You’ve got Mail! Explaining individual differences in becoming a phishing target

De Kimpe, Lies\textsuperscript{a*}, Walrave, Michel\textsuperscript{a}, Hardyns, Wim\textsuperscript{b}, Pauwels, Lieven\textsuperscript{b} and Ponnet, Koen\textsuperscript{ac}

\textsuperscript{a} Department of Communication Studies (MIOS), University of Antwerp, Antwerp, Belgium
\textsuperscript{b} Department of Criminology, Criminal Law and Social Law, Ghent University, Ghent, Belgium
\textsuperscript{c} Department of Communication Studies (mict), Ghent University, Ghent, Belgium

\textsuperscript{*}Corresponding author: Lies De Kimpe, Sint-Jacobsstraat 2, 2000 Antwerp – Belgium, lies.dekimpe@uantwerp.be, +32 655018

\textbf{Funding}

The authors gratefully acknowledge support of the Research Fund of the University of Antwerp. The interpretation of the data, the writing of the article and the decision to submit the article to Telematics and Informatics were the sole responsibility of the authors and were not influenced by the funding institution.
You’ve got Mail! Explaining individual differences in becoming phishing target

Abstract

Although phishing is a form of cybercrime that internet users get confronted with rather frequently, many people still get deceived by these practices. Since receiving phishing e-mails is an important prerequisite of victimization, this study focuses on becoming a phishing target. More precisely, we use an integrative lifestyle exposure model to study the effects of risky online routine activities that make a target more likely to come across a motivated offender. Insights of the lifestyle exposure model are combined with propensity theories in order to determine which role impulsivity plays in phishing targeting. To achieve these objectives, data collected in 2016 from a representative sample ($n = 723$) were used. Support was found for a relationship between both online purchasing behavior and digital copying behavior, and phishing targeting. Moreover, a relationship was found between all online activities (except for online purchasing behavior) and impulsivity. The present study thus suggests that especially online shoppers and users who often share and use copied files online should be trained to deal with phishing attacks appropriately.

Keywords: cybercrime; phishing targeting; lifestyle exposure model; impulsivity; online purchasing; digital copying
1. Introduction
In today’s society, the internet has become an integrated part of individuals’ lives. Also cybercriminals found ways to profit from the internet’s characteristics. Since it is easy to disguise one’s identity online, there are cybercriminals who pose as a trusted entity in order to deceive unsuspecting internet users into disclosing personal information (e.g., passwords, credit card details), which is called phishing (Lastdrager, 2014). Although phishing occurs rather frequently, internet users still seem highly susceptible to the deceiving messages and the fraudulent websites the cybercriminals create. These misjudgments can result in, amongst others, considerable financial losses or identity theft.

In the past few years, phishing victimization has increasingly gained attention as a research topic and several issues surrounding this subject have been addressed. For instance, previous studies have focused on the process of deception detection in phishing e-mails by users and by specialized software (Khonji et al., 2013; Wright et al., 2010). Also, the characteristics of phishing victims have been analyzed (Alseadoon, 2014; Halevi et al., 2015) and anti-phishing training, interventions and other tools have been evaluated (Kirlappos and Sasse, 2012; Purkait, 2012; Sheng et al., 2010). The present study adds to the existing literature by focusing on ‘phishing targeting’ (Reyns, 2015), which is an important prerequisite of victimization. Since sending only three phishing e-mails gives phishers a more than 50% chance of at least one click (Verizon, 2013), it could be argued that people receiving a lot of phishing e-mails are at higher risk of becoming a victim. More insight into the topic of targeting is thus required. More particularly, the characteristics that may influence internet users’ likelihood of receiving phishing e-mails are in need of further investigation. To identify dispositional and experiential factors related to phishing targeting, this study will test an integrated model, with key components derived from the lifestyle exposure model (Hindelang et al., 1978) and the routine activity theory of general deviance (Osgood et al., 1996). One of the key premises of this framework states that unstructured routine activities such as ‘hanging out on the street’ are conducive to crime. However, since this framework has hardly been applied in an online context, this study will investigate the relationship between several unstructured online activities (e.g., downloading and sharing files, risky online self-disclosure) and phishing targeting. Moreover, the extended lifestyle exposure model will be combined with the insights from propensity theories, which
stress the importance of impulsivity in predicting risk taking behavior. In sum, this study will integrate important criminological theoretical frameworks in order to predict phishing targeting.

2. Phishing

The term ‘phishing’ can be defined as “a scalable act of deception whereby impersonation is used to obtain information from a target” (Lastdrager, 2014, p. 8). Just like fishers, phishers use bait to increase the chances that their target will “bite”. This bait most frequently consists of e-mails seemingly sent by a trusted entity, for example a bank (Purkait, 2012; Purkait et al., 2014; Reyns, 2015). These e-mails often inform the user that a problem occurred that can only be solved if the e-mail’s recipient confirms some personal information (Hinde, 2004; Wright and Marett, 2010), or they promise tempting offers in exchange for personal details like a user ID or password (Hinde, 2004). Usually phishing e-mails do not ask for a direct reply, but contain a link to a fraudulent website, which is ‘the hook’ in the fishing metaphor. This website is very similar in look and feel to the official website it impersonates (Alseadoon et al., 2012; Hinde, 2004; Purkait et al., 2014; Wright and Marett, 2010). A recent report based on real-world online security incidents indicates that around 1 in 14 targets get successfully phished, either because they clicked the link or opened an attachment in a phishing e-mail (Verizon, 2017). Especially young people, between the ages of 18 and 25 seem to be a vulnerable target group (Sheng et al., 2010). An experimental study among university students shows that after two waves of attacks up to 83% of the young targets click the link mentioned in a phishing e-mail (Vishwanath, 2015a). Successful phishing attacks often result in identity theft and subsequently in financial gains for the offender. However, Purkait et al. (2014) stress the fact that money is not always the main objective for phishers. The collected information can also be used to harm the reputation of an individual or company, for example by spreading some controversial statements on behalf of another person. Besides identity, phishers can also steal intellectual property (Wright and Marett, 2010) or customer information from businesses (Hong, 2012).

A recent study by Graham and Triplett (2016) indicates that more than 30% of adults has received phishing e-mails in the past. Within a student population, this percentage even surpasses 50% (Ngo and Paternoster, 2011). It is thus safe to say that receiving phishing attempts via e-mail is a rather common phenomenon. This might be due to the specific characteristics of the internet. Wall (2007, p.70) notes that it is not difficult for a criminal to choose between either
planning a million dollar bank robbery or performing millions of $1 robberies in an online environment with a lot less risk and trouble. To obtain people’s e-mail addresses or complete e-mail lists, it suffices to just carry out a well-aimed Google search (Alazab and Broadhurst, 2015; TrendMicro, 2012). Once digital contact details are collected, phishers can start spreading mass e-mails to all addresses at once or carry out spear-phishing attacks to deceive specific victims with personalized e-mails (Alazab and Broadhurst, 2015; Hong, 2012).

The limited amount of risk related to performing a phishing attack stands in contrast to the serious losses and harm caused by these practices. Financial Fraud Action UK, for example, claims that between January and June of 2017 alone, a total of 366.4 million pounds was lost due to financial fraud. These losses are closely linked to personal and financial details stolen through online attacks and impersonation scams (i.e., phishing) (FFA UK, 2017). Moreover, recent reports by the Anti-Phishing Work Group (APWG) (2016) show that the amount of unique phishing sites detected was never higher than in the first half of 2016. Although most fraudulent websites are taken down once their illegitimacy has been confirmed, new ones are created every day (Kirlappos and Sasse, 2012). At the beginning of January 2018, more than 26,000 valid phishing web sites could still be found online and active (Phishtank, 2018). This proves that phishing is a still pressing problem.

2.1. Characteristics of the phishing message

Given the high amount of people that get victimized by phishers and the losses connected to this kind of online deception, researchers have tried to gain more insight into the characteristics of phishing e-mails that determine whether or not they will be successful. Two studies by Jakobsson (2007) indicate that e-mails offering a monetary price or asking for a password are more easily assessed as “phishy”, while messages only containing information (e.g., about an alleged security update) are more likely to be perceived as safe. This might pose a problem, as these apparent trustworthy e-mails might just as well contain a link to a phishing website. Further, mails containing spelling mistakes or unprofessional design tend to raise people’s suspicion (Furnell, 2007; Jakobsson, 2007). When phishers succeed in convincing the receivers that the mail is authentic, the next step is to persuade the recipient that sharing personal information is required. Here, social engineering strategies that have proven to be effective are ‘liking’ (i.e., pretending to be a person, organization or company the recipient likes and trusts)
(Jagatic et al., 2007; Wright et al., 2014), ‘reciprocity’ (i.e., giving people the impression they have to return a favor), ‘social proof’ (i.e., claiming other people have shared their personal details as well), ‘scarcity’ (i.e., giving the impression that an opportunity is limited) (Wright et al., 2014) and ‘authority’ (i.e., pretending to be an authority figure) (Butavicius et al., 2015).

2.2. Characteristics of the phishing target

Studies on fraud victimization in the past have also tried to gain more insight in the sociodemographic characteristics of victims (Holtfreter et al., 2008; Sheng et al., 2010; Titus et al., 1995). It soon became clear however, that creating demographic profiles for fraud victims is a complex undertaking (Holtfreter et al., 2008). In addition, this focus offers little insight into why people with certain demographics are more likely to become victimized and targeted (Pratt et al., 2010). Consequently, dispositional factors have been taken into account as well when fraud and phishing victims are studied. Research shows that individuals with high trust (Wright et al., 2010) and high submissiveness (i.e., the tendency to comply when faced with authority) tend to be more susceptible to phishing e-mails (Alseadoon et al., 2012). Higher susceptibility in turn, positively predicts responding to phishing e-mails (Alseadoon, 2014). Also high conscientiousness (i.e., the tendency to be dependable and hardworking) is claimed to be positively correlated with phishing victimization (Halevi et al., 2015), while ‘suspicion of humanity’ (i.e., the general idea that people do not have good intentions) has been linked to a decrease in phishing victimization (Wright and Marett, 2010).

Similarly, victims’ experience with computers and e-mails has been taken into account within phishing research. Several studies show that individuals with more internet experience (Wright et al., 2010; Wright and Marett, 2010) or technological knowledge (Downs et al., 2007; Sheng et al., 2010) are less susceptible to phishing. It could be assumed that habitual internet users have more experience with detecting inconsistencies in e-mails. However, a study by Pattinson et al. (2012) indicates that only for people who were aware that they participated in an experiment on phishing, familiarity with computers had a significant effect on how they managed phishing e-mails. This might implicate that even experienced internet users need a constant reminder of the risks they face. They might thus not be better equipped against potential harm at all time. This argument is supported by the reasoning of the criminological routine activities/lifestyle exposure perspective. This approach, which is based on the early work of
Cohen and Felson (1979) and Hindelang, Gottfredson and Garofalo (1978), claims that specific online activities are conducive to online victimization (Reyns et al., 2011). From this point of view, it is argued that avid internet users might in fact have a bigger, instead of a smaller chance to become a phishing target and/or victim, as it is more likely for frequent internet users to come across an online offender.

The present study builds on this reasoning to gain more insight into the relationship between different online routine activities and phishing targeting. More specifically, we will examine the link between becoming a phishing target and four different types of risky online exposure, which each stress a different aspect of the way in which internet users make themselves more visible and accessible online to phishers. First, (1) exposure to potentially illegal and/or infected files is considered, by examining people’s digital copying behavior. Also, we look at (2) risky disclosure of personal information on the internet in general (i.e., risky online self-disclosure), and on (3) social network sites (SNS) in particular (i.e., SNS use). Moreover, we take into account users’ (4) financial disclosure by looking at their online purchasing behavior. By taking these four risky routine activities into consideration, the lifestyle exposure framework can be adequately translated to study risk behavior in an online context and can be tested through a diverse range of exposure types.

The ideas of the lifestyle exposure approach will be combined with insights from propensity theories, that stress the importance of individual characteristics such as impulsivity (e.g., Gottfredson and Hirschi, 1990; Lahey et al., 2008; White et al., 1994) in the explanation of offending. Propensity theories argue that impulsive people self-select in risky routine activities and therefore increase their likelihood of becoming involved in cybercrime offending. Impulsivity has also been used to explain victimization, and cyber victimization in particular (Bosker and Holt, 2010; Ngo and Paternoster, 2011). However, studies have shown that there is no direct relation between (phishing) targeting and impulsivity (Holtfreter et al., 2008; Ngo and Paternoster, 2011). Therefore, by using an integrated model as shown in figure 1, this study will link impulsivity with phishing targeting in a more indirect manner. The theoretical frameworks this model builds upon, will be discussed in the remainder of the literature overview below.

[insert figure 1 about here]

3. Theory and hypotheses development
3.1. An integrative lifestyle exposure theory of general deviance and victimization

The framework the present study applies is inspired by the lifestyle exposure model of Hindelang et al. (1978) which has since been updated, extended and applied to explain individual differences in both victimization and offending (Miethe and Meier, 1994). According to the original lifestyle exposure theory of criminal victimization, differences exist in people’s leisure, vocational and professional activities, which is due to differences in their background characteristics (e.g., age, race, income). These characteristics are linked to differences in role expectations (i.e., what to expect in life) and structural constraints (i.e., what can be achieved in life). The variations in the structures individuals are part of, are linked to differences in lifestyles. In this theoretical model, a (risky) lifestyle refers to one’s daily routine activities (author 1). The lifestyle exposure model stresses the importance of lifestyles, since they can be linked to differences in exposure to environments that are conducive to crime (Meier and Miethe, 1993).

According to Osgood et al. (1996, p. 640), especially unstructured activities, or activities “that carry no agenda for how time is to be spent” (e.g., ‘going to a party’ or ‘hanging out on the street’) conducted with peers in the absence of authority figures or social control, are conducive to crime. Performing these unstructured activities or risky behaviors increases the likelihood of encountering offenders, who hang around in the same settings (Hoeben and Weerman, 2014). Engaging in risky routine activities thus increases the likelihood of victimization, but at the same time might also lead to situations where individuals become offenders themselves (Meier and Miethe, 1993; Ngo and Paternoster, 2011; author 1). While this theoretical framework has been empirically corroborated several times (Hoeben and Weerman, 2014; Osgood et al., 1996; author 2), it is surprising to see that it has hardly been applied in the context of online offending and victimization. Therefore, the present study will take into account a diverse range of risky online activities that make internet users more identifiable online and examine how these are related to phishing targeting. We take into account digital copying behavior, risky online self-disclosure, SNS use and online purchasing behavior.

The characteristics of the internet have made it easy for internet users to make copies of files and to use and share this content with other internet users. This behavior is legal when copies of software or multimedia content are made for private use. However, this is no longer the case when files are distributed more widely, for example when files are uploaded or downloaded through peer-to-peer (p2p) file sharing networks. These networks were estimated to grow at a
rate of 26% in the period between 2011 and 2016 (Cisco, 2012). Moreover, in the first three months of 2017 alone, almost 1 in 6 of all internet users in the UK (age 12 or older) had consumed illegal content online (Intellectual Property Office, 2017). This type of risky behavior is thus rather common and might increase the risk of becoming a phishing target. Multiple studies show that involvement in cyber deviance is significantly related with the risk of cyber victimization (Bossler and Holt, 2009; Holt and Bossler, 2008; Ngo and Paternoster, 2011; Reyns et al., 2011; van Wilsem, 2011). Therefore, it can be assumed that internet users involved in copyright infringement are at higher risk of getting victimized, because they expose themselves to risky online environments that are also used by other internet users who perform online deviant behavior. Moreover, it can be argued that even internet users who often legally use, copy and share files, might be at higher risk of getting victimized. As they more often open and/or download files, changes increase that they are confronted with illegitimate or infected files that are spread by online offenders. Therefore, we hypothesize (H):

*Individuals who often use, copy or share digital files (i.e., digital copying behavior) are more likely to become phishing targets (H1), since this behavior increases exposure in risky online settings.*

Not only can it be considered risky to share or use copied files with other internet users, also sharing personal information with others in a too generous way might increase someone’s victimization risk. For example, adding strangers as a friend to one’s online social network profile is significantly related with online victimization, as well as providing personal information on a SNS (Henson et al., 2011; Marcum et al., 2010). The same holds true for having one’s personal information posted online (Reyns, 2015). These risky forms of online self-disclosure make individuals more visible for cybercriminals and may give phishers more ammunition to create personalized scams. We therefore expect that:

*Risky online self-disclosure is positively related to phishing targeting (H2), because this increases the likelihood of sharing sensitive information with criminals.*

Personal information published on SNS is easily traceable. A study by the Pew Research Center indicates that 20% of SNS users have a completely public profile (Madden, 2012). These pieces of information might be used to contact targets via e-mail in a more personalized way, but in addition, potential victims can be reached directly through the social network platforms
themselves. On Facebook for example, 1.5% of the profiles is marked as ‘undesirable’, as they send spam, infected links and other unwanted content to the users of the platform (Cluley, 2012). Once a potential target accepts a friend-request sent by a fake user, the instant messaging system of the SNS is used to contact the target in order to collect personal details (Vishwanath, 2015b). Visiting and using SNS is thus not without risk. These platforms can be used by phishers to find personal information of potential victims, or to request additional information from them in a very personal and direct way. This leads us to the following hypothesis:

*People who use social network sites more often, have a higher likelihood of getting targeted by phishers (H3) because this online activity increases presence in settings that make them an easy target.*

Another risky routine activity that should be taken into consideration is online purchasing. The internet has become a popular way to buy and sell products. Today, almost eight out of ten Americans say they make online purchases, and 43% shop online at least a few times a month (Smith and Anderson, 2016). These new purchasing habits however, serve as an extra opportunity for online fraudsters (Holtfreter et al., 2008). For example, numerous fraudulent websites sell imitations of brand name articles or overprized concert tickets (Kirlappos and Sasse, 2012), and non-delivery (i.e., victims paying for an article or service but never receiving it) is one of the most frequently reported online offences according to the 2015 Internet Crime Report (Internet Crime Complaint Center, 2015). It is thus rather easy to come across online fraudsters while online shopping. Even if internet users don’t buy anything, but just share some personal information via these websites (e.g., to make an account), contact information can be used to approach targets and to phish for other delicate information. Because it is likely that those who shop online more often also come across fraudulent retailers more often, we hypothesize that:

*Internet users who make online purchases more often, are more likely to become a phishing target (H4), since this online activity increases exposure to online risky settings.*

### 3.2 Propensity theories and impulsivity

Besides being linked to risky routine activities, there is also evidence that the famous victim-offender overlap is associated with impulsivity (Schreck et al., 2006). Therefore, the present
study will integrate impulsivity into the proposed model (cf. figure 1). Gottfredson and Hirschi (1990) pioneered criminological research into impulsivity, criminal involvement and victimization. Their general theory of crime in its original form stated that an important factor in explaining criminal behavior is low self-control. When individuals have a lack of self-control, they are impulsive, not able to contain oneself and less inclined to consider the consequences of their behavior. Empirical tests of self-control theory have since demonstrated that impulsivity and thrill-seeking are the most important dimensions that are linked to offending and deviant behavior (Pratt and Cullen, 2000; Vazsonyi et al., 2017). Although the theory was originally developed to explain offending, it has been suggested that it has the potential to serve as a predictor for criminal victimization as well (Bossler and Holt, 2010; Ngo and Paternoster, 2011; Schreck et al., 2006). However, most studies are restricted to offline and traditional types of deviant behavior.

Given this clear link between impulsivity and imprudent behavior in an offline context (Forde and Kennedy, 1997), it can be assumed that impulsivity is also linked with the performance of online risk behaviors. For example, studies show that a lack of self-control is related to a variety of copying behaviors, such as movie piracy (Higgins et al., 2007), illegally uploading (Donner et al., 2014) and downloading files (Donner et al., 2014; LaRose et al., 2005). Therefore, we expect that there is a relationship between digital copying behavior and impulsivity. Furthermore, it can be assumed that there is an association between the use of SNS and impulsiveness. Those who use social media excessively, tend to have problems with spending their time effectively or to plan ahead. These are characteristics they share with impulsive individuals (Savcı et al., 2016). That’s probably why individuals with high impulsive tendencies use SNS more often (Wu et al., 2013). Given the link between SNS and online self-disclosure, we expect that also a relationship exists between impulsivity and risky online self-disclosure. Furthermore, research has demonstrated that impulsivity is associated with the decision-making process of consumers, both offline and online (Huang and Kuo, 2012). Consequently, online buying impulsiveness has been scrutinized several times (Chen and Lee, 2015; Huang and Kuo, 2012; LaRose and Eastin, 2002). Although these studies focus on when and how impulsive purchasing occurs, it is to be expected that a relationship between impulsivity and online purchasing behavior exists. Therefore, we expect:
Impulsivity is positively associated with digital copying behavior (H5a), risky online self-disclosure (H5b), SNS use (H5c) and online purchasing behavior (H5d).

In sum, this study will examine to which extent several unstructured routine activities explain getting targeted by phishers. These online activities in turn are expected to be influenced by internet users’ level of impulsivity.

4. Methodology
4.1. Participants and data collection
This study draws upon data from the interuniversity Social Capital in Neighborhoods (SCAN) project in which 819 respondents living in 41 neighborhoods of Ghent (Flanders, Belgium) participated. The SCAN is a yearly administered questionnaire that is part of an interuniversity cooperation between the University of Antwerp (MIOS, department of Communication Studies) and Ghent University (department of Criminology). During home visits in October and November 2016, face-to-face interviews were conducted using a structured questionnaire on online and offline social capital, health and risk behaviors.

The sampling design is based on a design applied by author 3. A sample of inhabitants from each neighborhood was selected based on the municipal registry of 2012. This sample was representative of the composition of each neighborhood and stratified by sex (male versus female), age (18–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75+) and nationality (Belgian versus non-Belgian). Moreover, for every inhabitant in the sample, three substitutes with the same sex, age and nationality were randomly selected. The backup respondents could be contacted after three unsuccessful home visits to the selected inhabitant, after a refusal to participate from the selected respondent or when the respondent did not meet the inclusion criteria (i.e., minimal age of 18, sufficient knowledge of the Dutch language and not residing in an institutional setting). When the interviewers ran out of substitutes, random inhabitants living in the same neighborhood were contacted. This happened in 29.4% (n = 241) of the cases. This rather high rate might be linked to the partial mismatch that existed between the data from the municipal registry of 2012 and the situation in the year 2016.

Of all respondents, 88.3% (n = 723) indicated they use the internet to look up information and/or to purchase goods or services. Given the focus of our study, those who did not use the
internet for either of those purposes were left out of further analysis as it can be assumed that these respondents do not use the internet for other activities either. Of the remaining 723 participants, 48.0% male and 52.0% female respondents between the age of 18 and 92 (M= 48.20; SD= 16.71) were interviewed. The majority of the respondents (62.66% or n = 443) has a university or college degree, 28.43% (n=201) graduated from high school, 5.95% (n= 42) completed the first three years of high school and the remaining 2.97% (n = 21) went to primary school or is not educated.

4.2. Measures
Note that the exact formulation of all items used to measure the following variables can be found in table 1.

Impulsivity
To measure impulsivity, items from the self-control scale developed by Grasmick et al. (1993) were used. This scale consists of five items that measure impulsivity (e.g., “If I can have fun I will, even though I will be in trouble later on”). Each item was scored on a five-point Likert scale ranging from totally disagree (= 1) to totally agree (= 5). Reliability analysis showed that the impulsivity scale was reliable (α = .76).

Digital copying behavior
To operationalize digital copying behavior, two items based on a measure of online copyright infringement were used (Bosler and Holt, 2010), which were then made broader applicable to all users who ever used, copied or shared (1) copies of official software and/or (2) copies of media files (e.g., music, films, games). Both items were measured using a 5-point scale ranging from never (= 1) to very often (= 5). The internal reliability proved to be good (α = .85).

Risky online self-disclosure
This concept was measured using two of the original five items used by Livingstone, Haddon, Görzig and Ólafsson (2011) to measure children’s actions in relation to online contacts. Respondents were asked to indicate to which extent risky forms of online self-disclose (e.g., “I have sent personal information to someone that I have never met face-to-face”) were performed in the last six months on a 5-point scale (never (= 1) to very often (= 5)). Reliability analysis
indicated the scale has a Cronbach’s alpha of .48, which is rather low. However, this is not uncommon, because the Cronbach’s alpha almost every time underestimates the true reliability of two-item scales (Eisinga et al., 2013).

**Social Network Site use**

A single item was used to measure the frequency of SNS use. Nine answering options (never (= 1), monthly (= 2), weekly (= 3), several days a week (= 4), once every day (= 5), 2 to 3 times a day (= 6), 4 to 5 times a day (= 7), 6 to 7 times a day (= 8) and more than seven times a day (= 9)) were offered to indicate how often one generally visits SNS. People without a profile on SNS (n = 212 or 29.3%), were not asked to answer this question.

**Online purchasing behavior**

Respondents were asked to indicate how often they use the internet to purchase goods or services. Again, nine answering categories were available (never (= 1) to more than seven times a day (= 9)).

**Phishing targeting**

Our outcome variable phishing targeting was measured by a single question from a study of Reyns (2015). Respondents were asked to which extent they ever were tempted by fraudulent e-mails and/or websites into sharing bank account details, passwords or other personal information with the offender. To ensure an unambiguous understanding of the question, the respondents were additionally informed that these e-mails and websites are usually sent by someone posing as a trustworthy organization, such as one’s bank or employer. A five-point Likert scale was used, anchored by never (= 1) and very often (=5).

**Control variables**

Sex (male (=0) and female (=1)), age, educational attainment and general internet use were included in our model as covariates. The age of the respondents was measured using seven age categories, namely 18–24 (=1), 25–34 (=2), 35–44 (=3), 45–54 (=4), 55–64 (=5), 65–74 (=6), 75+ (=7). To measure the highest educational attainment, respondents were offered four options: no education/primary school (=1); first three years of high school (=2); high school (six or seven years) (=3) and higher education (=4). To assess how much time the respondents spend
online in general, a single item was used to measure how often they use the internet to look something up or to look for information, as the internet is in the first place a gateway to information. Nine answering categories were offered (never (= 1) to more than seven times a day (= 9)).

4.3. Data Analysis

Structural equation modelling (SEM) was used to examine the hypothesized relationships between the components of our model. By means of Mplus 7.4. (Muthén and Muthén, 2012), first a measurement model was built to verify whether the observed variables were a reliable measure of the latent variables. Then, the structural model was tested with impulsivity as independent variable, digital copying behavior, risky online self-disclosure, SNS use and online purchasing behavior as endogenous variables, phishing targeting as our outcome variable and age, sex, educational level and frequency of internet use as covariates.

The fit of the model was estimated through multiple goodness-of-fit indices. Because the chi-square test is sensitive to sample size, its value is almost always significant (Byrne, 2012). Therefore, the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Squared Residual (SRMR) were taken into account as well. CFI ranges from 0 to 1.00. The closer to 1, the better the model fits. As a rule of thumb, .90 is often used as a cut-off value (Hu and Bentler, 1999). The value of RMSEA should be kept as low as possible with values below .08 representing a good fit and values up to .10 indicating a mediocre fit (Byrne, 2012). The SRMR ranges from 0 to 1, with values close to 0 indicating a good-fitting model (Byrne, 2012). More specifically, a good model fit is indicated when the SRMR is smaller than .08 (Hu and Bentler, 1999).

5. Results

5.1. Preliminary Analyses

Our preliminary analyses showed that of all internet users in our sample, 51.3% (n = 371) have been the target of phishing attacks in the past, although 19.4% (n = 140) claimed they get targeted ‘seldom’. Another 19.4% (n = 140) of people indicated they ‘sometimes’ receive phishing messages. The remaining 12.6% (n = 91) gets targeted often or very often. Other descriptive results can be found in Table 1.
The correlations among the constructs of our research model are shown in table 2. A significant positive correlation was found between impulsivity and the risky routine activities ($p < .01$; except for the correlation between impulsivity and online purchasing: $p < .05$). Also, the correlations between the online activities and phishing targeting were significant and positive ($p < .01$).

5.2. Measurement Model

First, we assessed the measurement model, with impulsivity, digital copying behavior and risky online self-disclosure as latent constructs. The measurement model indicated a good fit with the data: $\chi^2 (24) = 111.279 (p < .001)$, CFI = .946, RMSEA = .072 (CI: .059 - .086), SRMR = .048. All factor loadings were significant and above .52 (standardized values).

5.3. Structural Model

Subsequently, we determined whether age, gender, educational level and general internet use should be included as covariates in the analyses. The analysis revealed a number of significant associations among the variables considered. Educational level and gender were significantly related to digital copying behavior (respectively $\beta = .20$, $p < .001$; $\beta = -.19$, $p < .001$). Moreover, a significant association was found between educational level and impulsivity ($\beta = -.16$, $p < .01$) and between gender and the use of SNS ($\beta = .11$, $p < .01$). Therefore, these sociodemographic variables were included as covariates in the analyses. Note that the demographics were not regressed on the outcome variable phishing targeting, as we presume that sociodemographic factors are mostly unknown and of little importance to phishers. Moreover, significant relationships were found between general internet use and digital copying behavior ($\beta = .33$, $p < .001$), risky online self-disclosure ($\beta = .32$, $p < .001$), SNS use ($\beta = .41$, $p < .001$) and online purchasing behavior ($\beta = .32$, $p < .001$). Therefore, general internet use was also included as a covariate to the analyses.

The model fit indices proved that the structural model had an acceptable fit with the data: $\chi^2 (76) = 249.106 (p < .001)$, CFI = .891 and RMSEA = .068 (CI: 0.058 - 0.077), SRMR = .060. The analyses showed that the online activities online purchasing behavior ($\beta = .12$, $p < .01$) (H4)
and digital copying behavior ($\beta = .16, p < .01$) (H1) were significant and positive predictors of becoming a phishing target. Contrary to our expectations, risky online self-disclosure ($\beta = .11, p = .131$) (H2) and the use of SNS ($\beta = .02, p = .761$) (H3) were not significantly related to the outcome variable. The hypotheses regarding the relation between impulsivity and the online routines on the other hand, were confirmed (H5a to H5c), except for the relation between impulsivity and online purchasing behavior, which was insignificant ($\beta = .09, p = .054$) (H5d). A detailed overview of all estimates can be found in Figure 2.

6. Discussion and Conclusion

Phishers use the internet to create fake e-mail accounts and messages, in order to persuade their targets to disclose sensitive information. To prevent that people fall prey to such cons, a first step is to gain more insight into which specific internet users are more likely to become a target. With our integrated lifestyle exposure framework in mind, we hypothesized that risky online activities increase the likelihood that internet users come across phishers. More specifically, the relations between becoming a phishing target and four types of online exposure, namely digital copying behavior (H1), risky online self-disclosure (H2), SNS use (H3) and online purchasing behavior (H4), were examined. Moreover, the present study linked the insights from the extended lifestyle exposure model with those from propensity theories in order to investigate the relationship between impulsivity and the unstructured routine activities included in the study (H5).

First of all, our findings suggest that online purchasing behavior is related to phishing targeting. This result might imply that interacting with online retailers is not always a safe undertaking and exposes the internet user to some risk. Not only it is often unclear how businesses save and share the collected personal information, consumers should also be mindful of the fact that not every online retailer can be trusted. The link between online purchasing and phishing targeting however, should not be seen as an argument against online purchasing behavior, but rather as a warning sign for internet users who often buy products and/or services online. Especially these people should be actively trained in detecting phishing e-mails and phishing websites, as academics and other specialists in the field agree that security training is one of the most important countermeasures to tackle phishing (Jansson and von Solms, 2013). Easy accessible tools to train internet users already exist, such as the online game Phishing Phil,
that teaches users to discern phishing URL’s from legitimate ones (Sheng et al., 2007). It can be questioned however, if internet users spontaneously come into contact with such games during their daily online activities. Therefore, context-specific training, as suggested by Kirlappos and Sasse (2012), might be a better alternative. In the case of online shopping, context-specific training would imply that online retailers try to train their customers whenever they visit their website or purchase a product. For example, customers could be asked before paying to take part in a little quiz that checks if they can distinguish a phishing e-mail from a real e-mail sent by the retailer. By making their clients aware of the fact the difference is in the details, online retailers would contribute to an online environment where less of their own customers fall prey to phishers.

Second, our study suggests that a relationship exists between digital copying behavior and phishing targeting. Since a considerable amount of internet users are involved in copyright infringement (Intellectual Property Office, 2017), it might not surprise that these individuals are more visible and accessible to cybercriminals within online risky environments, such as p2p platforms. Given that previous studies have indicated that online offending and cyber victimization are closely related (Bossler and Holt, 2009; Ngo and Paternoster, 2011; Reyns et al., 2011; van Wilsem, 2011), it could be useful to raise the awareness of people involved in such practices about the more vulnerable position they put themselves in. However, it would be even more effective to discourage copyright infringement altogether. At the same time, it should be stressed that not all digital copying behavior is illegal. Still, our results suggest that copying, sharing and using copies of software and/or digital content increases one’s risk of becoming a phishing target. Also internet users often performing this type of online behavior in a legal way could thus benefit from increased awareness and following training. In addition, governments should also urge internet users to report the phishing attacks they experience. This is considered a key element in reducing the effectiveness of phishing (Verizon, 2017). For instance, setting up an online channel to report suspicious websites and e-mails anonymously would lower the threshold for users who are involved in copyright infringement to report these criminal practices (Europol, 2016).

Furthermore, this study indicates that it could be useful to apply the integrated lifestyle exposure model in an online context. Although this study only explains a limited amount of the
variance in phishing targeting, the results can still be interpreted as a confirmation of the link between risky daily routines and victimization in an online context. In addition, it is shown that there is a significant relationship between impulsivity and the online activities examined. This is an interesting contribution to the literature in this field, since it was up to now assumed that individual characteristics like impulsivity were not related to cyber victimization in situations where not a specific person, but any person and their computer could become the target of the online crime (e.g., in the case of phishing) (Bossler and Holt, 2010; Ngo and Paternoster, 2011). This study proves however, that impulsivity might be in fact related to phishing targeting, although in a more indirect way.

At the same time, it should be mentioned that the hypotheses presuming that SNS use and risky online self-disclosure would be predictive of phishing targeting, could not be confirmed in this study. A possible explanation for the lack of a significant relationship between SNS use and becoming a phishing target, might be that only the frequency of visiting SNS was taken into consideration. Perhaps phishing targeting has more to do with the privacy settings used and the specific amount of personal information shared on SNS platforms. Depending on how one’s settings are managed, exposure to motivated offender might differ. In the current study, the cautiousness by which internet users handle their personal details online in general was estimated by measuring risky online self-disclosure. However, again no significant relation with phishing targeting was found. This result could indicate that phishers prefer using large e-mail lists, for example found by hacking into databases of online retailers, rather than searching for individuals’ personal information on SNS or other websites. Therefore using SNS or being careless with personal information online might not be strongly related to phishing targeting.

7. Limitations

There is ample room left for the optimization of the proposed model, since only a limited amount of risky activities were considered and only a small part of the variance in phishing targeting was explained (7.2 %). The present study can serve as an encouragement to further expand our understanding of phishing targeting. Although we considered a diverse range of risky forms of online exposure, the four activities offer by no means a complete overview of all the risky online activities that could be linked to phishing targeting. Given that the data used for this study were part of the broader SCAN project, a limited amount of items and variables could be listed in the
survey. It might thus be interesting for future research to include a broader range of potentially risky online activities (e.g., hacking, online banking, online gambling). In addition, another limitation is that we did not consider internet users’ online skills, such as digital literacy, although some studies have provided evidence that these skills could serve as a form of personal guardianship that can protect internet users against phishing (Graham and Triplett, 2016). It might be interesting for future research to include digital literacy, as well as other online behaviors (e.g., frequency of e-mail use), online protective behaviors (e.g., privacy settings on SNS, the use of anti-virus software) and dispositional factors (e.g., morality). Moreover, the operationalization of some of the study variables could have been more elaborate (e.g., the scales measuring risky online self-disclosure and digital copying behavior only consisted of two items). The respondents who enrolled in this project were all aged 18 and older. In future waves of the project the online activities of minors (aged 16 to 18) will also be taken into account, since deviant behavior peeks during adolescence (Moffitt, 1993). Finally, it is important to mention that this study focused on targeting or attempt to victimize, which is not the same as actual victimization. Since the people who get targeted by phishers more often are not necessarily the same people who get victimized, it might be interesting for future research to take both phishing targeting and victimization into account, as Graham and Triplett (2016) did with relation to media literacy. Comparing the activities that predict targeting with those related to victimization, and how impulsivity plays a role in this, would result in a more detailed overview of the phishing process.
References


Alseadoon, I.M.A., 2014. The impact of users’ characteristics on their ability to detect phishing emails. Queensland University of Technology.


Cluley, G., 2012. Facebook: There are over 83 million fake accounts on our site. Naked Secur.


### Tables

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impulsivity (α = .76)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1 – I often do things without thinking it through.</td>
<td>2.32</td>
<td>.98</td>
</tr>
<tr>
<td>Item 2 – If I can have fun, I will, even though I will be in trouble later on.</td>
<td>2.35</td>
<td>.97</td>
</tr>
<tr>
<td>Item 3 - Sometimes I take risks for fun.</td>
<td>2.19</td>
<td>1.03</td>
</tr>
<tr>
<td>Item 4 – I speak my mind, even when that’s not a smart thing to do.</td>
<td>2.69</td>
<td>1.10</td>
</tr>
<tr>
<td>Item 5 – I often just do what I feel like doing immediately.</td>
<td>2.61</td>
<td>1.02</td>
</tr>
<tr>
<td><strong>Digital copying behavior (α = .85)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1 - I have copied, shared or used a copy of official computer software.</td>
<td>1.95</td>
<td>1.13</td>
</tr>
<tr>
<td>Item 2 – I have copied, shared or used a copy of music files, movies or games.</td>
<td>2.28</td>
<td>1.26</td>
</tr>
<tr>
<td><strong>Risky online self-disclosure (α = .48)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1 – I have added people to my contacts on social network sites (e.g., Facebook) whom I’ve never met in person before.</td>
<td>1.46</td>
<td>.85</td>
</tr>
<tr>
<td>Item 2 - I have sent my contact information (e.g., my full name, address or telephone number) to someone I’ve never met in person before.</td>
<td>1.33</td>
<td>.68</td>
</tr>
<tr>
<td><strong>Online purchasing behavior</strong></td>
<td>1.15</td>
<td>1.25</td>
</tr>
<tr>
<td><strong>SNS use</strong></td>
<td>4.47</td>
<td>2.34</td>
</tr>
<tr>
<td><strong>Phishing targeting</strong></td>
<td>2.03</td>
<td>1.17</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>48.20</td>
<td>16.71</td>
</tr>
<tr>
<td><strong>General internet use</strong></td>
<td>4.82</td>
<td>2.01</td>
</tr>
</tbody>
</table>

**Table 1.** Descriptives of the variables included in the study (n = 723)
**Table 2.** Correlations between the components of the research model. *p < .05; **p < .01; ***p < .001 (n = 723)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Impulsivity</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Digital copying behavior</td>
<td>.232***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Risky online self-disclosure</td>
<td>.228***</td>
<td>.332***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Online purchasing behavior</td>
<td>.093*</td>
<td>.257***</td>
<td>.164***</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 SNS use</td>
<td>.325***</td>
<td>.307***</td>
<td>.227***</td>
<td>.211***</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6 Phishing targeting</td>
<td>.017</td>
<td>.230***</td>
<td>.154***</td>
<td>.189***</td>
<td>.134**</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 1. Conceptual model of determinants of phishing targeting
Figure 2. Full model of determinants of phishing. Note: All reported coefficients are standardized values, adjusted for the influence of covariates. The dashed lines indicate non-significant paths. ** $p < .01$; *** $p < .001$. 
Highlights “You’ve got Mail! Determining Risk Factors of Becoming a Phishing Target“

- The integrated lifestyle exposure model is applicable in an online context.
- Individuals who make online purchases get targeted by phishers more often.
- Individuals who often share/use copied files get targeted by phishers more often.