate that the channel of distribution – and the costs it creates – correspond to the amount of people being addressed. While cheap distribution allows more people to be addressed and requires a different mode of (self-)selection, higher costs require a more careful selection of the target group.

The Hook Phase: The second phase of the scam, the hook phase, prepares the potential victim psychologically and mentally to transfer money to the scammers. This is primarily achieved by a modification of the relationship between the victim and the scammer. Therefore, the scammers send out signals (cf. Spence 1973, 356) which should lead victims to the assumption that they speak with a trustworthy person. One key tendency which helps the scammers in this phase – even though it expands through all phases of the scam – and significantly contributes to its success is the *truth default* (cf. Levine 2014). The truth default can be described as the tendency that humans generally assume that other people tell the truth, at least as long as the narrative follows certain conventions.

While the truth default certainly helps, the focus of this section will be set on the exploitation of irrationalities in the form of nudging as it is the scammer's main concern to appear trustworthy – a trait which cannot be proper evaluated through email communication. The different signals/nudges discussed in this section can be found either in the initial spam mail or in later messages as multiple variations of the scam exist. These variations are partly caused by the informational props the scammers use (*i.e.* references to current events) but also due to different realizations competing in the market. A scam can be conceptualized evolutionary as scammers try out different versions of fraud, wording, narratives etc. and continue the most successful ones while retiring the less functional versions. This would also explain why descriptive psychological research identified the approaches, settings, and scenarios in which predictable irrationality occurs and these scientific insights match with the observable praxis of scammers – both analyzed, yet with alternating intentions, the behaviour of actual human beings. The nudges discussed in this section (authority, scarcity, reciprocity, and likeability) are well-recorded and often applied cues, props, and narratives. Therefore, it can be assumed that they are, evolutionary speaking, relatively stable and hint at the more basic structures of human decision-making.

The first nudge almost all scammers use is the *authority* nudge. The majority of recorded scam mail, which could be classified as advance fee fraud (cf. Chawki 2009, 4), features characters of authority. These can be high ranking military staff, government officials, royalty, or clerical authorities (cf. Edelson 2003, 393/394). The authority nudge exploits a mental shortcut which has evolutionary roots. Human beings organized themselves in dominance hierarchies ever since. In modern times, these dominance hierarchies multiplied as the field in which one can compete diversified as well. The functionality of social entities (*i.e.* families, companies, states etc.) depends on these dominance hierarchies as – at least in functional hierarchies – skill, effort, and competence determine one's positioning in the hierarchy (cf.

Peterson 2018). This allocation mechanism promises that jobs and tasks are assigned based on competence. This should increase the probability that tasks and problems are handled by people who can actually solve them and thereby increase the overall well-being of the group. However, as human social structures increased in size and grew more complex, competence and skill could no longer be evaluated properly. One way to, at least in part, guarantee the allocation of duties based on merit was the certification of skills through diplomas, certificates, and job titles. Following this line of reasoning, it can be argued that – as long as the dominance hierarchy is functional – job titles, diplomas, and certificates are signals of competence and, as such, they are barely ever questioned (cf. Cialdini 1987, 176). Especially in times of and/or judgment under uncertainty, human beings show the tendency to turn to authorities. Also, almost all individuals have learned from experience that following the instructions of authority (*i.e.* a doctor or a teacher) leads to more satisfying results (cf. Cialdini 1987, 175). Especially in judgements outside one's own domain of competence, the mental shortcut which equals authority with competence has proven to be particularly potent. One such example is the (in)famous Milgram experiment in which a medical/scientific authority figure urged participants to administer electric shocks to another person (cf. Milgram 1965). The Milgram experiment underlines human's willingness to outsource decisions based on competence by proxy of authority. The same nudge is applied in the discussed spam emails as scammers assume that potential victims are more willing to follow instructions if an authority figure urged them to do so. The success of the scam partly depends on the credibility and authenticity with which the scammer can pretend to be an authority figure. This statement can be supported by observations from historical fraud research. Variations of the Nigerian Scam can be traced back to the 16th century (cf. Smith 2009, 28), back then being known as *The Spanish Prisoner*. The Spanish Prisoner was also an advance fee fraud in which an alleged Spanish royal needed money to get himself out of jail. In return for the generous gesture, the royal would make the

good-hearted helper a rich man. In the 16th century, few people ever received a mail and even fewer received letters written on an expensive, high-quality paper (cf. ibid.). The selection of the materials and the mode of communication were chosen on purpose to increase the credibility of the claims being stated in the letters. While multiple aspects of the scam have changed over time, some remained stable. One of the most resilient traits of the scam is the usage of authority figures – a powerful nudge with a recorded history of more than 400 years.

While the usage of alleged authority figures helps to perpetuate the authority aspects of the nudge, it also contributes to a second nudge, the *scarcity* nudge. Scarcity feeds two very basic human's needs: the desire to be special (cf. Cialdini 2003, 22) and regret aversion. The desire to be special is exploited by the scam as – at least for the naïve victim – it looks like as if s/he was chosen to help the authority figure in question. In the illusion created by the scam, the basic dynamic of competence hierarchies is reversed as the, based on position, a lower-level person is in charge to save the authority figure. The feeling of being needed by someone higher in standing creates a feeling of privilege. Dominance hierarchies also have the tendency that only very few people ascend to the top. As the scammers signal belonging to that precious and tiny group, scarcity is created and ultimately exploited. The power of the scarcity nudge can also be explained by evolutionary learning processes. In human history but also in almost every person's experience, scarcity signalled desirability (cf. Modic/Lea 2013, 5). May it be precious metals, prestigious institutions, or unique opportunities, scarcity has long been associated with desirability and thereby value. The well-known laws of supply and demand are a mathematical manifestation of this correlation. However, the scarcity nudge not just creates the illusion of being chosen for a rare and thereby precious opportunity but also triggers a mechanism called regret aversion. The analysis of spam mails has shown that many of these messages operated with words signalling urgency (cf. Bergiel/Bergiel/Balsmeier 2008, 137) and thereby multiplying the

potency of the scarcity nudge. The creation of urgency has two aims. Firstly, it actively pushes the decision into the field of System 1/*Humans* – even though System 2/*Econs* should be the victim's system of choice – as System 1 is responsible for fast and in this case urgent decisions. Secondly, the creation of urgency triggers *regret aversion*. Regret aversion (cf. Loomes/Sudgen 1983) describes the procedure in which humans imagine their future selves under certain conditions, *i.e.* realizing a scenario or spurning an opportunity. This procedure is a simplification mechanism to facilitate judgement under uncertainty. The incredibly complex System 2 question (Should I engage in this endeavour?) is replaced by the much simpler System 1 question (How would I feel if I miss out on this opportunity?). The imagined state of missing out on the opportunity results in a strong feeling of regret. The comparison of the anticipated future emotional states replaces the fact-driven analysis of the scenarios and results in higher degrees of participation. While the substitution of questions is against the laws of logic and analytical decision-making, the scarcity nudge – as one way to facilitate the processes described above – has proven to be a powerful nudge which can alter behaviour significantly.

Another relatively frequently occurring phrasing in these spam emails is the emphasis that the sender looks for a trustworthy individual to help him/her. Some scammers even go a step further and attach copies of confidential documents (*i.e.* ID cards, driver's licenses, or passports) to the initial or one of the later mails – the attempt to attach credibility to the signal (cf. Zahavi 1997, 228/229). While the documents are likely to be manipulated, stolen, or forged, both – the overemphasis on trust as well as the attached documents – try to exploit the same weakness: human's tendency for *reciprocity*. Reciprocity is a mechanism of mental bookkeeping which should tell the involved actors whether they are rather indebted or the debt holder in human relations (cf. Cialdini 2003, 20; Cialdini 1999). This form of mental bookkeeping is not limited to monetary or material units but is also used for the conceptualization of human relations. Historically speaking, reciprocity is the in-built mechanism which

urges humans to keep symmetrical relations with their environment and thereby enable long-term cooperation (cf. Cox 2004, 262). As all human beings – yet in varying degrees – have the desire to be in balance with their social environment, reciprocity makes human behaviour more predictable. Reciprocity may be the key factor facilitating long-term cooperation and thereby being a crucial building block for human development and progress as, from a game-theoretical/economic perspective, cooperation is crucial for utility maximization (cf. Hammerstein 2002, 84). The scammers try to exploit this mechanism by creating a scenario of asymmetrical reciprocity. As argued, mental bookkeeping is not an analytical tool which calculates amounts and compares these but should rather be considered an indicator whether an individual is indebted or not. Asymmetrical reciprocity aims at exploiting the differences between rational-based and feelings-based indebtedness. By providing immaterial and unsolicited gifts, such as trust, compliments, copies of forged documents etc., the scammers try to push the mental balance into the area of indebtedness. The victim's perceived indebtedness creates the urge to even out the balance (cf. Cialdini/Goldstein 2002: 43). Out of this asymmetrical relationship, the scammer approaches the potential victim and asks for the transfer of monetary resources. After having subconsciously consulted his/her mental bookkeeping, the victim arrives at the conclusion that s/he owes the scammer a favour and – to a higher degree than without the reciprocity nudge – decides to transfer money on a foreign account. Reciprocity is a well-known mechanism which has been exploited by multiple groups and organizations in real life, such as the *Osho* cult which asked its members to beg for donations. The members of the cult gave potential contributors a flower and a blessing before asking for a contribution – the creation and exploitation of asymmetrical reciprocity which resulted in high revenue for the group. As shown, the exploitation of asymmetrical reciprocity has a long and vivid history. However, as the digital world does not allow the instant transfer of material goods, the scammers confine themselves to

immaterial asymmetry through signalling trust and the transfer of (perceived to be) important documents.

The last nudge applied in the hook phase is the *likability* nudge. An increase in likability makes cooperation between the victim and the scammer more likely (cf. Cialdini/Goldstein 2002, 40). The mental shortcut which is exploited by the likability nudge can best be described by the concept of the *mental shotgun* (cf. Kahneman 2012). The mental shotgun effect takes place when a relatively difficult question is replaced by a related but significantly easier one. The question of whether the received email(s) are legit is replaced by the question of whether the reader likes the sender or not. Due to the limited communication, the basis on which the reader has to decide is relatively small and solely controlled by the fraudster. Especially with regard to the likability nudge, information about the target group is the key as scammers try to create an in-group feeling and/or in-group mentality (cf. Sagarin/Mitnick 2012, 33). The psychological basis for this approach can be found in the observation that human beings tend to evaluate others based on similarities; the more similarities two people share, the more positive they will evaluate each other. Such in-group mentality can be created by hinting at shared spiritual/religious beliefs, the belonging to a certain group of people (*i.e.* bereaved people), or struggling with a specific problem (*i.e.* the loss of a relative). Recorded cases of direct attempts to victimize religiously engaged people, 9/11 victims' relatives, or relatives of deceased soldiers (cf. Dyrud 2005, 6) illustrate that scammers are well-aware of the in-group feeling's potency to create social proximity and thereby likability. However, likability cannot just be created by social proximity and likeliness but also by an appeal to the potential victim's self-image. An insight from marketing says that humans do not buy items based on who they are but based on who they would like to be (cf. Landon Jr. 1974). The same can be said about the Nigerian Scam. Most people (want to) think of themselves as helpful, cooperative, and good-hearted – a positive self-image can be considered a basic human desire (cf. Hobson 2012, 170). The scammers try to exploit this

tendency by presenting themselves as ill-fated, out of luck, and struck by destiny. (In)Famous realizations of these attempts are severe illnesses (of relatives or the scammer him-/herself), unforeseeable accidents, or other strokes of fate (cf. Dyrud 2005, 7). These narratives try to appeal to the potential victim's concept of self by providing him/her the chance to act in a way which would underline his/her self-image.

This section highlighted the four main nudges used by advance fee scammers in the first stage of the scam. The four principles the fraudsters try to exploit in this phase of the scam are authority, scarcity, reciprocity, and likability. The main purpose of this phase – in this chapter conceptualized as the *hook* phase – is the transfer of money. As the following section will show, the initial amount is secondary as the follow-up phase, the *line* phase, has mechanisms installed which make it more likely that victims repeatedly transfer money and, over time, also increase the sum they are transferring – a steady cost/damage escalation. As the analysis of typical scam mail and the underlying psychological implications could show, the scammers employ the concepts of authority, scarcity, reciprocity, and likability as they are – even if differently pronounced in varying individuals – universals of human behaviour. These universals have a long-standing history as they were partially produced by basic human cognition and brain physiology, in part by culture, and in some cases by both.

The line phase begins after the first monetary transfer has taken place. This phase aims at the prolonging and escalation of money transfers. From a narratological standpoint, the scammers justify this on-going cost escalation with sudden and unforeseeable extra fees they have to pay. This phase makes use of three nudge-related scenarios, namely the consistency nudge, cognitive dissonance, and modified perception of risk. All three of these effects prolong and exacerbate the scam, often with detrimental effects.

Research has shown that people generally try to be consistent in their thinking, actions, and attitudes (cf. Guadagno/Cialdini 2010, 152/153). The desire to act in accordance with decisions made prior

is a relatively stable trait which can be found across cultures (cf. Petrova/Cialdini/Sills 2007). It seems to be the case that the initial decision – *i.e.* to transfer money to scammers or not – requires a relatively huge amount of mental resources. After that decision has been made, similar decisions or repetitions of the initial decisions require less mental effort; an on-going migration of the decision from System 2 to System 1. Also, it does not seem to be the case that, in the escalation of costs, the amount seems to be a factor. Once a decision has been made and the mental threshold has been passed, the probability to repeat the action becomes more likely. Therefore, the scammers apply all nudges outlined in the prior section trying to convince the victim to transfer a relatively moderate amount. After the initial threshold has been overcome and the decision steadily migrates from Systems 2 to System 1, the amounts steadily escalate. The consistency nudge highlights the importance of the temporal dimension – when to ask for which amounts – of the scam and aligns it with existing knowledge on decision-making in singular and repetitive decision scenarios.

The effect of *cognitive dissonance* hints at a phenomenon related to the consistency nudge; however, it plays a different role in the scam. Cognitive dissonance (cf. Festinger 1957) describes the difference between people's ideas about the world and the actual world. Festinger argues that humans are not well-equipped in case of major difference or dissonance between the two. While lower-level deviance between the two can be accommodated, dissonance reduction of more key elements works differently. One of Festinger's key insights is that, sometimes, if humans are really attached to a thought or concept, they do not change the concept if it collides with reality but rather alter (their perception of) reality (cf. ibid). For the Nigerian Scam, this means that victims, after they have convinced themselves that their online contact is a trustworthy person, show the tendency to rather alter their mental image of reality instead of changing their mental concept about their contact. This allows the scammers to perpetrate the fraud even longer and in part explains the potency of the consistency nudge as well.

From a rational choice perspective, it could be assumed that the accounting of losses and gains is similar in nature. The utility of a monetary unit gained is numerically the same as the negative utility of a unit lost. However, empirical research has shown that this is not the case. Losses are accounted at roughly twice the rate as gains (cf. Tversky/Kahneman 1991), a concept which should later be known as *loss aversion*. Loss aversion likely has its roots in evolution as losses threatened the survival of an individual or group much more than gains could guarantee future survival. As the current population is the descendants of the surviving populations of the past – a population which acted loss aversive – this trait can be found deeply ingrained in almost all of us (cf. Kahneman 2012, 435). Loss aversion alone would speak against the success of the Nigerian Scam. However, depending on whether humans situate themselves rather in the loss or gain spectrum, their sensitivity towards risk-taking changes as well. Kahneman and Tversky (cf. 1981, 453) tested this risk aversion by confronting two groups of participants with the same decision, yet differently framed. Participants were confronted with the scenario of a pandemic which infected 600 people and could end lethally. In scenario one, participants could choose between (A) the sure survival of 200 people or (B) a 33.3% probability that all 600 survive (and a 66.6% probability that all 600 die). The majority of participants (72%) choose option A. In a second trial, Kahneman and Tversky reversed the odds and confronted participants with either option (C) 400 people die or (D) a 33% probability that nobody dies and a 67% that all 600 decease – 78% choose option (D) over option (C), a complete reversal of the observations from the first trial. While this study underlines the potency of the framing effect, it also hints at another trait in human behaviour. Generally, humans are rather loss and risk aversive; however, when confronted with certain losses, the willingness to engage in high-risk scenarios increases (cf. Tversky/Kahneman 1992) – one explanation for high-stakes and high-risk gambling behaviour after initial losses. For the Nigerian Scam, this means that after the initial losses have been caused for the victim, its willingness

to invest in riskier endeavours increases. Just as a gambler who invests high stakes to win back his/her initial losses, the victims of the Nigerian Scam continue to pay higher and higher sums hoping to eventually get back the money they initially invested. Alongside the consistency nudge and the cognitive dissonance phenomenon, this change in risk aversiveness – caused by the victim's self-evaluation and situating in the loss or gain spectrum – reinforces the longevity and damages the Nigerian Scam can cause for victims.

The concepts presented in this section – the consistency nudge, cognitive dissonance, and modified risk aversion – partly overlap and partly cause each other. However, all four hint in the same direction: the continuation and escalation of the scam. The exploitation of repetitive and unreflective behaviour leads already victimized people into an even darker path of future victimization causing tremendous damages for individuals through the payment of multiple instalments. With regard to the initial observation of this paper – the procedure of target group optimization – a similar pattern can be recognized here. The majority of all victims gets away with moderate damages; however, few victims – led by the above-outlined effects and mechanisms – respond to the con financially and psychologically more potently and stack up tremendous amounts of damages.

The Sinker Phase: As outlined in the prior section, the victim and the fraudster continue the conversation as well as monetary transfers. They both participate in the same communication/signalling game; however, they play on different playing fields. While the victim tries to remain a good relationship with the fraudster – the victim still assumes that s/he will eventually be rewarded – the fraudster has a more realistic look at the game as s/he can opt-out and disappear with the money whenever the situation demands it. As only incorrect, stolen, or forged information and documents were shared with the victim, there is no chance that reimbursement or prosecution will ever take place. This asymmetrical power balance is contrasted by the illusion the fraudster has consciously and carefully constructed; while the

scammer emphasizes that s/he is dependent on the victim's goodwill, the tables have turned after multiple payments have been conducted. Often, scammers stop responding once unnegotiable problems arise. The naïve victim starts to research the presented narratives and finds out that s/he has been conned. Due to the feeling of shame, many of these frauds go unnoticed and do not appear in official statistics. Thereby, it can be assumed that the actual damage exceeds the official numbers by a substantial amount – a fact which again highlights the potency of the presented nudges and mechanisms. The sinker phase is characterized by the fact that the victim's illusion is disrupted by something (*i.e.* a piece of information, fact etc.) which causes the victim to eventually abandon their theories about their conversational partners – the end of the cognitive dissonance and often a tremendous shock.

3. Conclusion

As this paper could show, scammers apply a plethora of powerful nudges to persuade their potential victims. These communicative nudges are transferable to other areas of life and it is highly likely that they would – depending on the context – unfold their force there as well. Especially the nudges from the first section (authority, likability, scarcity, and reciprocity) are, at least in part, already used by marketers and advertisement specialists. Considering the expansion of information technology and the availability of texts as well as distribution channels, the potential to nudge audiences into a certain direction seems to be almost limitless. Thereby, it can be assumed that the modification of decision architecture will gain in importance in the future. Simultaneously, the Nigerian Scam also highlighted the advantages of applied Nudge theory for other settings: it is cheap, effective, and requires quasi no change in settings and/or procedures.

REFERENCES

Bergiel, B. J., Bergiel, E. B., & Balsmeier, P. W. (2008). Internet Cross Border Crime: A growing Problem. *Journal of Website Promotion, 3*(3/4), 133-142.

Binmore, K. (2013). *Spieltheorie/Game Theorie - A very short introduction*. Stuttgart: Phillip Reclam Jun. GmbH.

Chawki, M. (2009). Nigeria Tackles Advance Fee Fraud. *Journal of Information, Law & Technology (JILT), 1*, 1-20.

Cialdini, R. B. (1987). Compliance Principle of Compliance Professionals. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social Influence: The Ontario Symposium* (Vol. 5, pp. 165-184). Psychology Press.

Cialdini, Robert B. (1999). *Influence* (Rev. ed.). New York: Quill.

Cialdini, R. B. (2003). The Power of Persuasion - Putting the Science of Influence to work in Fundraising. *Stanford Social Innovation Review,* 17-27.

Cialdini, R. B., & Goldstein, N. J. (2002). The Science and Practice of Persuasion. *Cornell Hotel and Restaurant Administration Quarterly,* 40-50.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior, 46*, 260-281.

Dyrud, M. A. (2005). "I Brought You A Good News". An Analysis of Nigerian 419 Letters. *Proceedings of the 2005 Association for Business Communication (Annual Convention)*.

Edelson, E. (2003). The 419 Scam: Information Warfare on the spam front and a proposal for local filtering. *Computer and Security, 22*(5), 392-401.

Festinger, L. (1957). A theory of cognitive dissonance. Stanford, Calif: Stanford University Press.

Gladwell, M. (2005). *Blink - The Power of Thinking without Thinking*. New York, NY: Little, Brown and Company.

Guadagno, R. E., & Cialdini, R. B. (2010). *Preference for consistency and social influence: A review of current research findings.* Social Influence, 5(3), 152–163.

Haltermann, I. (2012). The perception of natural hazards in the context of human (in-)security. In U. Luig (Ed.), *Negotiating disasters: Politics, Representation, Meanings*. Frankfurt am Main: Peter Lang.

Hammerstein, P. (2002). Why is Reciprocity so rare in social animals? In P. Hammerstein (Ed.), *Genetic and Cultural Evolution of Cooperation* (pp. 83-93). Cambridge: MIT Press.

Harford, T. (2007). *The Undercover Economist* (Vol. 32). London: Abacus.

Herley, C. (2012). Why do Nigerian Scammers Say They are from Nigeria? *Microsoft Research: One Microsoft Way,* 1-14. Retrieved from https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/WhyFrom Nigeria.pdf.

Hobson, R. P. (2012). Emotion as Personal Relatedness. *Emotion Review, 4*(2), 169-175. doi:10.1177/1754073911430141

Kahneman, D. (2012). *Schnelles Denken, langsames Denken* (Vol. 8). München: Penguin Verlag/Random House.

Kahneman, D./Tversky, A. (1977). Prospect Theory. An Analysis of Decision Making Under Risk. doi:10.21236/ada04 5771

Kanich, C., Kreibich, C., Levchenko, K., Enright, B., Voelker, G. M., Paxson, V., & Savage, S. (2008). Spamalytics: An Empirical Analysis of Spam Marketing Conversion. *Proceedings of the 15th ACM Conference on Computer and Communications Security - CCS 08,* 1-12. doi:10.1145/1455770.1455774

Kich, M. (2005). A Rhetorical Analysis of fund-raiser-scam solicitations. *Cercles, 14*, 129-142.

Landon Jr., E.L. (1974). Self-Concept, Ideal Self Concept, and Consumer Purchase Intentions. *Journal of Consumer Research*. Volume 1 (2). 44–51.

Levine, T. R. (2014). Truth-Default Theory (TDT): A Theory of Human Deception and Deception Detection. *Journal of Language and Social Psychology, 33*(4), 378-392.

Levitt, S. D., & Dubner, S. J. (2016). *Think Like a Freak* (Vol. 1). München: Riemann Verlag/Goldmann.

Loomes, G./Sudgen, R. (1983): "A Rationale for Preference Reversal," American Economic Review, 73, 428-43

McNeil, B./Pauker, S./Sox, H., Jr./Tversky, A. (1982). On the elicitation of preferences for alternative therapies. New England Journal of Medicine, 306, 1259-1262.

Milgram, S. (1965). *Some Conditions of Obedience and Disobedience to Authority*. Human Relations, 18(1), 57–76.

Modic, D., & Lea, S. E. (2013). Scam Compliance and the Psychology of Persuasion. *SSRN Electronic Journal*.

Neuhaus, T. (2020). *Nudging Education – The (Potential) Role of Decision Architectures in Improving Educational Setting*. International Journal of Education and Research. Vol. 8 (9). 73-86.

Park, Y., Jones, J., McCoy, D., Shi, E., & Jacobsson, M. (2014). Scambaiter: Understanding Targeted Nigerian Scams on Craigslist. *NDSS Internet Society,* 1-15.

Peterson, J. B. (2018). 12 Rules For Life: Ordnung und Struktur in einer chaotischen Welt. München: Goldmann Verlag.

Petrova, P. K., Cialdini, R. B., & Sills, S. J. (2007). *Consistency-based compliance across cultures*. Journal of Experimental Social Psychology, 43(1), 104–111.

Sagarin, B. J., & Mitnick, K. D. (2012). The Path of Least Resistance. In D. T. Kenrick, N. J.

Goldstein, & S. L. Braver (Eds.), *Six Degrees of Social Influence - Science, Application, and the Psychology of Robert Cialdini* (pp. 27-38). Oxford: Oxford University Press.

Saini, J. R. (2012). Identification and Analysis of Most Frequently Occurring Significant Proper Nouns in 419 Nigerian Scam. *National Journal of Computer Science and Technology, 4*(1), 1-4.

Schaffer, D. (2012). The Language of Scam Spams: Linguistic Features Of "Nigerian Fraud" E-Mails. *etc - A Review of General Semantics, 69*(2), 157-179.

Smith, A. (2009). Nigerian Scam E-Mails and The Charms of Capital. *Cultural Studies, 23*(1), 27-47. doi:0.1080/09502380802016162

Spence, M. (1973). Job Market Signalling. *The Quarterly Journal of Economics, 87*(3), 355-374.

Stajano, F., & Wilson, P. (2011). Understanding Scam Victim: Seven Principles for system security. *Communication of the ACM, 54*(3), 70-75.

Straßheim, H./Jung, A./Korinek, R. L. (2015). Reframing Expertise: The Rise of Behavioral Insights and Interventions in Public Policy. In A. B. Antal, M. Hutter, & D. Stark (Eds.), *Moments of Valuation*. Oxford: Oxford Publication Press.

Thaler, R. H./Sunstein, C. R. (2017). *Nudge - Wie man kluge Entscheidungen anstößt* (Vol. 7). Berlin: Ullstein/Econ Verlag.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 184*, 1124-1131.

Tversky, A., & Kahneman, D. (1979). The Framing of Decisions and the Rationality of Choice. *Engineering Psychology Programs, Office of Naval Research,* 1-28.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, 210*, 453-458.

Tversky, A., & Kahneman, D. (1986). Rational choice and the Framing of Decisions (pp. 1– 47). Office of Naval Research.

Tversky, A./Kahneman, D. (1991): Loss Aversion in Riskless Choice: A Reference-Dependent Model, The Quarterly Journal of Economics, Volume 106, Issue 4, Pages 1039–1061

Tversky, A./Kahneman, D. (1992): Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertainty 5, 297–323.

Tversky, A. (1975). A Critique of Expected Utility. Erkenntnis, 9(2).

Von Neumann, J./Morgenstern, O. (1944). Theory of games and economic behaviour. Princeton: Princeton University Press.

Wainwright, T. (2017). *Narconomics: How to run a drug cartel*. London: Ebury Press.

Zahavi, A. (1997). A reply to John. *Animal Behaviour, 54*, 228-229.