DECRYPTING DIALOGUE: THE EFFECT OF EXCHANGING BEHAVIOUR IN RANSOMWARE NEGOTIATIONS

Leon Küsters

Supervised by MBA Michalis Georgiou, Prof. Dr. Ellen Giebels

University of Twente

Faculty of Behavioural, Management, and Social Sciences

Psychology of Conflict, Risk, and Safety

04.06.2024

Abstract

This paper examines the role of exchanging behavior in the context of ransomware negotiations. Ransomware is forecasted to be one of the most important cybersecurity topics for the coming years and an informed negotiation approach is an important aspect of any defense strategy. Inspiration was drawn from the related field of crisis negotiation to synthesize the hypotheses. These expected a beneficial effect of the IV exchanging behavior on the DV's likelihood of reaching a negotiated agreement and on the magnitude of discount cybercriminals are willing to offer. The dataset (N = 25) consisted of open-access negotiation transcripts. To test the hypotheses, a mixed methods approach paired qualitative coding with data analysis. A t-test probed for a difference in exchanging frequency between paid and unpaid transcripts and a regression analysis tested for an effect of exchanging frequency on ransom discount offered. While the directions of the effects were as expected, the findings failed to support the hypotheses in a statistically significant manner. However, other exploratory findings support the efficacy of exchanging. Upon reflection, the sample size severely limited the power of the test and the likelihood of finding significant findings. Nonetheless, the results point to a possibly different role of trust in the context of ransomware negotiations and provide a starting point for further research.

Keywords: Ransomware, Table of Ten, Crisis Negotiation, Cybercrime

DECRYPTING DIALOGUE: THE EFFECT OF EXCHANGING BEHAVIOUR IN RANSOMWARE NEGOTIATIONS

Cybercrime is one of the biggest threats to society for the foreseeable future. In 2020 alone, the FBI estimates that more than 4 billion US dollars have been lost as a consequence of cyberattacks in the US (Bureau of International Narcotics and Law Enforcement Affairs, 2023). EUROPOL and the World Economic Forum also see cybercrime as a critical issue. Both institutions have recently identified cybercrime as one of the top risks in their "Global Risk Report" and policy cycle (EU Policy Cycle - EMPACT | Europol, 2023; World Economic Forum, 2023). Responsible for this is often the attack pattern called Ransomware. Ransomware is a type of virus that holds digital files hostage by encrypting them and then withholding the decryption key unless a ransom is paid (Beaman et al., 2021). These attacks are extremely popular, causing ransomware to be one of the most potent and financially damaging cybercrimes in the world.

Once a network or device has been infected by such a virus, the attack enters a new stage. Threat Actors (TA) usually provide a link to a private chat where victims can converse with cybercriminals (Beaman et al., 2021). In these chats, victims can then negotiate a lower buyout, stall for time, or gain more information. When doing so, both parties unavoidably make use of expressions intended to influence their counterpart. Such influencing strategies are often critical to the negotiation outcome (Giebels & Noelanders, 2004). However, until now, no research exists that examines the effect of influence strategies in ransomware negotiations specifically. To gain a more intimate understanding of these Influence Tactics (IT) this paper will focus on the role of a specific influence strategy, *Exchanging*, that plays a crucial part in many (other)

negotiations (Vetschera, 2013; Olekalns & Smith, 2000). First, this paper will give an overview of the status in the research field of cybercrime, tap into the specific domain of ransomware attacks, and discuss related areas of research, particularly within the domain of crisis/hostage negotiation. Afterward, this research will analyze the effect of exchanging behavior in ransomware negotiations before finally reflecting on the practical and theoretical implications of the findings.

Theoretical Framework

Cybercrime encompasses a wide range of illegal activities. Generally, any illegal activity that uses a computer system in cyberspace can be classified as cybercrime (Chitadze, 2023). It can be anything from hacking, harassment, or online fraud to online child pornography. Researchers have theorized about the motives for cybercrime for a while. Popular theories that attempt to explain criminal behavior in cyberspace are Routine Activity Theory (RAT) and Social Learning Theory (SLT). RAT illustrates that criminal behavior is likely to occur when three factors converge: (1) A motivated offender, (2) a suitable target, and (3) the absence of a capable guardian. However, RAT was originally designed to describe offline behavior and its validity in cyberspace is a matter of debate (Leukfeldt & Yar, 2016). SLT, on the other hand, focuses on the role of socialization and observational learning. It emphasizes the importance of peers and authority figures that seem to have a significant influence on cybercriminals' values and beliefs (Holt et al., 2010). A more recent approach based on SLT is the focus on subcultures. Holt (2019) examined the formation and behavior of deviant subcultures in cyberspace. He found that many of these subcultures not only influence the norms and values of cybercriminals but are

also one of the reasons why they engage in criminal activities. The motives behind cybercrime are dictated by opportunity, the social environment, and deviant subcultures.

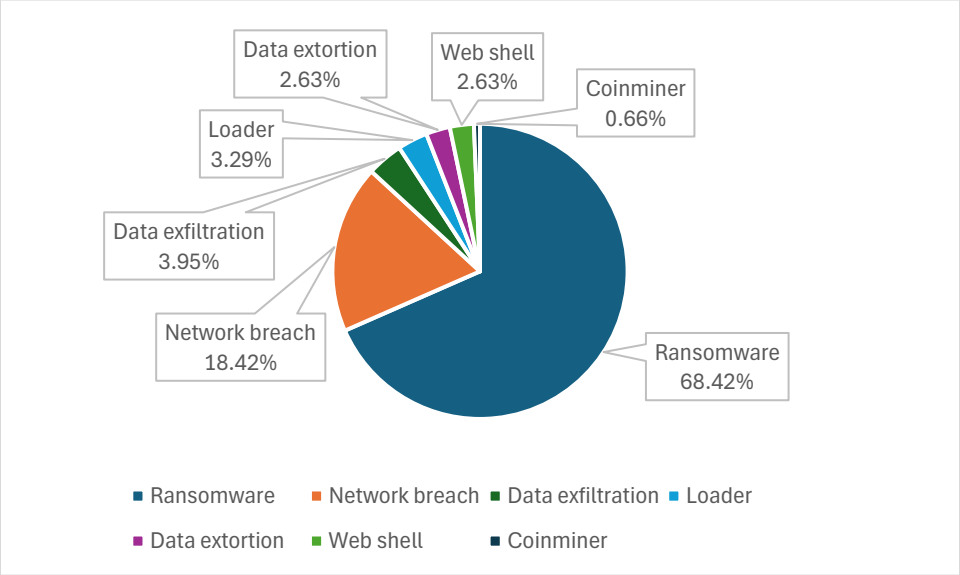
Wanting to examine these motives in practice, Lee et al., (2023) divided cybercrimes into three distinct groups (integrity-related, computer-related, content-related) and analyzed the reasons behind each cluster. The information was collected by interviewing police officers closely involved in the prosecution of cybercriminals and asking for their insights. They found that integrity-related crimes like hacking were supposedly treated as fun challenges, often out of a desire for recognition. Meanwhile, computer-related acts like fraud or ransomware attacks were reported to have financial motives and content-related cybercrimes such as cyberdefamation or cyberstalking seem to be motivated by more expressive reasons like the expression of anger or frustration. In practice, the aims cybercriminals hope to achieve vary based on the types of cybercrimes, with Ransomware likely to be financially motivated.

Technical Details of Ransomware

Ransomware is the leading type of attack compared to all other cyberattack types (see Figure 1). It accomplishes its goals by holding files hostage can be divided into three types of virus strains - scare, locker, and cryptoware (Andronio et al., 2015). First, scareware is a tool that floods the victim’s device with pop-up ads, hoping to manipulate the user to download further malicious software. lockerware, however, does interfere more directly with the target system. It can encrypt files or partially block the primary functions of a device. Then, cryptoware can completely block a device or network from being used by encrypting vital and/or sensitive files (Beaman et al., 2021). All types of ransomware are dangerous, but cryptoware is the most lethal of the three and gives the most leverage to attackers.

Figure 1

Distribution of detected cyberattacks worldwide (in 2022)



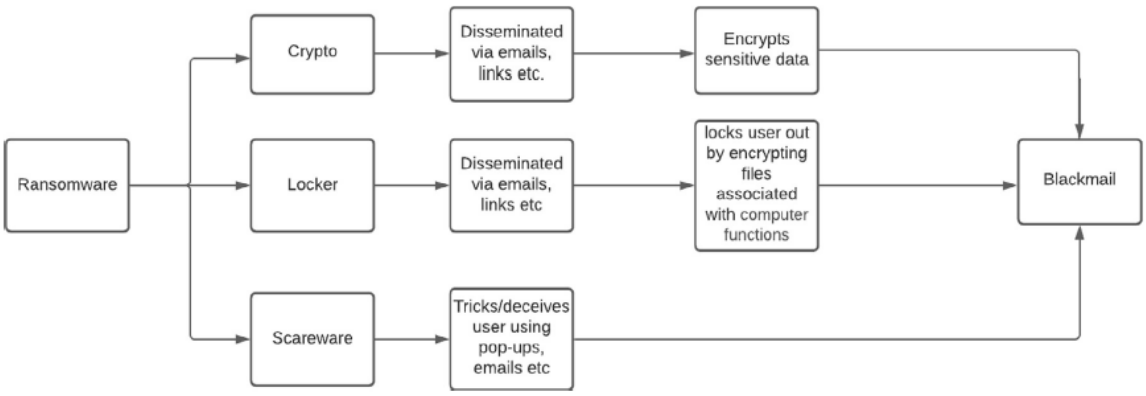
Note. From "Distribution of detected cyberattacks worldwide in 2022, by type", by Sophos X-Ops Incident Response detections, 2023, https://www.statista.com/statistics/1382266/cyber-attacks-worldwide-by-type/

That is because locker- and scareware are comparably easy to overcome, while crypto-ransomware is exceedingly difficult to solve. The cause of this is that crypto ransomware encrypts files on the target system using an encryption key which is then encrypted again using a different key (see Figure 2). This master key is only accessible to the attackers (Beaman et al., 2021). Without this key, modern encryption is virtually impossible to reverse (Gómez-Hernández et al., 2018). The only reliable way to recover the data is by paying a ransom to gain access to the master key. To make the victim pay the ransom, perpetrators typically leave behind a ransom note. On this note, the victim can find information regarding the state of the attack and steps describing how to proceed (Beaman et al., 2021). The sophistication of modern

Ransomware attacks gives victims little to no chance of recovering their data without paying a ransom or resetting their systems.

Figure 2

Details of Ransomware-Strains



Note. From Beaman et. al (2021)

This conundrum is exacerbated by new technical and strategic developments in the world of Ransomware attacks. First off, most Ransomware attackers now use a double extortion scheme where they not only encode data but also extract possibly sensitive files from their victim's devices. Threatening to publish this data gives criminals additional leverage in negotiations and allows perpetrators to set a higher ransom (Meurs et al., 2023). Additionally, the emergence of Ransomware as a Service (RaaS) has allowed even laypeople to efficiently employ Ransomware. RaaS is a relatively new business model in which criminal organizations use their expertise to build and sell their software to future perpetrators (Beaman et al., 2021). These development kits do not require technical expertise and can be executed with minimal effort.

Ransomware has not only become easier to access but has also found new avenues to increase the leverage of attackers in negotiations.

The consequences of a successful attack can be severe. “The state of Ransomware” an annual survey commissioned by the security firm Sophos (2023) showed that the mean recovery cost of a Ransomware attack was $1.82M in 2023, scaling with yearly revenue. This value excludes ransom payments, which were made in 47% of cases. Aside from the immediate costs of resetting the systems or paying the ransom, Ransomware attacks also have a direct impact on a business’s revenue. While under attack, systems are inoperable, and restoring them usually takes between a week and a month (Sophos, 2023). Thus, victims are faced with the possibility of losing out on business for a significant amount of time. For those reasons, targets of Ransomware attacks often have clear incentives to pay the ransom and do so in a significant number of cases.

The Victim’s Perspective

Nowadays, most businesses have strategies for combatting Ransomware attacks. These are primarily based on taking preventative action. Falco et al. (2019) interviewed seven Urban Critical Infrastructure Operators about their measures against Ransomware attacks. While a sample of seven might not be representative of the general CIO population, the results showed that operators generally focus on establishing leverage through backups or establishing connections to authorities. Post-attack the persons responsible reevaluate their systems and document the lessons learned. Meanwhile, the mid-attack phase is often neglected. This phase is characterized by executing damage control measures and possibly engaging in dialogue with the attacker. If done correctly, the ransom can often be negotiated. Numbers show that nearly three out of four hackers are willing to return the stolen data for a discounted price (Wade, 2021).

Negotiations also offer benefits aside from reducing the ransom amount. Buying time by engaging in dialogue can be crucial to accurately assess the damages and evaluate the value of the files that have been encrypted or stolen (Falco et al., 2019). Taking together, and given its lack of research on it, negotiation appears to be an underdeveloped tool capable of mitigating the consequences of Ransomware attacks.

Despite the benefits, government bodies and some companies categorically oppose paying ransoms to criminal actors or even entertaining the idea of a negotiation. For example, the US Cybersecurity & Infrastructure Security Agency (CISA) categorically discourages victims from paying ransom, going as far as to threaten sanctions (CISA, 2023). Europol concurs and does not recommend paying ransoms, as doing so finances criminal organizations and encourages further cyberattacks (Europol, 2015). The same ethical concerns are the primary reason for businesses categorically refusing to pay in the case of Ransomware attacks (Falco et al., 2019). In line with this, before conducting negotiations with cybercriminals, the ethical implications and the existence of possible regulations need to be considered.

In case the victim is willing to negotiate, inspiration for the best practices can be drawn from Game Theory. By providing a set of rules and levers administrators can effectively mimic zero-sum negotiations. It relies on participants to make rational decisions on how to leverage their advantages (Von Neumann & Morgenstern, 2007). By including as many relevant factors as possible, Game Theory aims to approximate a real-world scenario. This concept has also been applied to Ransomware negotiations. There, the two counterparts have to make tactical decisions regarding software and security investments, ransom demands or counteroffers, all while considering factors like time investment, legal fees, reputational damage, potential downtime, and more (Meurs et al., 2003; Ryan et al., 2022). Game Theory allows for a better understanding

of TAs and allows cybersecurity professionals to consider the effectiveness of their decisions, albeit in a simplified manner.

Through the Lens of Crisis Negotiation

A more in-depth analysis of a negotiator's decision-making could be done with insights gained from the Crisis/Hostage Negotiation literature. Essentially, Ransomware is a way of taking your files as digital hostages, not entirely dissimilar from "traditional" hostage situations. Crisis Negotiation is a more researched field than its ransomware counterpart. Existing research has investigated the most efficient ways of interacting with a person in crisis (PiC). PiC's are, on the one hand, individuals who experience emotional crises and threaten harm to themselves or others, but on the other hand also people like kidnappers or hostage takers who use other individuals as leverage (Vecchi et. al, 2005). Various experts and law enforcement agencies have developed strategies as to how negotiations with these people should be handled. Prominent examples are the Behavioural Change Stairway Model (BCSM), the Structured Tactical Engagement Process, and the S.A.F.E model (Coulthard et. al, 2020). To what extent these are translatable to ransomware negotiations is still a matter of debate.

Nonetheless, all these models have a similar approach to resolving crises. They identify three steps that need to be resolved successively, beginning by trying to shift the PIC to a more receptive state of mind and diffusing emotionality. Once that is established, the opportunity for more immediate relationship-building arises until ultimately negotiators can begin to reason with the subject and achieve behavioral change (Vecchi et. al, 2005; Kelln & McMurtry, 2007; Hammer, 2007). In comparison to traditional negotiations, these tactics are more directly focused on achieving behavioral change and diffusing emotionality, however, other factors sometimes necessitate a slightly different approach.

One of these is the difference between expressive and instrumental crisis negotiations. Expressive crisis situations often arise spontaneously and are classified by a high degree of emotionality and impulsivity. These are often volatile barricade situations, where a person is not taken hostage as a means to an end, but as an emotional, often irrational, reaction (Vecchi et. al, 2005). On the contrary, instrumental situations are rational, calculated acts more akin to bargaining situations (Giebels & Noelanders, 2004). Examples of instrumental crisis situations are extortion or kidnappings where a perpetrator hopes to enforce his demands (Vecchi et. al, 2005). As ransomware attacks are planned, deliberate actions driven by the desire to maximize financial gains, they are more akin to instrumental crisis negotiation scenarios.

An approach that is not based on a step-by-step philosophy but provides a more practical view of ITs is the Table of Ten by Giebels (2002). Created through interviews with Dutch and Belgian police negotiators and analysis of organizational change literature it classifies utterances of both negotiator and PiC as one of ten possible ITs (see Table 1). This classification is loosely inspired by research from Cialdini (2001), who identified six key principles of influence: (1) Reciprocity, (2) Consistency, (3) Social Proof, (4) Authority, (5) Liking, and (6) Scarcity. Beyond their theoretical application, IT's of the Table of Ten can be used to classify separate influence strategies in transcripts or recordings of crisis negotiations.

Giebels (2002) then further divides these ITs into relational and content-based strategies. Relational ITs are attempts to build a personal relationship with their counterpart. The three tactics that fall into this category are *Being Kind*, *Being Equal,* and *Being Credible*. *Being Kind* and *Being Equal*. Content-based ITs are focused on the contents of the message. The seven remaining tactics belonging to this category and their underlying principles can be found in Table 1. After coding and analyzing crisis negotiation transcripts using the Table of Ten, Giebels

& Noelander (2004) concluded that instrumental and expressive negotiation scenarios require different approaches regarding ITs. Kidnappings and extortions have a more symmetric interaction profile, meaning that ITs were more likely to be responded to with similar ITs, particularly content-based ones. Further, statements showing one's credibility seemed to be productive, increasing the PiC's willingness to cooperate. Ultimately, the ITs *Being Credible* and *Exchanging* were deemed the most beneficial in instrumental crisis negotiations.

Table 1

ITs of the Table of Ten

Strategy	Underlying Principle	IT
Being kind	Sympathy	All friendly, helpful behavior
Being equal	Similarity	Statements aimed at something the parties have in common
Being credible	Authority	Behavior showing expertise or proving you are reliable
Emotional Appeal	Self-image	Playing upon the emotions of the other
Intimidation	Deterrence/fear	Threatening with punishment or accusing the other personally
Imposing a Restriction	Scarcity	Delay behavior or making something unavailable
Direct Pressure	Power of repetition	Exerting pressure on the other in a neutral manner by being firm
Legitimizing	Legitimacy	Referring to what has been agreed upon in society or with others
Exchanging	Reciprocity	Give-and-take behavior
Rational persuasion	Consistency	Use persuasive arguments and logic

Note. From "Crisis negotiations: A multiparty perspective", by E. Giebels and S. Noelanders, 2004

The Effect of Culture

A factor that has been found to influence the optimal usage of ITs is culture. Particularly the cultural dimension of uncertainty avoidance has been found to significantly affect the response of PiCs (Giebels et al., 2017). Uncertainty avoidance is an individual's tolerance for uncertain or unknown situations. In their research, Giebels et al. (2017) found that German negotiators (high uncertainty avoidance) used more legitimizing utterances and formal language than their Dutch counterparts (low uncertainty avoidance). Brett (2000) also stresses the importance of culture. In their model of intercultural negotiation, they argue that differences in cultural values, communication styles, and interpretations of power pose significant strategic challenges to negotiation between two parties of different cultures. Culture causes people to see things from different perspectives and understanding these differences can be crucial to any type of negotiation.

Currently, there exists a significant amount of research on crisis negotiation, but little of it examines ransomware scenarios specifically. What does exist is either focused on the technical perspective or of a purely theoretical nature. Although the similarities between crisis- and ransomware negotiations are apparent, it is unclear whether the conclusions of theoretical and practical research can be universally applied. After all, there are some key differences between in-person high-stakes confrontations and remote financially motivated crimes such as Ransomware. This difference, however, is crucial to examine if negotiators want to keep up with the rapid development of ransomware attack strategies and hone their own dialogue-based countermeasures. To do so, negotiators require a clear idea of the effectiveness of certain ITs. As far as is currently known, there has been no empirical research conducted on ITs specifically in Ransomware negotiations.

The current study

This study aims to provide a first view into the usage of the Table of Ten's ITs in ransomware negotiation before concretely focusing on the effect the IT *Exchanging* has on the outcome of ransomware negotiations. Negotiating with an attacker is crucial, as coming to a negotiated solution is often the most financially advisable solution. In these negotiations, exchanging behavior plays a vital role. In a wider sense, exchange behavior is behavior motivated by expected returns (Ahmad et al., 2023). These can be acts like making concessions, offers & counteroffers, but also asking for something in return. In the Table of Ten specifically, *Exchanging* is conceptualized similarly. It is based on the feeling of obligation to respond in kind when a person is treated a certain way in line with the idea of reciprocity (Giebels & Noelander, 2004). Reciprocity makes others feel obligated to return a favor, gesture, or behavior, causing them to respond in a similar way (Cialdini, 2001). These behaviors are central to any negotiation, whether crisis-related or traditionally distributive as exchanging information, standpoints, or engaging in dialogue allows for a compromise to be found and negotiations to be resolved.

Providing empirical evidence for honing negotiators' usage of ITs could help improve the outcomes and likelihood of finding said compromise. Additionally, productively engaging in dialogue can not only reduce the ransom buyout but also buy time for a more informed response and prevent the leakage of personal data. Thus, this research will examine the following research question: *To what extent is the usage of the IT Exchanging related to the negotiation outcome?*

The exchange of information and personal standpoints is crucial for the completion of a negotiation, as providing an offer or counteroffer indicates a willingness to make concessions and productively participate in the negotiation (Vetschera, 2013). As shown in the article by Falco et. al (2019) not all negotiators are aiming to productively participate in negotiations.

Despite that, some of them might engage in a dialogue with the perpetrators anyway, f.e. for reasons besides financial ones. In these scenarios, negotiators plausibly use less exchanging behavior, as f.e. proposing (counter-)offers and trading standpoints would imply a willingness to pay the ransom. For that reason, in cases where there exists categorical opposition to paying a ransom, the negotiation could feature a lower relative frequency of exchanging behavior, particularly from the negotiator. For that reason, this paper will test the following hypothesis:

H1: *Negotiation transcripts where a ransom was paid show a higher frequency of the negotiator's usage of Exchanging than those where no ransom was paid*

Olekalns and Smith (2000) have examined exchanging behavior in a wider sense and determined it as one of the most beneficial tools for improving joint gains and negotiation outcomes in distributive negotiations. Another lesson from research into negotiation dynamics is that influence behavior is often symmetrical. Using a certain influence strategy makes it more likely for the other party to respond in kind (Adair & Brett, 2005). This occurrence is also related to the concept of reciprocity. For that reason, the usage of *Exchanging* is likely to cause more *Exchanging*. Concrete evidence of this is given by Giebels & Noelander (2004), who coded real crisis negotiation logs and found the described interaction pattern between the negotiator and PiC. With the previously established benefit of exchanging behavior, this repetition can be likened to a productive feedback loop potentially related to positive negotiation outcomes like finding an agreement or reducing the offered buyout. For that reason, the second hypothesis tested in this paper is:

H2: *A higher usage frequency of the IT Exchanging by the negotiator has a positive relationship with the amount of discount the TA is willing to offer.*

This research hopes to show practical and theoretical implications for Ransomware negotiations. Analyzing the effectiveness of negotiation strategies is a significant step toward optimizing negotiation tactics to achieve a satisfactory negotiation result. This information could lead to a better understanding of TAs and enable negotiators to consciously reflect on their usage of negotiation strategies depending on the scenario. With the emergence of Ransomware as one of the most prevalent cybercrimes negotiators need to be informed on the best practices to bargain with cybercriminals. This research could not only lead to a significant financial benefit for affected companies but also prevent the leakage of sensitive information to other cybercriminals or business rivals. Additionally, the results of this paper could provide a starting point for further research into negotiation strategies.

Method

Design

To investigate the research question "*To what extent is the usage of the IT Exchanging related to the negotiation outcome*?" a mixed methods approach was used. In this, deductive qualitative coding based on the Table of Ten (Giebels & Noelander, 2004) was paired with two quantitative statistical analyses. Coding the data using grounded theory allowed for an objective analysis of IT's and created a dataset suitable for statistical analysis. The Table of Ten was used as a scientifically validated but also practically applicable tool to code the negotiation transcripts. Then, the quantitative aspects of this study were set up to deepen the understanding of said

dataset and allow for drawing reliable, evidence-based conclusions. The study was conducted between the 5th of February 2024 and the 20th of June 2024.

Coding

Sample selection

First, for the qualitative analysis, 25 transcripts of ransomware negotiations were coded. These anonymized transcripts of ransomware negotiations were acquired from the website created by Marchive (n.d.). Out of all the available transcripts, a set of inclusion criteria was applied to exclude unsuitable data from the pool. These were: (1) A number between 35 to 80 messages, (2) the presence of TA groups roughly representative of the related logs in the database, and (3) approximately 50% of transcripts paid and not paid. The reasoning for the choices was to (1) guarantee analysis of productive exchanges and limit time investment for the researchers, (2) have a stratified dataset representative of the landscape of ransomware groups, and (3) ensure the possibility of reliable statistical analysis related to negotiation outcome.

After applying the criteria, a stratified sampling approach was used to select the final dataset. The goal of such an approach is a dataset that is representative of the population. Since the true population of ransomware attacks and chat logs is unknown, the closest estimate is the available transcripts from Marchive (n.d.). For that reason, the available transcripts were divided based on the criminal actors responsible for the attack. Each of these groups was then allotted a number of transcripts depending on the ratio of transcripts they were responsible for in comparison to the available total. However, it was necessary to ensure as close to a 50/50 split in terms of paid/not paid to ensure the sample size per group is acceptable. Hence, some of the initially selected datasets had to be replaced. To uphold the scientific integrity of the sample selection, unpaid transcripts were replaced by paid transcripts from the same criminal networks.

After conducting the sampling, the average number of messages per transcript was 55.52 (*SD* = 12.11). In twelve cases, the negotiation was classified as resolved, with a ransom being paid, while the opposite was the case in thirteen transcripts. The ransom amount ranged from $1300 to $15.000.000 with an average of $2.058.678. An Interrater-reliability coefficient (Cohen's Kappa) of .68 was achieved for the 25 transcripts analyzed in this study. For more information on the dataset's characteristics like threat groups and number of messages refer to Appendix A.

Coding Procedure

This part of the research made use of deductive coding, as the data needed to be systematically classified and synthesized in order to enable the quantitative analysis. Coding and analysis were handled by the author in conjunction with other graduate bachelor's students of the Psychology department of the University of Twente, the Netherlands. Before beginning the coding of the logs, roughly ten hours of training were conducted. To avoid the usage and resulting bias of real negotiation logs, generative AI was employed to create new, fictional, transcripts. The open-access model was trained with real negotiation logs and asked to provide transcripts based on the given information. All in all, 21 transcripts were analyzed for practice purposes. After the first practice session of coding, Cohen's Kappa was equal to .37.

The real sample, consisting of the 25 selected transcripts, was then downloaded and systematically analyzed in Atlas.TI. This was done according to a pre-established coding scheme and ruleset (see Appendix B and C). In total, 22 codes were applied, eleven each for perpetrator and victim. 20 of those represent the usage of ITs and two are about the relay of general information not attributable to any of the ITs of the Table of Ten (Giebels & Noelander, 2004). The decision was made to assign only one IT per message since many messages were

making use of different combinations and assigning only the most important central code allows for more reliable coding and increases intercoder agreement. All transcripts were coded by each researcher on an independent basis. Afterwards, every code in the transcripts was compared and discussed with a decision being made on which is the most appropriate choice. This was done to ensure agreement and make the data suitable for statistical analysis.

Contrary to similar research approaches (see Giebels & Noelanders, 2004; Beune et al.,2020), Codes were not assigned per speaking turn, but per message in the chat logs. This decision was made since in the chat logs, chatters would often send multiple messages in a row. Usually, after each speaking turn one would code the response - or lack thereof - of the other party. However, seeing as the lack of response was often unrelated to an intention to influence the counterpart, but to the difference in time zones, short time intervals between messages, or other lack of availability. For that reason, assigning the non-response of a party an IT would not be an accurate assessment of the negotiation dynamic. Assigning one IT per utterance is more representative of the employed negotiation strategy. This decision resulted in the necessity to calculate the frequencies of ITs respective to the negotiation party's total utterances to accurately represent the usage of ITs. Therefore, data related to discounts or IT usage is displayed in ratio variables or mentioned as percentages.

The dependent variable was calculated based on the information provided by Marchive (n.d.). For the first research question, the transcripts were coded as 1 (paid) and 0 (unpaid). For the second research question, the calculation included was based on initial ransom demand which was used to create a ratio variable (0 – 1) depending on what percentage of discount the TA was willing to offer. Notably, this methodology was applied for both paid and unpaid transcripts. So,

if a TA offered a discount, but the negotiator refused to pay, the offered discount was still used for the statistical analysis.

Data Analysis

After finishing the coding, the data was transferred and collected in Microsoft Excel and formatted to be suitable for statistical analysis. Then, to investigate the first hypothesis, a one-sided independent samples t-test was conducted to compare the transcripts wherein a ransom was paid to those where no payment was made. Under investigation was the frequency of the IT *Exchanging*. Conducting a t-test allows for the testing of significant differences between two groups. This way, the results show whether the difference between the usage of *Exchanging* in the transcripts could be an explanation of a ransom being paid.

For the second research question, a correlational analysis was paired with linear regression. The latter used the frequency of *Exchanging* as an IV and the percentage of ransom *Discount offered* by the TA as the DV. This approach allows for a general overview of the effect the IT's have on the willingness of the cybercriminal to offer a discount while also taking into account the magnitude of the effect a higher *Exchanging* frequency could have. The linear regression was conducted to examine the effects of using *Exchanging* on a more detailed level.

Upon completion of the coding, statistical analysis was conducted using the analytics software *R*. For the first research question a one-sided T-test was conducted to test for a difference in group means of *Exchanging* usage. The significance threshold was set to $\alpha = .05\%$. To answer the second research question, an exploratory correlation analysis was conducted to glean the possible relationship between the IV and DV. Then, a linear regression model was fitted to the data to examine the relationship between the IV *exchanging frequency (Victim)* and the DV *Discount offered (%)*. The DV was measured as a ratio of the initial ransom demand to

ensure that outliers with comparably large initial ransom demands or discounts do not skew the results. Afterwards, an exploratory mediation analysis was conducted with TA *Exchanging* frequency as the MV, the negotiator *Exchanging* frequency as the IV, and *Discount offered (%)* as the DV.

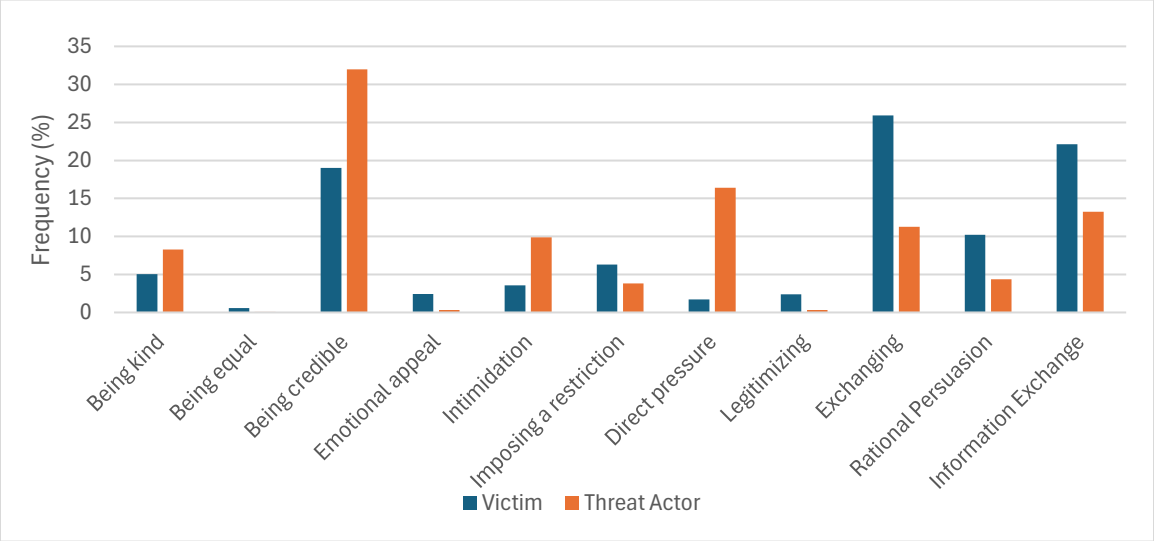
Results

Descriptive Statistics

In total, 1386 messages were coded. Of those, 1143 (82.5%) could be classified as IT's of the Table of Ten. 230 (16.6%) were Information Exchanges and 13 messages were not clearly identifiable, mostly due to sensible information being redacted. Of those 1386 ITs, 465 (40.7%) were of a relational nature, and 678 (59.3%) were of a content-based nature. Combining both negotiator and TA, the most-used IT was *Being Credible* with 367 codes. The least-used IT was *Being Equal* with only five clearly identifiable cases. An overview of how the usage of ITs differed between negotiators and TAs is shown in Figure 2. A detailed breakdown of the frequency of each IT is displayed in Table 2.

Figure 3

Usage of ITs in Ransomware Negotiation



Note. N = 25, All values in per cent relative to total summed messages per party

Table 2

IT Usage Frequency by Negotiator and TA per Transcript

	Negotiator			TA		
IT	*M*	*SD*	*Range*	*M*	*SD*	*Range*
Being kind	5%	5.4%	0-21.3%	8.3%	5.6%	0-22.2%
Being equal	0.6%	2%	0-7.9%	0.1%	0.4%	0-2.2%
Being credible	19%	9.6%	0-35.3%	32%	16.3%	0-66.7%
Emotional Appeal	2.4%	3.3%	0-10%	0.3%	1.2%	0-5.4%
Intimidation	3.6%	6.2%	0-20%	9.9%	8.3%	0-25.6%
Imposing a restriction	6.3%	8.5%	0-33.3%	3.8%	3.5%	0-11.5%
Direct pressures	1.7%	4.6%	0-21.1%	16.4%	10.4%	0-46.7%
Legitimizing	2.4%	3.4%	0-12.1%	0.3%	0.9%	0-3.3%
Exchanging	25.9%	12.5%	3.9-50%	11.3%	7.5%	0-26.9%
Rational Persuasion	10.2%	9.1%	0-35%	4.4%	4.3%	0-13.5%
Information Exchange	22.1%	12.9%	5-47.4%	13.2%	8%	0-31.6%

Note. N = 25, All values are calculated by examining relative frequencies per party per transcript

Of special interest for this research is the IT *Exchanging*, particularly the usage by the negotiator. Summed from all transcripts, *Exchanging* was used 237 times, 17.3% of total IT usage. Out of those, 154 were attributable to the negotiator, amounting to 25.9% of all influence strategies used by that party, the most of any IT. For the TA, 83 cases of *Exchanging* occurred, corresponding to the third most used negotiation strategy with a frequency of 7.5%. Notably, the negotiator used *Exchanging* in every transcript, with the minimum frequency being 3.9%. Remarkable is also the standard deviation of 12.5% for the negotiator's usage of *Exchanging*, indicating a significant amount of variance. This means that there are large differences in the amount of *Exchanging* the negotiator has used in the transcripts. *Exchanging* is one of the most prominent ITs, however, its usage rate is variable.

Next, the dynamics of ransoms & offered discounts were examined. The initial ransom demands ranged from $1300 as the lowest to $15.000.000 as the highest. The average initial ransom demand was $2.058.678 (*SD* = $6.163.743) and the average *Discount offered* post-negotiation was $475.586 (*SD* = $797.281). Notably, in one case there was no information about the magnitude of the ransom demand, this transcript was successively removed from all analyses relating to the ransom amount. In total, twelve of the negotiations were coded as resolved, with a ransom being paid and data being recovered. Respectively, in thirteen cases the negotiations were unresolved, with no ransom being paid and no data being recovered. More information on the IT's the negotiator used in paid and non-paid negotiation scenarios can be found in Table 3.

Table 3

Victim IT Usage Frequency by Payment Status on a per Transcript Basis

	Paid			Not paid		
IT	*M*	*SD*	*Range*	*M*	*SD*	*Range*
Being kind	6.1%	5.1%	0-18.8%	4.1%	5.7%	0-21.4%
Being equal	0%	0%	0-0%	1.1%	2.7%	0-7.9%
Being credible	21.1%	8.3%	4-34.2%	17.1%	10.6%	0-35.3%
Emotional Appeal	1.2%	2.2%	0-5.9%	3.6%	3.8%	0-10%
Intimidation	1.9%	5.4%	0-18.8%	5.1%	6.6%	0-20%
Imposing a restriction	6.2%	6.4%	0-20.7%	6.5%	10.4%	0-33.3%
Direct pressures	2.4%	6.1%	0-21.1%	1%	2.6%	0-8%
Legitimizing	1.2%	2.2%	0-6.3%	3.5%	4%	0-12.5%
Exchanging	30.5%	10.4%	15.8-5%	23.0%	13.5%	3.9-45.5%
Rational Persuasion	8.2%	6.1%	0-19.1%	12.1%	11.1%	0-35%
Information Exchange	21.3%	10%	6.3-40.7%	23%	15.5%	0-47.4%

Note. Paid (N =12), Not paid (N = 13), All values are calculated by examining relative frequencies per party per transcript

Beyond the difference between paid and unpaid negotiations, the data showed some relationships between IT usages (see Table 4). Notably, some of the investigated variables have very low occurrence rates. For that reason, the focus will be put on those, more reliable, relationships instead. First, the negotiator using *Intimidation* is related to a statistically significant increase in the amount of *Intimidation* used by the TA ($r = .48$, $p = .01$). Additionally, the negotiator using "Imposing a restriction" shows a significantly positive relationship with the TAs usage of *Being Kind* ($r = .50$, $p = .12$). Next, the negotiators usage of *Rational Persuasion* is significantly related to the TA imposing a restriction ($r = -.45$, $p = .02$). Despite occasionally low sample sizes, some variables show statistically relevant relationships.

Table 4

Correlation Table of IT's, Offered Discount, and Payment Status

	TA Being kind	TA Being equal***	TA Being credible	TA Emotional Appeal***	TA Intimidation	TA Imposing a Restriction	TA Direct Pressure	TA Legitimizing***	TA Exchanging	TA Rational Persuasion	TA Information Exchange	Discount (in %)	Paid (yes/no)
N Being kind	-.07	.63**	-.18	-.18	.13	.04	.36	.22	-.19	.09	-.12	.25	.19
N Being equal***	-.24	-.06	-.35	.66**	.11	.00	.26	-.11	-.10	.30	.28	-.29	-.28
N Being credible	-.04	-.10	-.06	.08	.03	.23	-.05	-.03	-.08	.14	.09	.12	.22
N Emotional Appeal***	.16	-.15	-.26	.22	.01	-.08	.11	-.27	.04	.36	.09	-.08	-.38
N Intimidation	-.35	-.12	-.19	.15	.48*	.02	.23	-.22	-.27	.34	-.09	-.22	-.26
N Imposing a Restriction	.50*	-.15	-.07	-.16	.22	-.30	-.29	.03	.04	-.13	.14	-.09	-.02
N Direct Pressure***	-.41*	-.08	.43*	.11	-.23	.19	-.02	-.14	.06	-.15	-.38	.02	.16
N Legitimizing***	-.29	.29	-.37	.09	.39	-.07	.43*	.07	-.27	.06	.20	-.54*	-.35
N Exchanging	.21	-.09	.33	-.01	-.55**	0.14	-.22	.00	.40*	-.22	-.28	.23	.31
N Rational Persuasion	-.09	-.23	-.20	.00	.19	-.45*	.11	-.09	.20	.30	.02	-.10	-.22
N Information Exchange	-.03	.22	.16	-.16	-.10	.15	-.07	.20	-.23	-.32	.21	.02	-.07
Discount (%)	.12	-.24	.32	-.11	-.34	.32	-.52*	.20	.38	-.11	-.15		.72**
Paid (yes/no)	-.06	-.20	.42*	-.27	-.46*	.21	-.41*	.19	.44*	-.31	-.10	.72**	

Note. N = 25, * = p < *.05,* ** = p < .005, *** = *Fewer than 20 instances found*, N = Negotiator, TA = T

Of those, *Exchanging* is the focus again. *Exchanging* shows two statistically significant correlations with ITs used by the TA. First, it is negatively correlated with the TAs' usage of *Intimidation* ($r = -.55$, $p < .005$). That means that when the negotiator uses more *Exchanging* the TA less frequently uses *Intimidation* in the same transcript. Next, the negotiator's usage of *Exchanging* was also significantly related to the same IT of the TA ($r = .40$, $p = .046$). This means that the more frequently the negotiator engaged in exchanging behavior the more often the TA did so as well. Also, *Exchanging* shows a relevant but statistically insignificant correlation with the *Discount offered* by the TA ($r = .23$, $p = .27$). Thus, in cases where the negotiator used more *Exchanging*, the negotiator occasionally offered a higher ransom discount. Although this correlation analysis is more exploratory, exchanging behavior by the negotiator seems to be related to increases in *Exchanging* as well as ransom *Discount offered* and decreases in *Intimidation* by the TA.

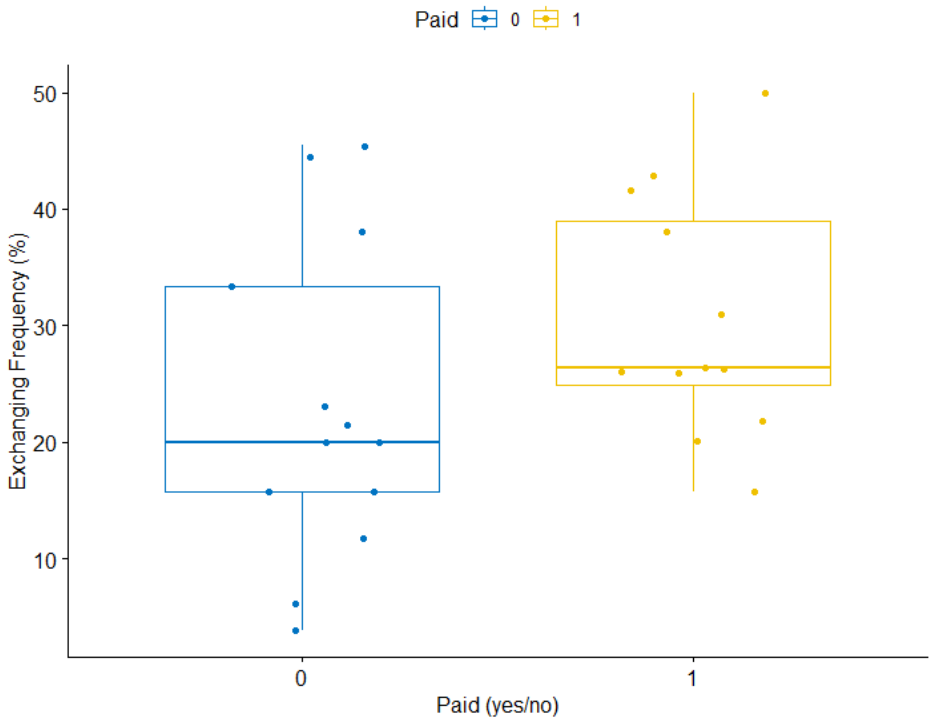
Hypothesis-Testing

To test these relationships more concretely and to examine the first hypothesis H_1, a one-sided two-sample t-test was conducted. The investigated variable was the average E*xchanging* frequency by the negotiator in paid/unpaid negotiations. However, before doing so, the data was tested for violations of the parametric assumptions. First, the assumption of normality was investigated using the Shapiro-Wilk's test. The results for the group "not paid" ($W = .93$, $p = .35$) showed no issues. For the group "paid" the results ($W = .93$, $p = .44$) also show no significant findings. This indicates no issues with skew and kurtosis, meeting the assumption of normality. Next, Levene's test was employed to check the assumption of equal variances. The results ($F(12, 11) = 1.69$, $p = .39$) showed no indication of unequal variances, making the dataset suitable for parametric testing. For that reason, a one-sided independent-sample t-test was

conducted to compare for *Exchanging* frequency in paid and unpaid negotiations. The results showed a difference in *Exchanging* frequency between the groups "paid" (M = .31, SD = .10) and "not paid" (M = .23, SD = .14). This test was not significant, $t(23) = -1.54$, p = .93. A visual representation of the two groups is shown in Figure 3. While there is a relationship in the expected direction, no statistical significance was found.

Figure 4

Boxplot representing the group differences in Exchanging frequency between paid and unpaid negotiations

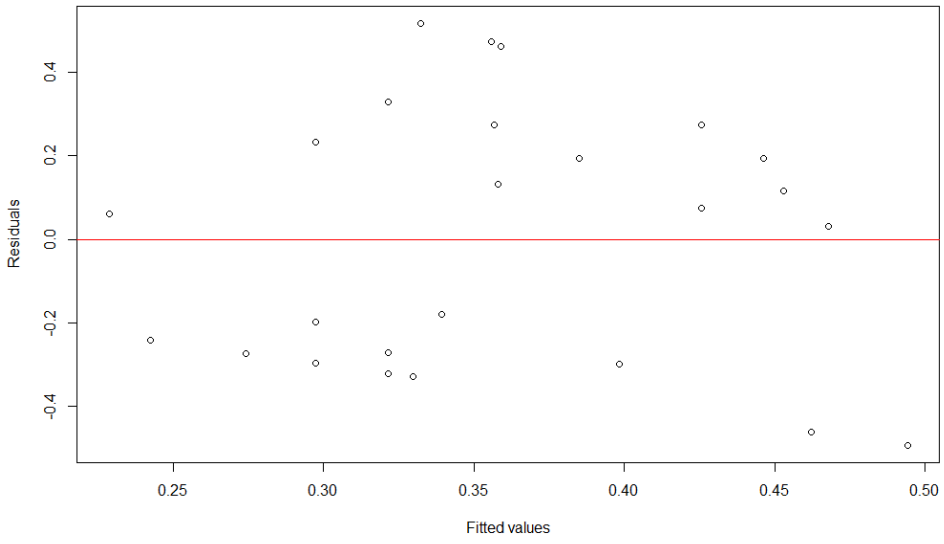


Note. $N_{paid} = 12$, $N_{unpaid} = 13$, *Exchanging* frequency is a ratio variable of total IT usage per transcript

For the second hypothesis, the relationship between the negotiator's usage of *Exchanging* and the magnitude of the offered discount by the TA was investigated. Before conducting a linear regression, the parametric assumptions were explored. An overview of the diagnostic plots can be found in Appendix D. First, the residuals of the model were checked for normality of variance. The Shapiro-Wilk normality test results (W = .93, p = .12) showed no issues with skew and kurtosis. Additionally, a visual inspection of the residuals plotted on the fitted values of the model was conducted to ensure that there was no violation of the assumption of the homogeneity of variance (see Figure 4). No clear patterns were evident upon inspection, thus upholding the homogeneity of variance. Additionally, due to having only one predictor variable, there cannot be any multicollinearity. With no issues regarding the parametric assumptions, the linear regression was performed.

Figure 5

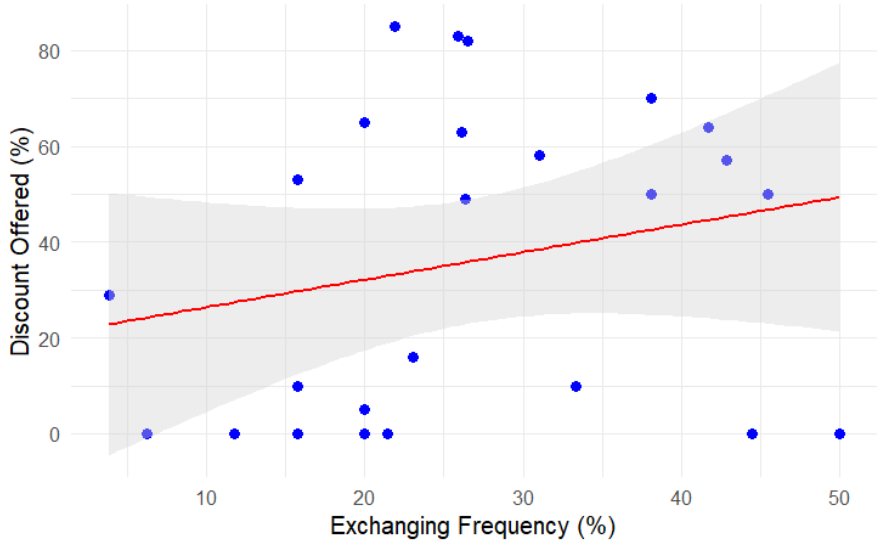
Model residuals plotted on fitted values



Simple linear regression analysis was conducted to evaluate the extent to which the victim's *Exchanging* frequency could predict the offered ransom discount. A significant regression was not found, B = 0.58, t(24) = 1.12, p = .27. The R2 was [.05], indicating that the negotiators *Exchanging* frequency explained approximately 5% of the variance in ransom *Discount offered* (F([1], [23]) = 1.26, p = .27). The regression equation was: *Discount offered* = 20.65 + 0.58 x (negotiator *Exchanging* frequency). That is, for each one percent increase in victims *Exchanging* frequency, the predicted offered ransom discount increased by approximately 0.58%. Confidence intervals indicated that we can be 95% certain that the slope to predict ransom *Discount offered* from negotiator *Exchanging* frequency is between -0.49 and 1.64.

Figure 6

Relationship between Exchanging Frequency and Discount Offered



Note. N = 25, Scatterplot with linear regression equation visualized (red), all values are ratio variables

Discussion

Summary

This paper set out to investigate the effect the IT *Exchanging*, of the Table of Ten, had on the outcome of ransomware negotiations. Specifically, it was assumed that negotiation transcripts that ended with a payment being made would feature a higher frequency of *Exchanging* than those where no agreement was reached. Additionally, it was theorized that *Exchanging* would be related to more positive negotiation outcomes, specifically to a higher ransom *Discount offered* by the TA. The means of the paid/unpaid transcripts were in the expected direction, but no statistically significant effect was found. Additionally, *Exchanging* showed the expected direction of the relationship with *Discount offered*, albeit non-significantly. No statistically significant evidence was found to support the hypotheses.

Hypothesis testing

First, the hypothesis regarding the difference in *Exchanging* usage between paid and unpaid transcripts was examined. There was a non-negligible, yet statistically insignificant, difference in the expected direction. However, the difference between the group means was not sufficient to confirm the hypothesis and its underlying reasoning. It was assumed that these group differences are attributable to the negotiators' intention to pay the ransom as documented by Falco et. al (2019). While that is still a plausible explanation, other factors could play a role. For example, the strongest point-biserial correlation of the dataset is *Discount offered* with “Ransom paid”. This relationship points to the willingness of the TA to offer a discount as a more important predictor. Meurs et. al (2022) also found the offered ransom post-negotiation to be one of the most important predictors of payment. It seems that the TAs willingness to negotiate at all is very relevant to the likelihood of a ransom being paid.

Then, the focus was put on the second hypothesis regarding the effect of the negotiators exchanging behavior on the offered ransom discount. As expected, there was a positive relationship between the usage of *Exchanging* and the offered ransom discount. However, while reciprocity and exchanging behavior seemed beneficial, they did not have the same magnitude of effect as Olekalns and Smith (2000). Olekalns and Smitch (2000) found the exchange of information as one of the most important predictors of joint negotiation gains. A possible alternative reason here could be the moderating effect of culture. Since *Exchanging* is a content-related IT, it is usually more efficient for individuals of low-context cultures (Giebels & Noelander, 2004). However, in ransomware negotiation, we do not have any knowledge of the TA's identity or cultural association. Thus, expectations based on a Western population might not hold true. Culture, along with methodological issues that will be examined below, could be responsible for the lack of statistically significant findings.

Additional Findings

Two other interesting findings are supporting the efficacy of *Exchanging*. First, the conducted correlation study showed that one of the most significant correlations was the negotiators *Exchanging* frequency with the TA's *Exchanging* frequency. This supports previous more general research into negotiation sequences, that showed that ITs are likely to be mirrored by their counterparts (Adair & Brett, 2005). Even in Crisis negotiations specifically, Giebels & Noelander (2004) found a similar pattern of a reciprocal use of *Exchanging*. In this regard, ransomware negotiations seem similar to crisis negotiation scenarios. What has to be recognized, however, is that a correlation does not inform about the specific sequence in which *Exchanging* was used. So, while the results suggest a relationship between the negotiators' usage of exchanging behavior and the TA's usage of the same IT, more research is necessary to obtain

definite proof.

Secondly, the negotiator's usage of *Exchanging* was significantly negatively correlated with the TA's usage of Intimidation. This means that the more *Exchanging* is used, the less is *Intimidation*. This finding is contrary to Giebels & Noelander (2004), who found only *Being equal* to have a positive interaction pattern with a TA in instrumental sieges. They also found no way to elicit or discourage *Intimidation*, whereas this research seems to indicate that *Exchanging* can indeed discourage TA *Intimidation*. Part of that relationship might be explained by the difference in coding style. In this paper's approach, all frequencies are percentages of the total utterances per party. Thus, if *Exchanging* elicits *Exchanging*, the comparative total of Intimidation might be lower than in other research, where each speaking turn is coded instead of each utterance. On the other hand, the relationship is very strong, making it unlikely to be entirely explainable by methodological differences. As a result, it seems that *Exchanging* can discourage the TA from using undesirable ITs such as Intimidation, and engage him in more productive dialogue.

Another notable finding was the high usage of *Being Credible* by both negotiators and TA. Other profiles of Influence strategies show no such numbers of *Being Credible*. The values found in this paper far exceeded both instrumental as well as expressive crisis negotiations (Giebels & Noelander, 2004). This could be related to the role of trust in negotiations. Previous research has outlined the importance of trust in negotiations (Kimmel et al., 1980). In crisis negotiations, the trust factor is already significantly lower than in "traditional" distributive negotiations (Giebels & Noelander, 2004). Add to that an environment where no personal contact is establishable and decryption is not guaranteed, even by payment, the trust level is

likely to be even lower. *Being Credible* is a tactic based on the principle of authority that TAs use to convince the other party that they can be trusted.

Limitations

The first and most obvious limitation is the sample. First, there is a high variance in the dataset. This is likely related to the sample size being only 25 negotiation transcripts. Having a sample of this size magnifies the effect of singular data points. In the dataset, there are some significant outliers regarding ransom amount and *Exchanging* frequency. This dynamic then influences the results at all levels, even though these outliers are not representative of the whole dataset. This is particularly true for the group mean comparison, where the investigated groups consisted of twelve or thirteen transcripts respectively. Another issue that might be directly caused by the sample size is the lack of statistical significance. Having a lower sample size decreases the power of the test and thus the likelihood of finding a significant p-value.

Additionally, the sample population is not all ransomware transcripts, but only those available from Marchive (n.d.). This introduces possible sampling bias, as it is unknown how exactly these transcripts were acquired. Conclusions drawn from this dataset are only valid for transcripts from Marchive (n.d.), or for the TA groups that were assessed. Then, some samples had to be replaced manually in order to uphold a 50/50 split of paid/unpaid negotiations. As some TA groups had no paid transcripts at all, they had to be replaced by paid transcripts from other groups. As a result, some groups hold a more significant share of paid/unpaid transcripts in the sample. The sample representativeness is not as good as it should be.

A different limitation is the amount of unknown variables that could not be controlled for. For example, Meurs et. al (2022) found four contextual predictors for the negotiation outcome of ransomware negotiations (Ransom requested after negotiation, data exfiltration,

targeted ransom note, days of negotiating, and blackmail). No contextual data was available for the sample due to anonymization. Another important factor that has already been mentioned is culture. It is impossible to account for cultural differences as the identity of the TA is unknown. The high variance and poor linear model fit are other indicators for underlying variables that might explain more variance in the negotiation outcome. Many contextual factors influence the negotiation outcome, which could not be analyzed.

Further, the reliability measure for the coding of the material was relatively low. The achieved value was .68, which is above the usual threshold of .65, but nonetheless concerning. These issues occurred due to differences between the AI-generated practice documents and the real negotiation logs. In practice, the kappa was often higher due to the similar structure of many of the transcripts. The real data was less predictable and more varied. This also leads to another issue, the underdeveloped coding ruleset. There were some reoccurring issues in codes, where the opinions of the coders diverged. An example of this was the difference between direct pressure and intimidation, where one coder tended to code a similar phrase as *Direct Pressure*, and the other coded it as *Intimidation*. The reason for this was that the practice logs did not provide sufficient variation to account for all of the different types of utterances present in the real transcripts.

Future Research

For future research, the most important factor would be a more data-rich sample. Currently, it is very difficult to acquire ransomware negotiation logs, as many companies would rather keep them private if a TA was able to infiltrate their systems. However, the lack of contextual information hinders the ability to create a model that can effectively reduce the effects to those attributable to negotiation strategies. For that reason, a databank with ransomware

negotiation logs that have been sufficiently anonymized, yet still consistently contain contextual variables would be essential to enable future, more in-depth, research into ransomware negotiations. Having the ability to account for more information, not only through sample size, could improve the validity of future research.

Another factor that could be interesting is the effect of trust in ransomware negotiation. The dataset showed a very high degree of *Being Credible* usage, an IT intended to increase the perception of reliability and trust. Trust has been shown to be incredibly important to negotiation success. Even for crisis negotiation, trust is crucial, as many crisis negotiation strategies focus on relationship building and increasing the trust the TA can put into said relationship (Vecchi et. al, 2005; Kelln & McMurtry, 2007; Hammer, 2007). Nonetheless, considering the much higher frequency of *Being Credible* and the incentive for the TA to show reliability, trust could be subject to a different dynamic in ransomware negotiations. Future research should examine how the role of trust in ransomware negotiations compares to other kinds of (crisis-)negotiations.

Meaningful Contributions & Practical Recommendations

This research provided a first exploratory view into a field in which very little research exists. It highlighted the importance of exchanging behavior. Despite the lack of significant findings to support the hypotheses, there are indications that it could have a beneficial effect on the negotiation outcome. Additionally, this research showed some clear similarities with research in other contexts like the symmetrical effect of *Exchanging*, but also differences like the frequency of occurrence of being credible. The latter also highlights the possibility of a different role of trust in ransomware negotiations.

While there are some takeaways, a lot of information regarding the effects of *Exchanging* remains unknown. At least, exchanging behavior is related to more *Exchanging* from the TA,

however, it is unclear whose *Exchanging* is being reciprocated. Keeping this in mind, it might be possible that using *Exchanging* is related to a reciprocal exchange of offers & counteroffers and thus beneficial to the negotiation. Additionally, the TA seems to be very interested in demonstrating his credibility. This is an avenue that could inform the choice of tactics used in negotiations. Asking for more proof or voicing one's concerns of the reliability of the TA could stall for more time and allow for more information to be collected. In a field where very little research exists, this paper informs future research and gives tentative estimations on possible similarities and differences between Ransomware and other (crisis-) negotiation scenarios.

References

Adair, W. L., & Brett, J. M. (2005). The Negotiation Dance: time, culture, and behavioral sequences in negotiation. *Organization Science*, *16*(1), 33–51. https://doi.org/10.1287/orsc.1040.0102

Andronio, N., Zanero, S., & Maggi, F. (2015). HelDroid: Dissecting and Detecting Mobile Ransomware. In *Lecture notes in computer science* (pp. 382–404). https://doi.org/10.1007/978-3-319-26362-5_18

Beaman, C., Barkworth, A., Akande, T. D., Hakak, S., & Khan, M. K. (2021). Ransomware: Recent advances, analysis, challenges and future research directions. *Computers & Security*, *111*, 102490. https://doi.org/10.1016/j.cose.2021.102490

Beune, K., Giebels, E., & Taylor, P. J. (2010). Patterns of interaction in police interviews. *Criminal Justice and Behavior*, *37*(8), 904–925. https://doi.org/10.1177/0093854810369623

Brett, J. M. (2000). Culture and negotiation. *International Journal of Psychology*, *35*(2), 97–104. https://doi.org/10.1080/002075900399385

Bureau of International Narcotics and Law Enforcement Affairs. (2023, June 14). *Cybercrime - United States Department of State*. United States Department of State. https://www.state.gov/cybercrime

Chitadze, N. (2023). Basic principles of information and cyber security. In *Advances in human and social aspects of technology book series* (pp. 193–223). https://doi.org/10.4018/978-1-6684-5760-3.ch009

Cialdini, R. B. (2001). *Influence: Science and Practice*. Allyn & Bacon.

Coulthard, M., May, A., & Sousa-Silva, R. (2020). Police Crisis Negotiation: An assessment of Existing models. In *The Routledge Handbook of Forensic Linguistics* (2nd ed.). https://e-space.mmu.ac.uk/625110/3/Archer%20and%20Todd%20ch%20for%20RHFL%20final%20-%20with%20refs%20and%20figures.pdf

Cybersecurity and Infrastructure Security Agency [CISA]. (2023, October 19). *#StopRansomware Guide | CISA*. CISA. https://www.cisa.gov/resources-tools/resources/stopransomware-guide

EU Policy Cycle - EMPACT | Europol. (2023, October 4). Europol. https://www.europol.europa.eu/crime-areas-and-trends/eu-policy-cycle-empact

Europol. (2015, May 7). *Tips & advice to prevent ransomware from infecting your electronic devices*. https://www.europol.europa.eu/operations-services-and-innovation/public-awareness-and-prevention-guides/tips-advice-to-prevent-ransomware-infecting-your-electronic-devices

Falco, G., Noriega, A., & Susskind, L. (2019). Cyber negotiation: a cyber risk management approach to defend urban critical infrastructure from cyberattacks. *JOURNAL OF CYBER POLICY*, *4*(1), 90–116. https://doi.org/10.1080/23738871.2019.1586969

Giebels, E. (2002). Beïnvloeding in gijzelingsonderhandelingen: de tafel van tien. Nederlands Tijdschrift voor de Psychologie, 57, 145 - 154.

Giebels, E., & Noelanders, S. (2004). *Crisis negotiations: A multiparty perspective* [Print]. Universal Press.

Giebels, E., Oostinga, M. S. D., Taylor, P. J., & Curtis, J. L. (2017). The cultural dimension of uncertainty avoidance impacts police–civilian interaction. *Law And Human Behavior*, *41*(1), 93–102. https://doi.org/10.1037/lhb0000227

Gómez-Hernández, J. A., Álvarez-González, L., & García-Teodoro, P. (2018). R-Locker: Thwarting ransomware action through a honeyfile-based approach. *Computers & Security*, *73*, 389–398. https://doi.org/10.1016/j.cose.2017.11.019

Hammer, M. R. (2007). *Saving Lives: The S.A.F.E. Model for Resolving Hostage and Crisis Incidents*. Greenwood Publishing Group.

Holt, T. J. (2020). Subcultural theories of crime. In *Springer eBooks* (pp. 1–14). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-90307-1_19-1

Holt, T. J., Burruss, G. W., & Bossler, A. M. (2010). SOCIAL LEARNING AND CYBER-DEVIANCE: EXAMINING THE IMPORTANCE OF a FULL SOCIAL LEARNING MODEL IN THE VIRTUAL WORLD. *Journal of Crime and Justice*, *33*(2), 31–61. https://doi.org/10.1080/0735648x.2010.9721287

Kelln, B. R. C., & McMurtry, C. M. (2007). STEPS–Structured Tactical Engagement Process. *Journal of Police Crisis Negotiations*, *7*(2), 29–51. https://doi.org/10.1300/j173v07n02_03

Lee, S., Kang, I., & Kim, H. (2023). Understanding cybercrime from a criminal's perspective: Why and how suspects commit cybercrimes? *Technology in Society*, *75*, 102361. https://doi.org/10.1016/j.techsoc.2023.102361

Leukfeldt, R., Stol, W., & Kleemans, E. (2016). Origin, growth and criminal capabilities of cybercriminal networks. An international empirical analysis. *Crime Law and Social Change*, *67*(1), 39–53. https://doi.org/10.1007/s10611-016-9663-1

Leukfeldt, R., & Yar, M. (2016). Applying Routine Activity Theory to Cybercrime: A Theoretical and Empirical analysis. *Deviant Behavior*, *37*(3), 263–280. https://doi.org/10.1080/01639625.2015.1012409

Marchive, V. (n.d.). *Negotiation with ransomware groups*. Retrieved April 24, 2024, from https://ransomware.live/#/negotiations

Meurs, T., Cartwright, E., & Cartwright, A. (2023). Double-Sided Information asymmetry in double extortion ransomware. In *Lecture Notes in Computer Science* (pp. 311–328). https://doi.org/10.1007/978-3-031-50670-3_16

Olekalns, M., & Smith, P. (2000). Understanding optimal outcomes. Human Communication Research, 26(4), 527–557. https://doi.org/10.1111/j.1468-2958.2000.tb00768.x

Ryan, P., Fokker, J., Healy, S., & Amann, A. (2022). Dynamics of targeted ransomware negotiation. *IEEE Access*, *10*, 32836–32844. https://doi.org/10.1109/access.2022.3160748

Sophos. (2023, May). *The State of Ransomware 2023*. SOPHOS. https://www.sophos.com/en-us/content/state-of-ransomware

Vecchi, G. M., Van Hasselt, V. B., & Romano, S. J. (2005). Crisis (hostage) negotiation: current strategies and issues in high-risk conflict resolution. *Aggression and Violent Behavior*, *10*(5), 533–551. https://doi.org/10.1016/j.avb.2004.10.001

Vetschera, R. (2013). Negotiation processes: an integrated perspective. *EURO Journal on Decision Processes*, *1*(1–2), 135–164. https://doi.org/10.1007/s40070-013-0006-5

Von Neumann, J., & Morgenstern, O. (2007). Theory of Games and Economic Behavior (60th anniversary Commemorative Edition). In *Princeton University Press eBooks*. https://doi.org/10.1515/9781400829460

Wade, M. (2021). Digital hostages: Leveraging ransomware attacks in cyberspace. *Business Horizons*, *64*(6), 787–797. https://doi.org/10.1016/j.bushor.2021.07.014

World Economic Forum. (2023, January 11). *Global Risks Report 2023 | World Economic Forum*. https://www.weforum.org/publications/global-risks-report-2023/digest/

Appendix A

Sample Characteristics Overview

Ransomware (Threat Group)	Name	Number of Utterances	Initial Ransom	Negotiated Ransom	Paid	LINK
AKIRA	20230929	58	$300,000	$250,000	0	https://ransomware.live/#/negotiation/akira/20230929.html
AKIRA	20240129	70	$275,000	$140,000	1	https://ransomware.live/#/negotiation/akira/20240129.html
AKIRA	20230616	80	$160,000	$75,000	1	https://ransomware.live/#/negotiation/akira/20230616.html
AVADDON	20210512	35	$1300	N/A	1	https://ransomware.live/#/negotiation/avaddon/20210512.html
BLACKBASTA	20230410	57	$400,000	$150,000	1	https://ransomware.live/#/negotiation/blackbasta/20230410.html
CONTI	20210812	46	$300,650	$150,000	1	https://ransomware.live/#/negotiation/conti/20210812.html
CONTI	20210820	50	$980,000	$350,000	1	https://ransomware.live/#/negotiation/conti/20210820.html
CONTI	20210517	56	$400,000	$200	0	https://ransomware.live/#/negotiation/conti/20210517.html
CONTI	20211205	63	$950,000	$170,000	1	https://ransomware.live/#/negotiation/conti/20211205.html
CONTI	20210611	48	$600,000	$256,000	1	https://ransomware.live/#/negotiation/conti/20210611.html
DARKSIDE	20210413	63	$600,000	$250,000	1	https://ransomware.live/#/negotiation/darkside/20210413.html
HIVE	20211026	46	$3,500,000	$2,500,000	0	https://ransomware.live/#/negotiation/hive/20211026.html
LOCKBIT3.0	wabteccorp.com	39	$25,000,000	N/A	0	https://ransomware.live/#/negotiation/lockbit3.0/wabteccorp_com.html
LOCKBIT3.0	chsf.fr	42	$1,000,000	N/A	0	https://ransomware.live/#/negotiation/lockbit3.0/chsf_fr.html
LOCKBIT3.0	millennia.pro	43	$300,000	N/A	0	https://ransomware.live/#/negotiation/lockbit3.0/millennia_pro.html
LOCKBIT3.0	gocontec.com	52	$4,000,000	$3,600,000	0	https://ransomware.live/#/negotiation/lockbit3.0/gocontec_com.html
LOCKBIT3.0	msim.de	54	$2,000,000	$1,900,000	0	https://ransomware.live/#/negotiation/lockbit3.0/msim_de.html
LOCKBIT3.0	okcu.edu	56	$1,000,000	N/A	0	https://ransomware.live/#/negotiation/lockbit3.0/okcu_edu.html
LOCKBIT3.0	sirva.com	78	$15,000,000	N/A	0	https://ransomware.live/#/negotiation/lockbit3.0/sirva_com.html
REVIL	20210603	63	$2,500,000	$400,000	1	https://ransomware.live/#/negotiation/revil/20210603.html
REVIL	20210622	52	$100,000	$35,000	1	https://ransomware.live/#/negotiation/revil/20210622.html
REVIL	20210609	58	$300,000	$50,000	1	https://ransomware.live/#/negotiation/revil/20210609.html
BLACKMATTER	20210829	44	$5,000,000	$1,5000,000	0	https://ransomware.live/#/negotiation/blackmatter/20210829.html
BLACKMATTER	20210907	77	$15,000,000	$13,500,000	0	https://ransomware.live/#/negotiation/blackmatter/20210907.html
CLOAK	20230802-2	66			0	https://ransomware.live/#/negotiation/cloak/20230802-2.html

Appendix B

Coding Guidelines

Coding guidelines
Tabe of Ten

REFERENCES

Giebels, E. (2002). Beïnvloeding in gijzelingsonderhandelingen: De tafel van tien. Nederlands Tijdschrift voor de Psychologie, 57, 145-154. (first publication)

Other overviews

Giebels, E. & Noelanders, S. (2004). *Crisis negotiations: A multiparty perspective.* Veenendaal: Universal Press. (Available in English, Dutch and German). https://research.utwente.nl/en/publications/crisis-negotiations-a-multiparty-perspective

Giebels, E. & Taylor, P.J. (2010). Communication predictors and social influence in crisis negotiations. In R.G. Rogan & F. J. Lanceley (Eds.) *Contemporary Theory, Research, and Practice of Crisis and Hostage Negotiation, pp 59-77.* Cresskill, New Jersey: Hampton press.

Euwema, M.C. & Giebels, E. (spring, 2024). Conflict management and mediation. Edgar Elgar Publishing. ISBN 978 1 0353 3154 3.

WHAT TO CODE (latest version; adapted from Euwema & Giebels, forthcoming).

STRATEGY	Underlying principle	Examples of behavior
1.Being kind	Sympathy	A. Active listening B. Show empathy C. Kindly offer something
2. Being equal	Identification	A. Use 'We' instead of 'I/You' B. Stress something you have in common (background, family circumstances, hobbies) C. Emphasize mutual goal/dependence/enemy
3. Being credible	Authority	A. Show reliability (do what you say) B. Emphasize your expertise/experience (you know what you are doing)

		C. Show you are transparent
4. Emotional appeal	Self-image (heart)	A. Touch upon feelings/ask for sympathy *(how it affects you/victims)* B. Praise other's behavior C. Boost other's self-respect
5. Intimidation	Deterrence	A. Warnings B. Threats C. Condemn transgression
6.Imposing a restriction	Scarcity	A. Postpone an answer B. Ignore other/not being available C. Offer limited choice (A or B)
7. Direct pressure	Power of fact/repetition	A. Repeat request (planting the seed) B. Share fact C. Give instruction
8. Legitimizing	Legitimacy (external)	A. Reference to formal rules/the law B. Reference to procedures C. Mentioning of moral/social codes
9. Exchanging	Reciprocity	A. Ask for something in return B. Lower your bid C. Exchange proposal
10.Rational persuasion	Consistency (head)	A. Use of arguments B. Provide logic C. Confront with inconsistencies

---

Add code 11 – INFORMATION EXCHANGE
11A: ASK information, e.g. What is your name?
11B: GIVE information: my name is Pim

So, in total you have a max of 32 codes to choose from (but only choose an A-C if it is really evident).

Appendix C

Notes Coding

question: eher intention oder text coden

What is the difference between asking for some time (postponing answer) and asking for more time (asking for something in return)

> Figure out when to take postpone an answer in general

"We need more time" -> Implied threat to end negotiations -> Deterrence (5)

"We will get back to you" -> Postpone an answer 6

Agreements:

If a question references something as a condition for proceeding, then it's 9

If the reason for utterance is to prove credibility it is 3

Consider the reason for utterances more than focusing on a specific part of an utterance

If utterance is based on self-image (in an appeal to logic) then it's 10 (consistency)

> Example: "Given our financial situation, especially post-pandemic, we can only offer a fraction of your demand." -> 10

If a mention of the process is more related to the thought process it's 3c (transparency). When it's about hard coded written process it's 8 (reference to process)

> Examples:
>
> Our stakeholders are involved now. This is more complex than a simple transaction. -> 3
>
> We need to ensure that our actions don't set a precedent or encourage further attacks. (this one is more about social rules and not procedure) -> 8
>
> We're discussing the situation internally. This isn't a straightforward decision. -> 3

For 11b it needs to be an undeniable fact or a promise (clearly defined intention)

When between 11 and an influence tactic, use IT in most cases

Code an utterance as 8b when the proceeding of the negotiations is directly dependent on the referenced procedure

Only code 7 when there's direct pressure

The first time the ransom demands are mentioned by the threat actor it's code 7

Only use 6 if the principle of scarcity applies

Only code as 6c if there is an explicit mention of two options

If victims mention (threaten with) alternative solutions or refusal of payment code as 5

If (after the first occurrence) the time limit (i.e., number of hours) is mentioned unprompted again it is coded as 7 and not 6 / But if the victim asks for an extension and the response is to keep the original time limit it is still 6 / If they extend the deadline it is 9

When the Threat actor accepts a diminished proposal from the victim we code as 9b

"Thank you" at the end of the transcript is 1

Code 1-word utterances like "Hello" or "Ok" as 1 (Active Listening) as long as there is not negative undertone

If an entire utterance is redacted it will not be coded / If an utterance is partly redacted but it is still possible to recognize the intention, it will be coded normally / If a message is partly redacted and it is impossible to tell the intention, it will be coded as 11

Uploads of files for demonstrations of decryption by the victim are coded as 11 (information exchange), while the return of the decoded files is coded as 3 (being credible)

Appendix D