UNIVERSITY OF TWENTE.

Faculty of Behavioral, Management and Social Sciences

The role of social projection in direct and indirect estimations of cybercrime among young adults

Antonia Wieschebrock
B.Sc. Thesis
June 2017

Supervisors:
S. Zebel
E.G. Ufkes

Department of Conflict, Risk & Safety
Faculty of Behavioral, Management and Social Sciences
University of Twente
7500 AE Enschede
The Netherlands

Abstract

Indirect estimations are used to examine the prevalence rates of criminal behavior by asking people how many percent of their group members they think have committed any form of cybercrime or digitized crime in the past. It is assumed that these estimations reduce the occurrence of social desirability. In the current study, a higher (compared to a lower) identification with the ingroup is expected to lead to a greater use of social projection in these indirect estimations, as people base the prevalence rates more on their own criminal behavior and experiences. Because of this process, indirect estimations would provide a clearer picture of actual prevalence rates. In turn, greater projection should predict higher indirect estimations. Therefore, the use of social projection was hypothesized to mediate the positive association between identification and indirect estimations that was found in previous research.

A between-groups design was employed with an experimental manipulation of group identification (high versus low). Participants were recruited among members of a sports group or among the first-year computer science students at the University of Twente and answered both indirect as well as direct estimations about different cybercrime offenses such as hacking and committing identity theft. In addition, the use of social projection in the indirect estimations was assessed.

The mediation hypothesis could not be supported because no association was found between the group identification manipulation and indirect estimations. Additionally, no consistent link between social projection and the indirect estimations was found either. A possible explanation of the unexpected findings could be that the characteristics of the groups that are used in indirect estimations play a significant role. Hypothetically, the more the identity content of the group is clearly defined, the less social projection would be used by participants, meaning that the proposed model could possibly still be applied while using larger, unclearly defined groups. Finally, it is possible that social desirability also plays a role in indirect estimations, when using small groups as in the current study. Consequently, it is important to further examine the conditions under which social desirability, social projection, and identification play a role in indirect estimations to determine the validity of this form of questioning in finding actual prevalence rates of cybercrime and digitized crime.

Abstract (Dutch)

Om de prevalentie van crimineel gedrag in te schatten worden indirecte schattingen gebruikt waarbij mensen gevraagd worden hoe veel percent van hun groepsleden in het verleden een delict van cybercriminaliteit of gedigitaliseerde criminaliteit gepleegd hebben. Er wordt ervan uitgegaan dat deze schattingen het optreden van sociale wenselijkheid verlaagd. In dit onderzoek wordt verwacht dat een hoge (in vergelijking tot een lage) identificatie met de ingroep leidt tot meer gebruik van sociale projectie in deze indirecte schattingen aangezien mensen de schattingen dan meer op hun eigen crimineel gedrag en ervaringen zullen baseren. Vanwege dit proces zullen indirecte schattingen een duidelijker beeld van de werkelijke prevalentie opleveren. Meer gebruik van sociale projectie zal dan een hogere indirecte schatting voorspellen. Het gebruik van sociale projectie werd veronderstelt om de positieve samenhang tussen identificatie en indirecte schattingen te mediëren.

Een “between-groups”-design was toegepast met een experimentele manipulatie van groepsidentificatie (hoog versus laag). Participanten werden verzameld onder leden van zowel een sportvereniging als ook eerstejaars informatica studenten van de Universiteit Twente. Zij beantwoordden indirecte vragen en directe vragen over verschillende cybercriminaliteit delicten zoals hacken. Bovendien werd het gebruik van sociale projectie in de indirecte schattingen gemeten.

De mediatie hypothese kon niet bevestigd werden omdat er geen samenhang tussen de identificatie manipulatie en de indirecte schattingen gevonden werd. Verder werd er ook geen consistente samenhang tussen de sociale projectie en de indirecte schattingen gevonden. De kenmerken van de groepen zouden een mogelijke verklaring voor de onverwachte resultaten zijn; een duidelijk gedefinieerde groepsidentiteit zou eventueel tot minder gebruik van sociale projectie kunnen leiden. Het veronderstelde model zou dan wellicht toegepast kunnen worden bij grotere, niet duidelijk gedefinieerde groepen. Ten slotte kan het mogelijk zijn dat sociale wenselijkheid ook een rol speelt in indirecte schattingen als er gebruik gemaakt wordt van kleinere groepen zoals in dit onderzoek. Als gevolg daarvan is het belangrijk om de condities waaronder sociale projectie, sociale wenselijkheid en identificatie een rol spelen in indirecte schattingen verder te onderzoeken. Op die manier zou de validiteit van deze methode bij het vinden van werkelijke prevalenties of cybercriminaliteit en gedigitaliseerde criminaliteit bepaald kunnen worden.

The role of social projection in direct and indirect estimations of cybercrime among young adults

Cybercrime is a serious, growing problem and it is therefore important to find actual prevalence rates of offenders in order to correctly deal with this problem. A way of defining the prevalence rate of cybercrime is to examine the opinion of people about how many other members in their social group have committed different forms of crimes (Fischer, 1993; John, Loewenstein & Prelec, 2012). An example of this indirect questioning method is: "How many percent of young adults do you think have hacked someone else in the previous year?". This method is an innovative method because an important problem with determining prevalence rates based on self-reports is social desirability: a tendency of people to represent themselves in a positive way and to avoid being judged negatively. Self-reports are a way of direct questioning, in which people are asked if they themselves committed a specific crime. In this context, social desirability leads them to deny having committed the crime, due to the fear of not being accepted by the others and society as well as protecting one's own self-image (Fisher, 1993).

Van Batenburt-Eddes, Bute and Van de Looij-Jansen (2012) examined the direct questioning method and concluded that it has insufficient validity. They compared actual police statistics of a group with the self-reported crime and found that 62% of the group admitted their criminal behavior. Consequently, self-reports do not provide a complete picture and other methods are necessary in order to find actual prevalence rates. Thus, it is proposed that using indirect estimations could be a solution to the problem of social desirability occurring when people are asked directly (Fisher, 1993). Empirical studies comparing indirect versus direct estimation methods of cybercrime consistently show that the reported prevalence rates are significantly higher when asked indirectly (Hengemühle, 2016; Wieschebrock, 2016). This is interpreted as support for the idea that social desirable responding is weakened through indirect questioning. That is, it is assumed that with indirect estimations, people no longer have the need of positive self-presentation. It seems that people tend to have fewer concerns about reporting their group's crimes rather than their own crimes.

However, there are still some questions concerning the mechanisms that play a role in indirect questioning. Social projection is the process through which people use their own behavior as a reference point and expect that other people behave in the same way as themselves if they make estimations about other people (Robbins & Krueger, 2005). It is therefore likely to be an important mechanism in indirect estimations. Until now, the relationship between social projection and indirect estimations of cybercrime has not been investigated. Thus, the aim of this study is to specifically examine the role of social projection as a mechanism in making indirect estimations about the prevalence of cybercrime among young adults.

Cybercrime and Digitized Crime

The use of Information- and Communication Technology (ICT) is steadily growing and for most people, its use is important in their everyday lives. Besides many advantages of ICT, it also offers an opportunity for criminals to commit different forms of crime online (Van der Laan & Goudriaan, 2016). At the moment, the Dutch police makes a distinction between cybercrime and digitized crime: On the one hand, cybercrime describes all offenses in which ICT itself is the modus operandi and at the same time the target, so cybercrime "cannot take place offline and did not exist before ICT emerged" (Van der Laan & Goudriaan, 2016, p.20). On the other hand, digitized crime includes mostly traditional crime whereby ICT is merely a new means by which the offense is carried out (Politie, n.d.). Some examples of cybercrime include ransomware, malware like the spreading of viruses or worms, hacking and DDoS attacks. Offenses that can be described as digitized crime are cyber mobbing, identity theft, phishing, bargain and sale fraud and child pornography (Politie, n.d.; Van der Laan & Goudriaan, 2016). For the sake of convenience, in the following, cybercrime and digitized crime are both referred to as cybercrime only.

Cybercrime is not uncommon among juveniles and young adults. Approximately 28% of young adults (18-22 years old) reported that they committed any of the cyber- or digitized delinquency that was asked in the previous year (Van der Laan & Goudriaan, 2016). Furthermore, in a study with Dutch and German students, almost 60% of them admitted to having illegally downloaded media and almost 15% admitted to having logged into the computer of somebody else without this person

knowing, during the past year (Wieschebrock, 2016). Currently, 17.9% of the Dutch population are registered as victims of any form of cybercrime or digitized crime (Centraal Bureau voor de Statistiek [CBS], 2017). The greatest proportion of these people are victims of some form of hacking (7.4%), whereas cyber mobbing accounts for the second highest number of victims (6%), bargain and sale offenses account for 4.1% of victims and identity theft accounts for 0.4% of victims (CBS, 2017).

Since the number of victims suffering from cybercrime and the prevalence rates of offenders found with self-report questionnaires is that high, it seems that these forms of crimes perhaps are not underestimated. However, these numbers may not reflect actual prevalence rates of offenders who commit cybercrime. Particularly, it is impossible to infer a specific number of perpetrators merely on the basis of the number of victims or offenders' self-reports (Van Batenburt-Eddes et al., 2012). Therefore, a dark number of the prevalence rate of perpetrators among juveniles and young adults exists that still has to be identified (Tcherni, Davies & Lizotte, 2015). The mechanisms of social projection and identification in indirect estimations will be examined in this context, while attempting to define the additional value of this method in unravelling the dark number to some extent.

The Use of Social Projection and Identification

Research shows that people base their indirect estimations on knowledge about themselves, which means they think that others behave like themselves (Robbins & Krueger, 2005). This process is called social projection and makes indirect estimations a valuable form of measuring the prevalence rates of specific attributes and behaviors in a group. Applying this phenomenon to the context of cybercrime prevalence rates, people are asked how many percent of people they think committed any cybercrime in the past year. However, they cannot know everyone else's behavior. Thus, they must have something to base their answer on. Studies establishing the variables that are associated with indirect estimations showed that by doing these estimations, information about people's own behavior is more accessible than information about other individuals of the group. This is a result of people not having relevant information about everyone else. Therefore, strong indications exist that the indirect estimations must be based on their own experience and behavior (Krueger & Clement, 1994; Nisbett

& Kunda, 1985; Ross, Greene & House, 1977), which is also called egocentric bias in the perceptual account of social projection (Krueger & Stanke, 2001).

Social projection works best if the estimation is about a group to which the person belongs but the person does not have the relevant information about all other group members (Krueger, 2012). Indirect estimations are therefore suitable for the application with ingroups in the context of cybercrime because ingroup means that the person is a member of the relevant group and is unlikely to know whether every person of the group did commit any form of cybercrime in the past. Thus, it seems that in indirect estimations, social desirability occurs less, but due to the process of social projection, people still provide an indication of the prevalence rate of their own criminal behavior. In addition, by examining the mean indirect estimations of all group members, they provide an estimation of the prevalence rate of each cybercrime offense in their group. In order to clearly examine the role of social projection in indirect estimations experimentally, it would be useful to manipulate the use of social projection. However, it is difficult to directly manipulate the use of social projection because it is described as a rather unconscious, automatic and perceptual process (Ames, 2004) and therefore, it is necessary to put social projection into a more comprehensive model.

Hengemühle (2016) showed that identification with the group is positively associated with the indirect estimations, which means that people who identify themselves with the relevant group more also estimated the prevalence rate of cybercrime in that group higher than people who identify themselves less with the group. On the one hand, identification with a group describes the affective and emotional value that is attached to the membership of a specific group (Leach et al., 2008). On the other hand, identification with a group also occurs when people perceive a mental overlap between themselves and the group. The higher the mental overlap, the higher the identification with the group (Coats, Smith, Claypool & Banner, 2000).

There are two mechanisms through which people create a correspondence between themselves and the group in order to identify with this group: self-stereotyping and social projection, which is also called self-anchoring (Cho & Knowles, 2013). Van Veelen, Otten and Hansen (2013) found that self-stereotyping, which means that people define themselves in terms of the group characteristics, occurs in clearly defined groups. Social projection or self-anchoring, on the other hand, means that people

define the group in terms of their own characteristics. This latter process occurs in unclearly defined groups (Van Veelen et al., 2013). In the context of cybercrime, it is unlikely that a person knows whether each group member committed the offenses because it contains a sensitive topic and therefore, the group is not clearly defined for him or her. Accordingly, people can create a mental bond with the group through a bottom-up mechanism like social projection (Van Veelen, Otten, Cadinu & Hansen, 2015). Otten (2004) showed that social projection does not only occur in minimal groups, but also in real groups. Social projection is positively associated with social identification, especially with affective components of identification (Van Veelen, Otten & Hansen, 2011). Therefore, social projection could be a possible mechanism through which identification is associated with indirect estimations.

One study found that among some cybercrimes and digitized crimes, there are differences between the prevalence rates of ingroup and outgroup estimations, with the ingroup prevalence estimate being significantly higher than the outgroup prevalence estimate for the same crime (Wieschebrock, 2016). It is possible that this phenomenon occurs due to the fact that social projection is used much less in outgroup estimations than in ingroup estimations because people would also identify themselves more with ingroup members than with outgroup members (Robbins & Krueger, 2005). However, it might also be that some cybercrime offenses are perceived less negatively and even as 'cool' to some extent (Zebel, de Vries, Kuttschreuter & Stol, 2013), such that people strategically overestimate the prevalence of such offenses in the ingroup compared to the outgroup. Thus, in the current study, merely ingroup estimations are examined to test the proposed role of social projection and the alternative explanation of attitude in the indirect estimations.

Hypotheses

The process through which people create a mental bond between themselves and the group provides the opportunity to stimulate the use of social projection indirectly via social identification. This goes along with the finding that simple variations in the personal membership of the group being estimated are sufficient for the person to no longer use social projection (Clement & Krueger, 2002). First, they found that people project their own characteristics to ingroup members but not to outgroup members

(a group to which they do not belong). Second, they changed the membership of the people so that participants were first assigned to a specific group, but at a second time were corrected as not being a member, or the other way around. They found that people use more social projection by estimating the group when they are described as a member than when they were not described as a member, whether they were first a member and then non-member or the other way around. This difference in the use of social projection was additionally larger for people who perceived a threat towards their self-image through the membership of the group that was assigned (Clement & Krueger, 2002).

In the context of cybercrime, if the need for identification with the group could either be decreased or increased respectively, there would be a low identity and a high identity condition. People in the high identity condition are expected to use more social projection. Therefore, the first hypothesis states that there is a positive effect of identification on the indirect estimations. People can create a mental overlap between themselves and the group by defining the group in terms of their own behavior. Contrary, people in the low identity condition, wishing to be distinct from the group, are expected to use less social projection while making the indirect estimations. Thus, if a person does not have the desire to identify himself in terms of the group, the person would base his estimations to a lesser extent on his own characteristics. The second hypothesis states that there is a positive effect of identification on social projection.

Thus, it is expected that, due to indirect estimations being based less on social desirability than direct estimations, in the condition where a high need of identification with the group is aroused, people use more social projection. In line with the previous finding that indirect estimations are higher than direct estimations and indirect estimations about an ingroup are also higher than indirect estimations about an outgroup, it is also assumed that people commit more cybercrime and digitized crime than they admit in direct estimations (Wieschebrock, 2016). Therefore, a greater use of social projection of people in the high identity condition is expected to lead to greater prevalence rate estimations of cybercrime in indirect estimations. Accordingly, the third hypothesis states that there is a positive association between social projection and the indirect estimations.

Putting these associations together, a conceptual model of identification, social projection and indirect estimations can be proposed (see figure 1). Accordingly, the fourth hypothesis states that the

positive effect of identification on indirect estimations is mediated by social projection. To sum it up, the four hypotheses are:

H1: There is a positive effect of the identification manipulation on the indirect estimations.

H2: There is a positive effect of the identification manipulation on the use of social projection.

H3: There is a positive association between the use of social projection and the indirect estimations.

H4: The positive effect of identification on indirect estimations is mediated by social projection.

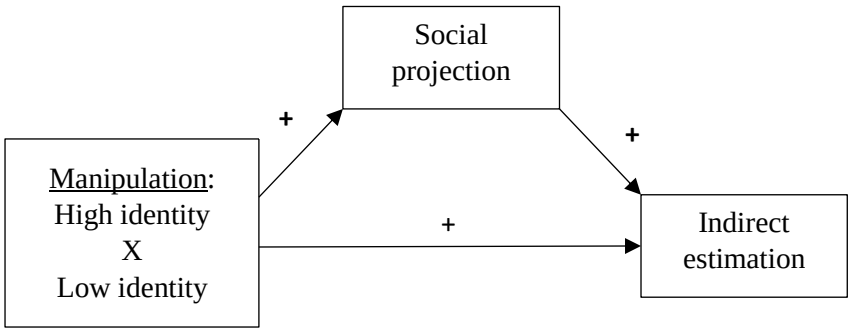

Figure 1. Proposed conceptual model describing the mediating role of social projection in the effect of identification on indirect estimations of cyber- and digitized crime.

Method

Participants

In total, 144 participants volunteered for the study, each of them was a member of either a student sports association (in the following called sports group) or first-year students of computer science (in the following called CS group) at the University of Twente. Attempting to conduct the study among nearly all members of these two existing groups, the goal was to calculate actual (self-reported) prevalence rates of the different forms of cybercrimes and digitized crime among and between these

groups. This would be an advantage because the prevalence rates would not have to be generalized toward a larger population.

Of all respondents, 50 participants did not complete the whole questionnaire (39 people stopped during the first question, the other ones stopped later in the questionnaire) and were therefore excluded from further analysis. Additionally, three people did not fill in all demographic variables and two people did not fill in the attitude scale. These people were not excluded from further analysis because the answers are not essential for further analysis.

After receiving permission from the ethical commission of the University of Twente, the researcher asked the board of the student sports association and the responsible professor of the first-year computer science students approval to conduct the study among their members. Then, all group members were approached after a meeting/training of the whole group and were asked to fill in the questionnaire. Also, an announcement on the Facebook-page of the sports group was placed in which the members were requested to take part in the survey. Some people were additionally personally approached and asked to fill in the questionnaire. As a result, there were 50 first year computer science students (response rate approximately 45.5%) and 44 members of the student sports association (response rate approximately 49.4%) that completed the survey. All participants filled in an informed consent before they took part in the study voluntarily and at the end of the survey, they were thanked and debriefed.

Of the 94 participants, 63 were male, 30 were female and one person did not report his/her gender. The participants were between the age of 17 and 28, with a mean age of 20.46 (SD = 2.26) and most of them were Dutch (more than 69% in both groups). In Table 1, a summary of the demographic data per group is given. The percentage of males is in the CS group higher (84%) than in the sports group (47.7%) and the highest reached education of most people in the CS group is a secondary school (94%) while this is more distributed among different education levels in the sports group. Also, the mean age is somewhat higher (approximately two years) in the sports group than in the CS group.

Table 1.
Frequencies, Percentages, and Means per Group of the Demographic Variables

		Sports group (%)	CS group (%)
Identification manipulation	High identity	26 (59.1)	24 (48)
	Low identity	18 (40.9)	26 (52)
Nationality	Dutch	37 (84.1)	35 (70)
	German	6 (13.6)	3 (6)
	Other	1 (2.3)	11 (22)
	Missing	-	1 (2)
Education	Secondary school	21 (47.7)	47 (94)
	Vocational education	1 (2.3)	1 (2)
	Bachelor degree	17 (38.6)	-
	Master degree	5 (11.4)	-
	Missing	-	2 (4)
Gender	Male	21 (47.7)	42 (84)
	Female	23 (52.3)	7 (14)
	Missing	-	1 (2)
Mean age (*SD*)		21.9 (2.26)	19.19 (1.32)
Missing		2	2

Note. sports group: $n = 44$; CS group: $n = 50$.

Overview and Design

A between-groups design was employed. There were two levels of the independent variable 'identification manipulation': high vs. low. The participants were randomly assigned to either of these conditions. Of the 94 participants, 44 were in the low identity condition and 50 were in the high identity condition. After the manipulation, measures of indirect and direct estimations of cybercrime were administered, as well as social projection.

Manipulation of Social Projection. The manipulation of social projection via the need for identification with the group was a between-subjects variable. Sahdra and Ross (2007) conducted a study about the relation between group identification and historical memory in which they triggered either the affiliation need, which is the wish of being a member of the group and the possibility to

identify oneself with the group, or the differentiation need, which is the wish to be distinct and independent from the group. According to optimal distinctiveness theory (Brewer, 1991), people strive for a balance between these opposite needs. Thus, by triggering one of these needs, the need for identification with the group can be either increased or decreased. The manipulation of Sahdra and Ross (2007) was used to trigger the affiliation need in the high need for identity condition. This was done by asking the participants to recall situations where they felt overly dissimilar to the group so that they experience a need to be more like the group. In the low need for identity condition, the differentiation need was triggered, which should decrease the importance of the social identity. The differentiation need was aroused through asking the participants to recall situations where they felt overly similar to the group. This was done so that they experience the need to be distinct from the group (Sahdra & Ross, 2007).

In their manipulation control, Sahdra and Ross (2007) found that identification was successfully lowered in the differentiation need arousal condition, but there were no significant differences between the assimilation need arousal condition and the control condition that they included. They attributed this to a possible ceiling effect in the identity of the control group with their nationality which was the group of interest in this study (Sahdra & Ross, 2007). Therefore, no control condition was added in the current study because it is likely that this would resemble another high need for identity condition. The instructions for the different groups were thus adopted from Pickett, Silver and Brewer (2002) after the modifications of Sahdra and Ross (2007). Additionally, the instructions were slightly changed so that the task was about the relevant ingroup of the participant, and not about other people in general:

Low identity (Differentiation Need Arousal) Condition:

Please take a moment and think of times when you felt too similar to other sports group members/first-year computer science students, so similar that you felt uncomfortable. In other words, think of times and situations where you felt that you were so much like the other sports group members/first-year computer science students that you could not “stick out,” that your own self was hidden. Please write a brief description of two memories of such times.

High identity (Assimilation Need Arousal) Condition:

Please take a moment and think of times when you felt too different from sports group members/first-year computer science students, so different that you felt uncomfortable. In other words, think of times and situations where you did not feel that you fit in with the other sports group members/first-year computer science students and that you "stuck out." Please write a brief description of two memories of such times.

Dependent variables. In total, there were nine different offenses that participants were asked about. The offenses of cybercrime included ransomware, other malware, like the spreading of viruses or worms, hacking and DDoS attacks. Digitized crimes were cyber mobbing, illegally downloading media, identity theft, bargain and sale fraud and child pornography (see appendix for all questions). The participants were provided with short descriptions about what is generally understood under each offense. This was to make sure that there are no misconceptions and that the estimations can accurately be compared among participants. The definition of child pornography was for example: "Child pornography includes viewing, having, selling, spreading or produce pornographic materials that exploit or portray a minor (under the age of 18), for example having a selfie that involves underage nudity or sexual situations on your cellphone." Additionally, one item was added that covers the overall execution of cybercrime and digitized crime. Here, the definition included: "Now think of all possible forms of cybercrime like the offenses previously mentioned in addition to other forms that you can imagine." (see appendix for all questions and definitions).

Of each offense, all participants filled in both the direct and indirect measurements. With respect to the indirect measurement, participants were asked about how many of their peers they thought were offenders of each cybercrime or digitized crime in the previous 12 months. An example of such question is: "How many percent of other sports group members/first-year computer science students do you think have hacked someone else in the previous year?" The possible range of answers was thus between 0% and 100%.

The second type of estimation was the direct measurement, the participants answered for each cybercrime and digitized crime separately whether they committed the crime in the previous 12

months themselves. Possible responses were either “yes” or “no”. An example of the direct measurement is: “Did you illegally downloaded media from the internet in the previous 12 months?”

Control measures. A few scales were included in order to test whether the manipulation was successful and examined the use of social projection, the occurrence of social desirability and the attitude towards the offenses.

Manipulation check. The identification scale from Leach et al. (2008) was used to test whether the increase of the differentiation need arousal or assimilation need arousal decreased identification in the low identity condition and increased identification in the high identity condition respectively. The scale contains 14 items, which covers the following five aspects of identification: Solidarity, satisfaction, centrality, individual self-stereotyping and in-group homogeneity. The items consist of statements for which the participants each indicate their agreement on a 5-point Likert scale. An example item of the centrality aspect is: “I often think about the fact that I am a member of my group.” All items of the scale are presented in the appendix. The reliability of this scale is good, with a Cronbach’s alpha value of 0.88.

Social projection (mediator). To test whether the manipulation truly affected the use of social projection, participants indicated their extent of agreement on a 5-point Likert scale on four items. Two items examined the extent to which participants based the indirect estimations on knowledge and experience of themselves. Furthermore, people were asked to what extent they estimate their own behavior equal to the behavior of other group members. Another item examined the use of self-competition, which is a sort of opposite to social projection (“By estimating the prevalence of the different cyber delicts, I expected that the other group members behave unequal to me”). Two additional items measured the extent to which people are certain about their answers to the indirect questions.

Principal component analysis to this scale including all six items indicated that the scale measures three factors. Therefore, only the first two items ($r = .44$, $p < .001$) were used as an indicator for social projection because they loaded strongly on one factor and likely indicated best the use of social projection with regards to their content. An example of one of these two questions is: “By

estimating the prevalence of the different cyber delicts, I used knowledge of my own behavior in order to make assumptions about the behavior of the other group members." (see appendix for the whole scale). The other items of the scale were considered separately in the following analyses, for example, the item: "My own behavior is similar to the behavior of the other group members".

Additionally, two items were added where participants indicated to what extent they have knowledge about the criminal behavior of the other group members. It would likely support the idea that people use social projection in indirect estimations if they have no relevant information. One item is: "I know all of my group members well." The correlation between the two items is nearly good (r = .39, p < .001).

Social desirability. Four items were added examining the perceived social desirability in the estimations. Respondents indicated their agreement on a 5-point Likert scale. Two of the items were related to the direct estimations and the other two items to the indirect estimations. Therefore, the extent of social desirability could be measured on the individual level and on the group level. An example of these items is: "I wanted to represent my group in a positive way by answering the questions about other group members' behavior." The two items measuring social desirability for the group level correlate weakly with each other (r = .25, p = .015). Therefore, these items were considered separately in the following analyses. The items measuring the social desirability on the individual level revealed a rather high correlation (r = .50, p < .001) and could therefore be combined in a scale.

Attitude. To test whether the attitude towards the offenses also influenced the indirect estimations, two items were added per offense that estimate first the extent to which the respondents think positive or negative about the offenses. They indicated their answer on a 5-point Likert scale ranging from "Extremely negative" to "Extremely positive". Second, they indicated to what extent they estimate these offenses as normative on a 5-point Likert scale ranging from "Totally unacceptable" to "Totally acceptable". The correlations between the two items were for each offense above .57 (all values of p < .001). The mean score of the two items per offense was therefore calculated to indicate the attitude towards the offenses.

Participation and honesty control. Finally, there was also a scale of control questions examining the degree to which people participated seriously in the study adopted from Hengemühle (2016). The respondents could give their extent of agreement towards five different statements like "Did you participate seriously in this research?" on a 5-point Likert scale (see appendix).

Procedure

At the very beginning, people were asked to which of the two examined groups they belong to. Then, the participants had to complete the "personal memory task", which included the manipulation of identification. Next, the participants first filled in the indirect estimations for each form of cybercrime so that they immediately took place after the manipulation. Following that, participants filled in the identification scale, social projection items and the items measuring social desirability on the group level. After that, the respondents made the direct estimations for each form of cybercrime and digitized crime. Later, the other control measures were examined including items measuring social desirability on the individual level, the attitude towards the offenses measure and the honesty control measure. At the end of the survey, the demographic variables of the participants were asked. These included age, gender, nationality and highest level of education.

Results

Overview of direct and indirect estimations

A summary of the comparison between the indirect and direct estimations per group is given in table 2. The comparison revealed that only for the offense of cyber mobbing, findings of earlier research could be replicated, in which the indirect estimation in both groups is significant higher than the direct estimation. In the CS group, this is also true for the offense of ransomware. For example, in the CS group, 2% of the people admitted that they spread and/or created ransomware in the previous 12 months, while the mean indirect estimation was approximately 8%. Unexpectedly, the opposite pattern occurs in the offenses including hacking, committing a DDoS attack and identity theft in the estimations of the sports group. Here, the indirect estimations are significant lower than the direct

estimations in the sports group. For example, approximately 18% admitted that they hacked someone else in the previous 12 months, while people estimated the prevalence rate in their group approximately just 9%. Also in contrast to previous findings, for the remaining offenses, there are no significant differences at all between the indirect and direct estimations.

When comparing the indirect estimations between the two groups, mean indirect estimations are significant higher in the CS group for the offenses including hacking, committing a DDoS attack, cyber mobbing, cybercrime in general and the mean indirect estimation of all offenses (see table 2). Also, the maximum reported individual indirect estimation reached four times 100% in the CS group, while in the LR group, only for one offense 100% is reported. This is in line with the idea that CS group members in general deal more often with topics such as ICT and cybercrime.

The difference in the direct estimations between the groups is only significant for the offense of committing a DDoS attack, 22% of the CS group reported that they committed a DDoS attack, while in the sports group, only 4.5% admitted that they committed a DDoS attack in the previous 12 months.

In comparison with previous studies, indirect estimations in the current study were lower for the offenses including hacking, committing cyber mobbing and spreading viruses and/or worms. Only for the offense of illegally downloading media, the indirect estimation was higher in the current research than in previous research.

Table 2.

Summary and Comparison of the Direct and Indirect Estimations per Group and per Cybercrime Offense

	Sports group			CS group			Difference between groups in indirect estimations	Difference between groups in direct estimations
	Direct question	Indirect estimation	Difference indirect & direct estimation	Direct question	Indirect estimation	Difference indirect & direct estimation		
Offenses	Yes (%)	*M* (*SD*)	*t*	Yes (%)	*M* (*SD*)	*t*	*t*	χ^2
Ransomware	4.5	4.84 (10.07)	0.23	2	7.82 (16.27)	2.53*	-1.05	0.49
Viruses/worms	6.8	6.14 (14.55)	-0.30	6	7.92 (11.38)	1.19	-.67	0.03
Hacking	18.1	9.25 (14.55)	-4.03***	20	20.74 (22.30)	0.24	-2.99**	0.26
DDoS attack	4.5	1.59 (2.37)	-8.16***	22	20.92 (22.83)	-0.34	-5.95***	5.98*
Cyber mobbing	2.3	7.27 (9.91)	3.33**	4	14.06 (19.74)	3.60***	-2.14*	0.23
Illegally downloading media	84.1	81.80 (14.98)	-1.02	84	87.96 (21.24)	1.32	-1.61	0.08
Identity theft	4.5	1.91 (2.86)	-6.01***	6	4.36 (10.37)	-1.12	-1.6	0.1
Bargain/sale fraud	4.5	5.05 (8.56)	0.42	6	7.54 (17.6)	0.62	-0.86	0.1
Child pornography	2.3	3.09 (11.3)	0.46	4	4.14 (10.71)	0.09	-0.46	0.23
Cybercrime general	38.6	29.59 (33.41)	-1.79	60	57.84 (40.19)	-0.38	-3.72***	4.27
Mean cybercrime	-	15.17 (5.91)	-	-	23.33 (10.97)	-	-4.55***	-

Note. *p < .05; **p<.01; *** p <.001; sports group: n = 44; CS group: n = 50

Manipulation check on identification

Many participants (41.49%) either did not answer the personal memory task at all (eleven participants) or reported that they have not had such a memory (28 participants). That is true especially for people in the low identity condition (50% of them reported that they haven't had such an experience). This is a first indication that the manipulation was less than successful.

A 2x2 ANOVA, measuring the effects of group membership (sports group vs. CS group) and identification manipulation (high vs. low) on the identification scale as dependent variable, revealed that there is no significant main effect of the identification manipulation on the whole identification scale ($F(1,90) = 0.52$, $p = .47$). Unexpectedly, the main effect for group membership was significant ($F(1,90) = 19.99$, $p < .001$) with the mean identification in the sports group ($M = 3.78$, $SD = 0.56$) higher than the mean identification in the CS group ($M = 3.23$, $SD = 0.62$). There is also no significant interaction effect ($F(1,90) = 0.15$, $p = .7$). In addition, it was also tested whether the manipulation had an effect on the different subscales of the identification scale (solidarity, satisfaction, centrality, individual self-stereotyping and ingroup homogeneity) using the same procedure. In these analyses, there was also neither a significant main effect of the identification manipulation, nor an interaction effect found, indicating that the manipulation on identification was not successful at least on the used measurement.

Hypotheses testing

To test whether the manipulation had an effect on the indirect estimation as proposed (hypothesis 1), a 2x2 MANOVA analysis was executed with group membership (sports group vs. CS group) and identification manipulation (high vs. low) as independent variables and all indirect estimations as well as the mean of all indirect estimations together as dependent variables. Overall, the MANOVA analysis revealed no significant effect of identification manipulation ($F(10,80) = 1.48$, $p = .161$) or interaction ($F(10,80) = 1.58$, $p = .127$) on the indirect estimation. Only the effect of group membership was significant ($F(10,80) = 3.75$, $p < .001$).

Consequently, the manipulation of identification had no overall effect on the indirect estimations, as opposed to the expectation of hypothesis 1. However, when looking at the individual

ANOVA between-subjects analysis per dependent variable, there was one significant interaction effect, namely for the offense of spreading viruses and/or worms ($F(1,89) = 8.85$, $p = .004$). The CS group showed the expected pattern of a higher indirect estimation in the high identity condition than in the low identity condition (High identity $M = 11.25$, $SD = 14.48$; Low identity $M = 4.85$, $SD = 6.37$), whereas the opposite pattern can be seen in the sports group (High identity $M = 2.35$, $SD = 3.94$; Low identity $M = 11.61$, $SD = 21.43$). Regarding the overall test of the model, this effect should be interpreted with great caution, as it might very well be an effect by chance (type 1 error). Thus, the manipulation of identification did not influence the indirect estimations about the cybercrime offenses. Consequently, the first hypothesis cannot be supported.

For the offenses including hacking, committing a DDoS attack, cyber mobbing, cybercrime in general and the mean indirect estimation there was a significant main effect for the group (for all the offenses mentioned: $F(1,89) > 7.61$, $p < .01$), with the indirect estimations in the CS group being higher than in the sports group (see also overview of direct and indirect estimations above).

Furthermore, it was tested whether the manipulation influenced the use of social projection (hypothesis 2). A 2x2 ANOVA, measuring the effects of group membership (sports group vs. CS group) and identification manipulation (high vs. low) with the mean score of the first two items of the social projection scale as dependent variable revealed no significant main effect of identification manipulation ($F(1,90) = 0.26$, $p = .61$), indicating that there is no clear association between identification and social projection as proposed in the hypothesis. There was a significant main effect for group ($F(1,90) = 10.06$, $p = .002$) with the mean social projection in the sports group ($M = 3.81$, $SD = 0.82$) higher than the mean social projection in the CS group ($M = 3.28$, $SD = 0.87$). The interaction effect was also significant ($F(1,90) = 5.8$, $p = .018$). The CS group showed the expected pattern: more use of social projection in the high identity condition than in the low identity condition (High identity $M = 3.54$, $SD = 0.69$; Low identity $M = 3.04$, SD = 0.96), whereas the opposite pattern occurred in the sports group (High identity $M = 3.67$, $SD = 0.96$; Low identity $M = 4$, $SD = 0.51$; see figure 2). Hypothesis 2 could therefore only be supported for the CS group and not for the sports group. This, in addition to the finding that the identification manipulation did not affected the

identification scale, also indicates that the manipulation influenced another process than identification that accounts for the change in the use of social projection.

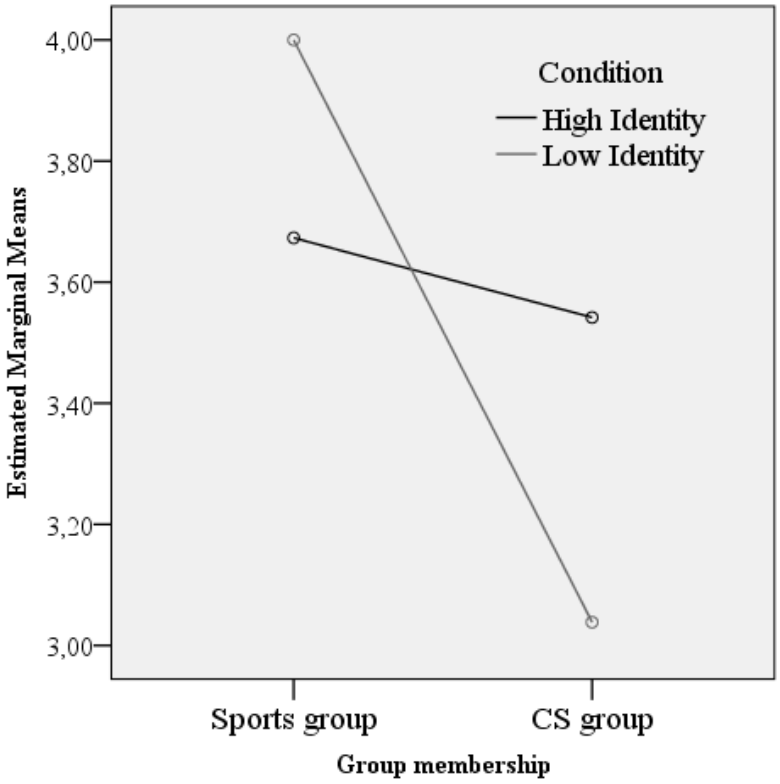

Figure 2. Interaction effect of group membership and identification manipulation on social projection.

Mediation analysis

Despite no significant differences of the identification manipulation (high vs. low) on the identification scale, mediation analyses were executed to examine whether non-experimental effects can be found. This analysis was executed for both groups apart, because previous findings revealed significant differences between the groups, for example on the social projection scale. A first mediation analysis revealed only two significant main effects of identification on the indirect estimation. First, in the CS group, there is a significant positive effect of identification on the indirect estimation about hacking (b = 10.20, $t(48)$ = 2.05, p = .046). In the sports group, this effect was significant positive for the indirect estimation about bargain/sale fraud (b = 5.67, $t(42)$ = 2.58, p = .013). Thus, the higher the identification with the group, the higher is the indirect estimation about these offenses in the groups. Because this positive association is only found for one offense in each

group, hypothesis 1 cannot be held in a generalized context and the other hypothesis are tested further only for the offense of hacking in the CS group and committing bargain/sale fraud in the sports group.

The second mediation analysis revealed the effects of identification on the first two items that measured social projection. In the sports group, there was a positive significant effect ($b = 0.45$, $t(42) = 2.10$, $p = .042$) of identification on social projection. In the CS group, this effect was not significant ($b = 0.05$, $t(48) = 0.26$, $p = .796$). As a consequence, the higher the identification, the more social projection was used in the sports group. This is not true for the CS group. Therefore in contrast to the findings from the ANOVA analyses, hypothesis 2 can be held for the sports group, but not for the CS group. Thus, only for the offense bargain/sale fraud in the sports group, the hypotheses could be tested further.

The last mediation analysis with identification and social projection both as independent variables and the indirect estimation of bargain/sale fraud as a dependent variable was executed for the sports group. It revealed that the effect of identification on the indirect estimation remained significant ($b = 5.85$, $t(41) = 2.50$, $p = .016$) and no significant effect of social projection on the indirect estimation was found ($b = -0.40$, $t(41) = -0.25$, $p = .806$). Consequently, hypothesis 3 and 4 cannot be supported. There is no association between social projection and indirect estimations and there is also no mediation between identification and indirect estimation via social projection. Also, regarding the large amount of mediation analyses that were executed, the little significant effects should be interpreted with great caution, as they might as well be an effect by chance (type 1 error).

Explorative analysis

To test whether there are other possible effects of the group membership and identification manipulation on the indirect estimations, the variables of attitude, group knowledge, and social desirability were further examined in another 2x2 MANOVA analysis. Some significant effects were found, but they showed no clear pattern that conceivably explained other possible effects occurring when making indirect estimations. Therefore, only some of the findings regarding the attitude towards the offenses are discussed briefly because it was proposed that this could also play a role in indirect estimations.

The analysis revealed that for the attitude towards spreading ransomware and viruses/worms, there were significant interaction effects of group membership and identification manipulation (for both offenses: $F(1,87) > 3.80$, $p < .010$). This means, that in the CS group, the attitude towards the offenses is more negative in the low identity condition than in the high identity condition, whereas the opposite is true for the sports group. Here, the attitude towards the offenses is more positive in the low identity condition than in the high identity condition. In figure 3, this pattern can be seen. For the attitude towards committing a DDoS attack, there was a significant main effect of the group, indicating that the attitude towards that offense is more positive in the CS group than in the sports group. In general, the attitude towards the offenses is quite negative in both groups, except the offense illegally downloading media. The attitude towards this offense is in both groups about two points higher than for the other offenses. These findings indicate that attitude does play a role in indirect estimations, but the nature of this association cannot clearly be stated out on the basis of the current study.

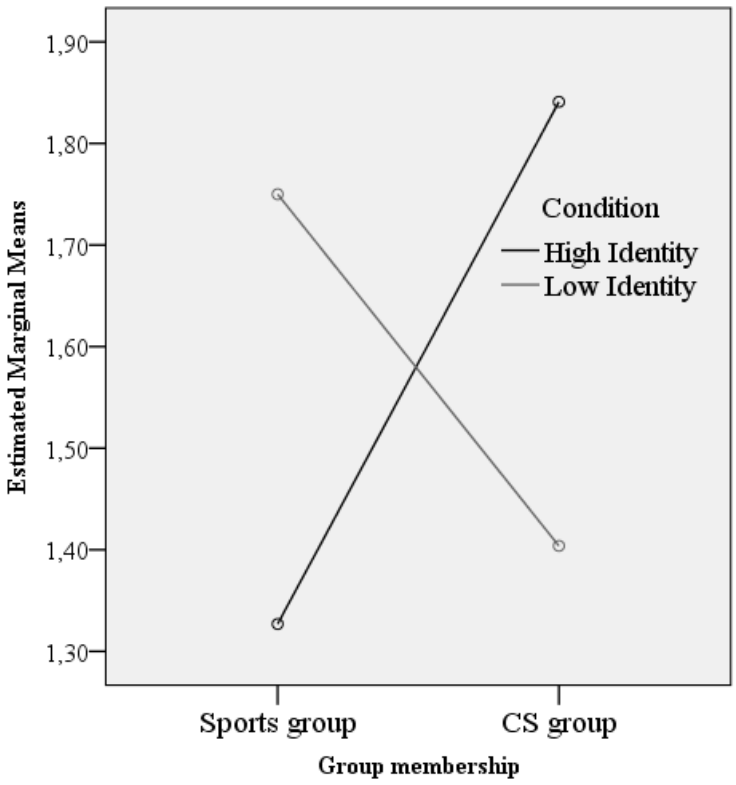

Figure 3. Interaction effect of group membership and identification manipulation on the attitude towards spreading viruses and/or worms.

The answers on the honesty scale are summarized in Table 3. In general, people answered the questions mostly honest and participated seriously. Despites that, in the CS group, two people did not understand the questions at all and one person indicated that he or she did not answer the questions about their own criminal behavior honestly. One person also indicated the same in the sports group. Furthermore, six people did not believe that their answers stayed anonymous, two of them were in the sports group and four of them in the CS group. The mean score indicating the extent of confidence that the data is treated anonymously and trustworthy differs significantly between the groups. In the sports group, people have less confidence than people of the CS group ($t(90) = -3.14$, $p = .002$, see table 3). In the CS group, two people also did not fill in the honesty scale.

Table 3.

Means and Standard Deviations for the Honesty Items per Group.

	Sports group *M* (SD)	CS group *M* (SD)
Did you understand questions?	1.64 (0.57)	1.56 (0.82)
Did you participate seriously in this research?	1.32 (0.52)	1.23 (0.43)
Did you honestly answer the questions about your own criminal behavior?	1.18 (0.54)	1.25 (0.67)
Did you honestly answer the questions about other people's criminal behavior?	1.57 (0.63)	1.60 (0.57)
Do you have confidence that your answers stay anonymous and are treated trustworthy?	1.61* (0.84)	2.21* (0.97)

Note. Answers to the questions range between 1 = Definitely yes and 5 = Definitely not; sports group: $n = 44$; CS group: $n = 48$; *Significant difference between mean score of the groups, $p < .01$.

Discussion

The goal of the current study was to examine the role of social projection and identification in making indirect estimations in order to approach actual prevalence rates of cybercrime and digitized crime. It was expected that more identification with the group leads to higher indirect estimations about the prevalence rates of cybercrime in that group. Furthermore, it was assumed that this association could be explained by a greater use of social projection, meaning that people who identify more with the

group, also base their estimations more on their own behavior and therefore estimate the prevalence rates higher than people who identify less with the group.

Putting all findings into a nutshell, a few conclusions can be drawn. First, there were many differences between the examined groups, which made it difficult to draw general conclusions at all. However, against expectation, it was found that the identification with the group had no general positive effect on the indirect estimation meaning that people who identify more with their group did not estimate the prevalence rates higher. Therefore, the mediating role of social projection in the relation between identification and indirect estimations could not be tested in this context either.

Relating to the influence of the manipulation of identification with the group on the use of social projection, conflictive results were found. On the one hand, when regarding the identification manipulation conditions, it can be pointed out that as expected, the first-year computer science students who identified more with the group based the indirect estimations more on their own behavior than the first-year computer science students who identified less with the group. However, against expectation, the opposite pattern was found in the sports group. On the other hand, when considering the identification scale that was additionally examined, a positive link between identification and the use of social projection was found only for the sports group. Therefore, there could be an association between the extent of identification and social projection but the current data is not sufficient clear-cut to define the nature of this association.

With regard to the association between the use of social projection and the indirect estimations, it was found that the high identifiers of the first-year computer science students used more social projection, and they also estimated the prevalence rate of spreading viruses and/or worms higher than the low identifiers. In contrast, the high identifiers in the sports group used less social projection but estimated the prevalence rate of spreading viruses and/or worms in their group lower than the low identifiers. Therefore, the findings suggest that there is a positive association between the use of social projection and the indirect estimations; greater use of social projection would lead to higher estimations in both groups. Taking all these findings together, it can be concluded that social projection likely plays a role in indirect estimations. However, in the current study, there is not enough

evidence to support the proposed model of the mediating role of social projection between identification and indirect estimations thoroughly.

Despite that, another important finding is that in contrast to earlier research, for the first time it was found that the indirect estimation was lower than the direct estimation. In the sports group for the offenses hacking and committing a DDoS attack and identity theft, people estimated the prevalence rate of their group lower than the prevalence that comes from the self-reported cybercrime rates. As found in previous research, the indirect estimations were higher than the direct estimations only for the offense of spreading ransomware among the first-year computer science students and the offense of cyber mobbing among both groups.

There are a few possible explanations for the current findings. A first possible explanation for the unexpected finding that no association between identification and indirect estimations could be found appears if the attitude towards spreading ransomware and viruses and/or worms would be taken into account. Assuming that people want to represent their group in a positive way (Hogg, 2013), a more positive attitude towards these offenses would lead to higher estimations especially in the high identity condition because people who identify the most with their group are the most motivated to show a positive image of their group (Hogg, 2013). This is in line with the finding of Hengemühle (2016) that a more positive attitude towards an offense is associated with a higher indirect estimation. In the current study, it was found that a less negative attitude towards the offenses of spreading ransomware and viruses and/or worms goes hand in hand with a higher prevalence rate of these offenses and a more negative attitude goes along with a lower prevalence rate especially for the people who identify more with their group. It was also found that the attitude towards committing a DDoS attack is more positive among the first-year computer science students than among the sports group, and members of the first-year computer science students group reported both more often that they themselves committed that crime, and their estimated prevalence rate of the group was also higher than in the sports group. This supports the idea that a more positive attitude towards a specific crime leads to higher prevalence rate estimations of that crime but this association has to be further examined, for example using experimental manipulation of the attitude.

A second possible explanation for the current findings deals with the group clarity. The fact that there was no clear overall association between identification and social projection is in contrast to the study of Van Veelen et al. (2015). They found a positive link between social projection and social identification, especially with the affective components of identification. Van Veelen et al. (2013) found that self-anchoring, meaning the same as social projection, occurs in unclearly defined groups, but not in clearly defined groups. It was assumed that in the present context, the identity content of the groups with regard to cybercrime is not clearly defined because it contains a sensitive topic. Nevertheless, the study was conducted among two existing groups, which were small at size. Therefore, it is possible that the identity contents of the current groups were perhaps clearly defined. If this is true, it is reasonable that social projection did not play a significant role. That implicates that the proposed model could not be applied in the context of clearly defined groups, but arguably still with unclearly defined groups like the large population of young Dutch adults for example.

In line with this reasoning, the characteristics of the groups that were used could also explain other conflictive findings between the current research and previous research. For example, it was found that the indirect estimations in the current research were lower for the offenses of hacking, committing cyber mobbing and spreading viruses and/or worms than in previous research. Only for the offense of illegally downloading media, the indirect estimation was higher in the current research than in previous research. In the previous research, the group of interest covered the large population of young adults of one whole country, of which the identity content likely is unclear. In contrast, smaller, real groups, which are possibly more clearly defined, were used in the current research.

Therefore, the participants could have perceived a greater threat to the positive image that they have about their group while making the direct and indirect estimations because it is more difficult for them to differentiate themselves from the group. In consequence, people see themselves as a group member, in terms of their social identity while making the indirect estimations (Hogg, 2013). Accordingly, with regard to the direct estimation of the self-reports people do not perceive a decrease in social desirability when making the indirect estimations about the group. Rather, the same social desirability can occur on a group level, meaning that the participants wanted to protect a positive image of themselves and their group while keeping the indirect estimations low. Because people were

approached personally and knew who got their answers, they could have perceived even a greater threat to the positive image of their group, especially in the sports group because the researcher was a member of the sports group herself. This could be supported with the finding that sports group members reported that they have less confidence that their answers stay anonymous and are treated trustworthy than the first-year computer science students. That social desirability would play such a meaningful role in indirect estimations would be an interesting finding because it would contradict the finding of Fisher (1993) that there occurs less social desirability while asking indirect questions. And therefore, it would partly undermine the main reason why the indirect estimations are used.

Extending this argumentation, social desirability on a group level could also explain why some indirect estimations are lower than the direct estimations. In the sports group for some of the offenses, people estimated the prevalence rate of their group lower than the prevalence that comes from the self-reported cybercrime rates. In earlier research, the indirect estimation was constantly higher than the self-reported cybercrime rates in the direct estimations (John et al., 2012; Hengemühle, 2016; Wieschebrock, 2016). In the current research, this is only true for the offense of ransomware in the group of first-year computer science students and the offense of cyber mobbing in both groups.

Perhaps the sports group members, who identified more with their group than the members of the CS group, were more inclined to keep a positive image of their group. This idea can be supported taking the findings of Sahdra and Ross (2007) into account. They found that people who identify more with their group reported less violent acts of their group than did low identifiers, possibly because of the effects of their social identity on their schema-consistent memory recall. That means that people who identify more with their group have a positive image of their group and therefore it is easier for them to recall good memories and acts of their group than negative acts, as for example criminal behavior. Voci (2006) also found that if the positive image that people have about their ingroup is threatened, a bias occurs that lead people to answer in a social desirable way so that the positive image of the group can be protected.

Strengths and weaknesses of the current study

A great advantage of the current study is that for the first time small existing groups were used in contrast to a group that includes a broad social category like nationality. By doing this, it was possible to unravel that there must be essential differences in the processes that play a role while people make indirect estimations about the prevalence rates of cybercrime in a large population compared to a small group to which they both belong to. This finding is very crucial because it has important implications for the use of indirect questions. Next, while using two groups, it was possible to uncover some differences between the members of a student sports association and first-year computer science students with regard to their cybercrime behavior. Also, the study focused on some more cybercrime offenses than previous studies, and included definitions of the offenses for the participants, thereby obtaining a clearer picture of the cybercrime behavior on the examined groups.

Nevertheless, there are also some limitations of the current study. First, the manipulation of identification seemed to be unsuccessful. The reason for this could be either the manipulation itself because people refused to answer the manipulation task or the identification scale that was used to indicate whether the manipulation was successful. In contrast to the scale designed by Leach et al. (2008), it is possible that the Cameron measure of identity (Cameron, 2004) which was also used by Sahdra and Ross (2007) as a manipulation control, would capture better the identification manipulation. This measurement includes reversed items to measure identification with a group and captures other subscales as the measurement of Leach et al. (2008). Accordingly, if another identification scale would capture differences between the conditions, the findings would be interpreted differently.

Second, the results regarding the use of social projection should be interpreted carefully because the social projection measurement can be criticized. A few questions were supposed to measure social projection. The participants were asked directly whether they used knowledge about their own behavior and experiences while making the assumptions about other people's behavior. It is important to note that social projection is a rather unconscious process (Ames, 2004) and overt questions like these may not indicate the actual use of social projection accurately. Other studies in the past mostly used correlations between self-judgments and judgments about others to indicate the

extent of social projection (see for example Ames, 2004; Krüger & Stanke, 2001; Van Veelen et al., 2011). Therefore, it could be true that the self-reported use of social projection is not in accordance with the actual extent of social projection used. The correlational type of social projection measurement cannot easily be applied in this context because it is assumed that there are differences between the self-reports and indirect estimations due to social desirability. However, it is possible to include prevalence rates estimates of character traits with a neutral valence. These traits could be measured through both individual self-report and indirect estimations for a group, and correlational analyses would reveal a correlation score of the two measures that indicates the use of social projection. This correlation score could also be used for the prevalence rates about cybercrime offenses.

Finally, a lot of first-year computer science students either refused to participate in the study or did not finish the survey. Evaluation with some members revealed that the group members of first-year computer science students did not want to participate because they detected some mistakes in the definitions of cybercrime and did not regard the research as seriously anymore or they found it too "psychological" meaning that they were confused about the questions regarding the manipulation or identification measurement. Two people of the first-year computer science students who completed the questionnaire even reported that they did not understand the questions at all. These facts could also distort the data to a large extent.

Suggestions for further research and implications

Taking the strengths and limitations of the current study into account, a few suggestions for further research can be made. It is strongly recommended to conduct a study, examining the different processes and mechanisms that play a role while making indirect estimations about different groups. Applying a within-subjects design, participants should make indirect estimations about a large group to which they belong (for example a broad social category like nationality) as well as a smaller group to which they also belong. It is assumed that the large group would be an unclearly defined group in contrast to the small group which would be a clearly defined group. The order of presenting these groups to participants should be counterbalanced.

Using a design like this, the role of identification and social projection in indirect estimations should be further examined. It would also be interesting to use another manipulation that would successfully increase identification in one condition and decrease identification in another condition. Additionally, another manipulation control measurement should be used. The identification scale should include reversed items, like the scale from Cameron (2004). Next, it would be important to measure the unconscious use of social projection in a more covert fashion providing the opportunity to compare the two indirect measurements and to infer which processes and group characteristics such as group clarity or group size account for the differences that were found in the current study.

Regardless of the limitations of the current study, there are some important implications. First of all, indirect estimations in general should be interpreted carefully because it seems as if they depend on the groups about which the estimations are made. Furthermore, it seems as if the indirect estimations are not constantly influenced by the identification with the group as stated in previous research (Hengemühle, 2016). Additionally, it appears that the use of social projection as well as the attitude towards the offenses play a role in indirect estimations, but it is still unclear how and under which conditions. Finally, the findings suggest that social desirability could also play a larger role in indirect estimations if smaller, precisely defined groups are used. It would also be of additional value if indirect estimations would be compared with actual police statistics of cybercrime offenses that are reported in the relevant group. Like in the work of Van Batenburg-Eddes et al. (2012), the comparison would help determining the validity of indirect questions in the context of cybercrime. Combined with previous suggested research, it would be possible to find out whether, and under which conditions, the indirect estimations clearly add something in finding actual prevalence rates of cybercrime and digitized crime.

References

Ames, D. R. (2004). Strategies for social inference: A similarity contingency model of projection and stereotyping in attribute prevalence estimates. *Journal of Personality and Social Psychology, 87*(5), 573-585. doi:10.1037/0022-3514.87.5.573

Van Batenburg-Eddes, T., Butte, D., van de Looij-Jansen, P., Schiethart, W., Raat, H., de Waart, F., & Jansen, W. (2012). Measuring juvenile delinquency: How do self-reports compare with official police statistics? *European Journal of Criminology, 9*(1), 23-37. doi:10.1177/1477370811421644

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17*(5), 475-482.

Cameron, J. E. (2004). A three-factor model of social identity. *Self and Identity, 3*(3), 239-262. doi:10.1080/13576500444000047

Centraal Bureau voor de Statistiek StatLine – Ondervonden delicten; persoonskenmerken. (2017, March 1). Retrieved from http://statline.cbs.nl/statweb/publication/?vw=t&dm=slnl&pa=83095ned&d1=0,5,29-32,37,40,45-46,51-52,69,81,99,104,114,124,135&d2=0&d3=0&d4=a&hd=151216-0953&hdr=g1,g2,g3&stb=t

Cho, J., & Knowles, E. (2013). I'm like you and you're like me: Social projection and self-stereotyping both help explain self-other correspondence. *Journal of Personality and Social Psychology, 3*, 444-456. doi:10.1037/a0031017

Clement, R.W., & Krueger, J. (2002). Social categorization moderates social projection. *Journal of Experimental Social Psychology, 38*, 219-231. doi:10.1006/jesp.2001.1503

Coats, S., Smith, E. R., Claypool, H. M., & Banner, M. J. (2000). Overlapping mental representations of self and in-group: Reaction time evidence and its relationship with explicit measures of

group identification. *Journal of Experimental Social Psychology, 36*, 304–315. doi:10.1006/jesp.1999.1416

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of consumer research, 20*(2), 303-315. Retrieved from http://www.jstor.org/stable/2489277

Hengemühle, L. (2016). *Zijn indirecte schattingen bruikbaar om de prevalentie van strafbaar gedrag te achterhalen?* (Bachelor thesis, University of Twente, Enschede, The Netherlands)

Hogg. A. (2013). Intergroup Relations. In: J. DeLamater and A. Ward (eds.), *Handbook of Social Psychology* (pp. 533-561). doi:10.1007/978-94-007-6772-0_18.

John, L.K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science, 23*(5), p. 524-532. doi: 10.1177/0956797611430953

Krueger, J. I. (2012). Social projection as a source of cooperation. *Current Directions in Psychological Science, 22*(4), 289-294. doi:10.1177/0963721413481352

Krueger, J. I., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology, 67*(4), 596-610. doi:10.1037/0022-3514.67.4.596

Krueger, J. I. & Stanke, D. (2001). The role of self-referent & other-referent knowledge in perceptions of group characteristics. *Personality and Social Psychological Bulletin, 27*, 878-888. doi:10.1177/0146167201277010

Leach, C.W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., & Ouwerkerk, J.W. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology, 95*(1), 144-165. doi:10.1037/0022-3514.95.1.144

Nisbett, R. E., & Kunda, Z. (1985). Perception of social distributions. *Journal of Personality and Social Psychology, 48*(2), 297–311.

Otten, S. (2004). Self-anchoring as predictor of in-group favoritism: Is it applicable to real group contexts? *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 22*(4-5), 427-443.

Pickett, C. L., Silver, M. D., & Brewer, M. B. (2002). The impact of assimilation and differentiation needs on perceived group importance and judgments of ingroup size. *Personality and Social Psychology Bulletin, 28*(7), 546-558.

Politie. (n.d.). Cybercrime. Retrieved from https://www.politie.nl/themas/cybercrime.html

Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review, 9*(1), 32-47. doi:10.1207/s15327957pspr0901_3

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias and social perception and attribution processes. *Journal of Experimental Social Psychology, 13*(3), 279–301. doi:10.1016/0022-1031(77)90049-X

Sahdra, B., & Ross, M. (2007). Group identification and historical memory. *Personality and Social Psychology Bulletin, 33*(3), 384-395. doi:10.1177/0146167206296103

Tcherni, M., Davies, A., & Lizotte, A. (2015). The dark figure of online property crime: Is cyberspace hiding a crime wave? *Justice Quarterly 33*(5), 890-911. doi:10.1080/07418825.2014.994658

Van der Laan, A.M., & Goudriaan, H. (2016). *Monitor Jeugdcriminaliteit 2015. Ontwikkelingen in de jeugdcriminaliteit van 1997 tot 2015*. Den Haag: WODC/CBS.

Van Veelen, R., Otten, S., Cadinu, M., & Hansen, N. (2015). An integrative model of social identification: Self-stereotyping and self-anchoring as two cognitive pathways. *Personality and Social Psychology Review, 20*(1), 3-26. doi:10.1177/1088868315576642

Van Veelen, R., Otten, S., & Hansen, N. (2013). Social identification when an in-group identity is unclear: The role of self-anchoring and self-stereotyping. *British Journal of Social Psychology 52*(3), 543-562. doi:10.1111/j.2044-8309.2012.02110.x

Van Veelen, R., Otten, S., & Hansen, N. (2011). Linking self and ingroup: Self-anchoring as distinctive cognitive route to social identification. *European Journal of Social Psychology, 41*(5), 628-637. doi:10.1002/ejsp.792

Voci, A. (2006). The link between identification and in-group favouritism: Effects of threat to social identity and trust-related emotions. *British Journal Of Social Psychology,45*(2), 265-284. doi:10.1348/014466605X52245

Wieschebrock, A. (2016). *Wat zijn de factoren die de Indirecte Schattingen van Jongvolwassenen over Cybercriminaliteit beïnvloeden?* (Research Project, University of Twente, Enschede, The Netherlands).

Zebel, S., de Vries, P., Giebels, E., Kuttschreuter, M., & Stol, W. (2013). *Jeugdige daders van cybercrime in Nederland: Een empirische verkenning.* WODC.

Appendix

Introduction
Welcome!
Before starting the questionnaire, I would like to request your consent for the participation in this study, which is part of a Bachelor thesis research in Psychology of the University of Twente. In general, the survey involves a short personal memory task in addition to questions about your opinion of cybercrime and its prevalence rates.
Participation in this study is completely voluntary and will require approximately 15 minutes.
All information you provide will remain confidential and will not be associated with your name. The data is also exclusively used for this research.
If you agree to participate, please be aware that you are free to withdraw at any point throughout the duration of the experiment without any negative consequences.
If you have any further questions concerning this study, please feel free to contact me!
(Contact information)

I have read the description of the study and agree to the terms as described above.

Group membership
To which group/association do you belong? (CS group or sports group)

Personal memory task (Identification manipulation)
Low identity (Differentiation Need Arousal) condition:
Please take a moment and think of times when you felt too similar to other members of…, so similar that you felt uncomfortable. In other words, think of times and situations where you felt that you were so much like the other members of … that you could not "stick out," that your own self was hidden. Please write a brief description of two memories of such times.

High identity (Assimilation Need Arousal) condition:
Please take a moment and think of times when you felt too different from other members of…, so different that you felt uncomfortable. In other words, think of times and situations where you did not feel that you fit in with the other members of … and that you "stuck out." Please write a brief description of two memories of such times.

Indirect estimations (0-100%)
The questions that follow refer to your opinion about the criminal behavior of all other group members of (group) in the previous 12 months.
If you do not know what the described behavior means, please look at the description above every question.

How many percent of your group members do you think have created and/or spread ransomware in the previous 12 months?

> Ransomware is a type of malware that prevents or limits users from accessing their system, either by locking the system's screen or by locking the users' files unless a ransom is paid.

How many percent of your group members do you think have created and/or spread other malware like viruses or worms in the previous 12 months?

> A computer virus is a program or piece of code that is loaded onto a computer without knowledge and brings the system quickly to a halt while using all available memory. A computer worm does not need to copy itself to a host program or require human interaction to spread.

How many percent of your group members do you think have hacked someone else in the previous 12 months?

> Hacking includes the unauthorized intrusion into a computer/phone/tablet or a network. An example would be to log in the Facebook account of somebody else.

How many percent of your group members do you think have committed a DDoS attack in the previous 12 months?

> A DDoS attack makes an online service (server, website or other network resource) unavailable by overwhelming it with traffic from multiple sources which forces it to slow down or even crash and shut down and thus denying service to legitimate users or systems.

How many percent of your group members do you think have committed cyber mobbing in the previous 12 months?

> Cyber mobbing includes for example threatening someone else via internet, like social media. Threating someone else with the spreading of erotic photos of this person would also be cyber mobbing, just like posting mean comments on social media directed to another person.

How many percent of your group members do you think have illegally downloaded media (videos/music etc.) from the internet in the previous 12 months?

> Illegally downloading media violates a copyright law through new technologies, including peer-to-peer sharing of digital files. For example, if you did not pay for a song, movie or other media file that has a copyright, then downloading that file is a crime.

How many percent of your group members do you think have committed identity theft in the previous 12 months?

> Identity theft includes the illegal use of personal data (like the PIN code or log-in data) or for example paying bills in the name of the victim.

How many percent of your group members do you think have committed a bargain and sale fraud in the previous 12 months?

Bargain and sale fraud includes for example not paying for something that you bought online or selling something but after receiving the price not sending the product.

How many percent of your group members do you think have committed child pornography in the previous 12 months?

Child pornography includes viewing, having, selling spreading or produce pornographic materials that exploit or portray a minor (under the age of 18), for example having a selfie that involves underage nudity or sexual situations on your cellphone.

How many percent of your group members do you think have committed any form of cybercrime in general in the previous 12 months?

Now think of all possible forms of cybercrime like the offenses previously mentioned in addition to other forms that you can imagine.

Manipulation control – Identification scale (Leach et al., 2008)

Please indicate to what extent you agree with the following statements. . (5 point Likert scale ranging from strongly agree to strongly disagree)

The group to which the following questions are referred to includes all members of (group).

(Group-Level) Self-Investment

Solidarity

1. I feel a bond with my group.
2. I feel solidarity with my group.
3. I feel committed to my group.

Satisfaction

4. I am glad to be a member of my group.
5. I think that my group has a lot to be proud of.
6. It is pleasant to be a member of my group.
7. Being a member of my group gives me a good feeling.

Centrality

8. I often think about the fact that I am a member of my group.
9. The fact that I am a member of my group.is an important part of my identity.
10. Being a member of my group.is an important part of how I see myself.

(Group-Level) Self-Definition

Individual Self-Stereotyping

11. I have a lot in common with the average person of my group.
12. I am similar to the average person of my group.

In-Group Homogeneity

13. People of my group have a lot in common with each other.
14. People of my group are very similar to each other.

Group knowledge

Please indicate to what extent you agree with the following statements. (5 point Likert scale ranging from strongly agree to strongly disagree)

The group to which the following questions are referred to includes all members of (group).

I know all of my group members well.

I know only few members of my group really well.

Social projection

Please indicate the extent of your agreement to the following statements, which refer to the earlier questions about the behavior of your group members. (5 point Likert scale ranging from strongly agree to strongly disagree)

1. By estimating the prevalence of the different cyber delicts, I used knowledge of my own behavior in order to make assumptions about the behavior of the other group members.
2. I based the estimations about other people of my group on my own experiences.
3. My own behavior is similar to the behavior of the other group members.
4. By estimating the prevalence of the different cyber delicts, I expected that the other group members behave unequal to me.
5. I know for sure whether the other group members committed each offense.
6. I randomly answered the estimations about other peoples' criminal behavior, because I do not know what the other people did in the past.

Social desirability group level
Please indicate to what extent you agree to the following questions. (5-point Likert scale ranging from strongly agree to strongly disagree)

I wanted to represent my group in a positive way by answering the questions about other group members' behavior.
I thought that my group will be judged negatively if I have indicated that many people have committed cybercrime.

Direct estimations
Please indicate whether you committed the following offenses in the previous 12 months.
Below, there is a brief description of the specific crimes, if you do not know what they imply. (yes/no)

Did you …
… created and/or spread ransomware?
… created and/or spread malware like viruses or worms?
… hacked someone else?
… committed a DDoS attack?
… committed cyber mobbing?
… illegally downloaded media (videos/music etc.) from the internet?
… committed identity theft?
… committed a bargain and sale fraud?
… committed child pornography?
… committed any form of cybercrime or digitized crime in general (like the offenses previously mentioned in addition to other forms that you can imagine)?

Social desirability individual level
Please indicate to what extent you agree to the following questions. (5-point Likert scale ranging from strongly agree to strongly disagree)

I wanted to represent myself in a positive way by answering the questions about my own behavior.
I thought that I will be judged negatively if I have indicated that I committed cybercrime.

Control measure – Attitude
Please indicate to what extent you think positive or negative about the following offenses: (5-point Likert scale ranging from extremely negative to extremely positive)
… creating and/or spreading ransomware

… creating and/or spreading malware (like viruses or worms)
… hacking someone else
… committing a DDoS attack
… committing cyber mobbing
… illegally downloading media from the internet
… committing identity theft
… committing bargain and sale fraud
… committing child pornography
… committing any form of cybercrime or digitized crime in general

Please indicate to what extent you find the following offenses acceptable or unacceptable: (5-point Likert scale ranging from totally acceptable to totally unacceptable)
… creating and/or spreading ransomware
… creating and/or spreading malware (like viruses or worms)
… hacking someone else
… committing a DDoS attack
… committing cyber mobbing
… illegally downloading media from the internet
… committing identity theft
… committing bargain and sale fraud
… committing child pornography
… committing any form of cybercrime or digitized crime in general

Control measure – Honesty (5-point Likert scale)
Did you understand all questions?
Did you participate seriously in this research?
Did you honestly answer the questions about your own criminal behavior?
Did you honestly answer the questions about other people's criminal behavior?
Do you have confidence that your answers stay anonymous and are treated trustworthy?

Demographic variables
What is your gender? (Male/Female)
What is your age?
What is your nationality? (Dutch, German, Other)
What is the highest level of your education? (Secondary school: e.g. HAVO, VWO, Abitur, highschool; Vocational education: e.g. MBO, Ausbildung; Bachelor degree; Master degree)

Debriefing
Thank you for your participation!

This study was designed to understand to which extent indirect measurements, like the questions asked about the criminal behavior of others, are helpful in finding actual prevalence rates of cybercrime offenses.
The personal memory task, that you did in the very beginning of the survey, took two different forms. The task was designed so that its goal was to either increase the identification of the participants with their group in one half of the participants, and decrease the identification with the group in the other half of the participants. The aim of this study is to examine whether this was of particular influence in making the estimations about other people's criminal behavior.
If you have any questions concerning the survey and/or are interested in the results, feel free to send me an e-mail: (Contact information)