Counter Strategies in Ransomware Negotiations:

Victim Responses to the First Ransom Demand

Melisa Imsak (s2872811)

University of Twente

Bachelor Thesis in Psychology of Conflict, Risk, and Safety

Faculty BMS

First supervisor: Michalis Georgiou, Msc.

Second supervisor: prof.dr. Ellen Giebels

June 24, 2025

Abstract

Ransomware negotiations are often framed in technical or legal terms, but behind these attacks lie urgent and emotionally charged human decisions. This thesis examines how victims respond to the first ransom demand, focusing not just on what they say, but how they navigate the uncertainty, pressure, and power imbalance of the situation. Analyzing 134 negotiation transcripts, the study identifies eleven distinct counter strategies and categorizes them by communication style. The primary aim was to explore whether these responses exhibit structured patterns, and whether communication style relates to ransom payment outcome. While no statistically significant link between communication style and payment was found, the findings reveal repeated behavioral patterns that suggest interactional structure rather than randomness. Victim replies often reflect attempts to delay, challenge, or probe the attacker, indicating that their role in negotiations may be more active and patterned than previously assumed. These insights contribute to a deeper understanding of crisis communication and offer a human-centered perspective on one of today’s most urgent cyber threats.

Introduction

Ransomware attacks have become one of the most disruptive forms of cyber extortion in recent years. In these incidents, attackers encrypt files or threaten to leak sensitive data unless the victim agrees to a ransom payment. While the technical, legal, and economic aspects of ransomware have received increasing attention in the literature (e.g., legal frameworks for reporting obligations and sanctions, or economic tools like cryptocurrency tracking and ransom insurance; see Anderson et al., 2021; Connelly & Wall, 2019; Wilner et al., 2019), the communication processes that unfold during these attacks, particularly from the victim's side, remain underexplored. Yet these interactions are far from marginal. Victims are confronted with psychological pressure, time constraints, and uncertainty, all while having to respond to a first ransom demand under asymmetric power conditions.

A feature of ransomware negotiations that puts victims further at a disadvantage is the fact that the negotiation typically begins with a concrete monetary demand from the attackers. In negotiation literature, this first offer is known to function as a critical anchor that influences subsequent decisions (Galinsky & Mussweiler, 2001). In ransomware negotiations, this offer takes the form of a fixed monetary demand, usually accompanied by threats and deadlines. The victim's initial set of responses to this demand – whether questioning, stalling, negotiating, or pleading – likely influences the trajectory of the entire exchange. These early reactions are especially valuable to study because they are made under the greatest emotional and informational pressure. Rather than viewing them as isolated replies, this study considers them as counter strategies: a cluster of messages sent by the victim in response to the attacker's first demand.

Existing frameworks in the field of negotiation, such as the Table of Ten influence strategies (Giebels & Taylor, 2010), have provided useful ways of categorizing intentional communication designed to persuade or shift the other party. However, the present study does not examine influence strategies in the traditional sense, commonly understood as deliberate efforts to persuade or shift the other party's position (Giebels & Taylor, 2009). Instead, it focuses on counter strategies used by victims in response to the first ransom demand. These strategies are examined not in terms of their persuasive intent, but as observable communicative patterns that emerge under pressure. While some responses may carry implicit intent or strategic features, the study does not assume that they are consciously designed to influence the attacker. Rather, the goal is to identify and categorize the distinct ways in which victims react in the early stages of ransomware negotiations, including both reactive and more proactive forms of engagement.

The first aim of this study was exploratory and focused on developing a typology of victim counter strategies in ransomware negotiations. The goal was to identify and categorize distinguishable types of responses that victims tend to use when confronted with an initial ransom demand. This approach was informed by prior research in crisis and hostage negotiation, which suggests that individuals under high pressure often rely on structured communicative responses – such as delay tactics, appeals, or clarification-seeking – to manage uncertainty and emotional strain (Giebels & Taylor, 2009; Rogan & Hammer, 1994). Applying an inductive, qualitative approach to real-world negotiation transcripts, the study aimed to capture the range of counter strategies used by victims and to organize them into a coherent typology based on their communicative function.

As a second exploratory aim, the study sought to structure the developed typology by assigning strategies to one of two analytically useful categories: direct or indirect. This distinction is grounded in negotiation and communication theory, where direct communication typically involves assertive, explicit, and task-oriented behavior aimed at influencing the outcome (Olekalns & Smith, 2000; Giebels & Noelanders, 2004). It often includes clear proposals, firm positions, or demands. In contrast, indirect communication reflects a more relational or face-saving orientation and is characterized by politeness strategies, emotional expression, or ambiguity (Brown & Levinson, 1987; Mannix

& Brett, 2002). Indirect messages may serve to reduce tension, avoid escalation, or preserve the social relationship between parties (Ting-Toomey, 1999). Establishing this distinction provided a theoretical foundation for organizing the strategies and enabled further analysis of how broader communication styles may relate to negotiation outcomes.

Building on the distinction between direct and indirect communication styles, this study examined whether the dominant communication style within each negotiation was associated with its outcome – specifically, whether the case resulted in payment (*a deal*) or non-payment (*no deal*). Each log was categorized based on the overall style that appeared most frequently (direct versus indirect), and this was compared to the final outcome of the negotiation. This difference is expected to be important, because research suggests that direct communication, defined by clarity, assertiveness, and goal orientation, can reduce ambiguity and support progress toward agreement in negotiation settings (Olekalns & Smith, 2000; Maddux & Kim, 2020). Indirect communication, while often used to preserve relationships or manage social dynamics (Brown & Levinson, 1987), may appear less decisive or action-oriented in high-stakes, time-sensitive interactions. Related findings in crisis negotiation emphasize the importance of tactical clarity and momentum under pressure (Giebels & Taylor, 2009). Based on these insights, it was hypothesized that negotiations dominated by direct communication would be more likely to result in payment, whereas negotiations dominated by indirect communication would be more likely to result in non-payment.

Method

Design

This study followed a mixed-methods exploratory design, combining qualitative content analysis with descriptive and inferential statistics. The primary aim was to investigate the communicative patterns through which victims respond to the initial ransom demand in ransomware negotiations. Rather than applying pre-existing coding frameworks, an inductive approach was used to develop a nuanced codebook based on observable behaviors.

The qualitative phase involved identifying and assigning behavioral codes to each victim-side message written in response to the attacker's first ransom offer. Importantly, the unit of analysis was not a single message but the initial response cluster, defined as the full sequence of consecutive victim messages that followed the attacker's first demand and preceded the attacker's next reply. This ensured that the analysis captured not just an isolated reaction but the broader early communicative stance.

Once coded, these responses were categorized into three broader communication styles – Direct, Indirect, or Mixed – based on the dominant tone and function of the strategies used. A log was classified as *Direct* or *Indirect* if one of those categories represented at least 60% of the applied codes. If neither category reached that threshold, the log was labeled *Mixed*.

Materials

The dataset consisted of ransomware negotiation transcripts collected from a publicly available GitHub (Casualtek, n.d.) repository that archives leaked conversations between cybercriminal groups and victim organizations. The transcripts included attacker demands, victim responses, bargaining attempts, and final outcomes.

Transcripts were excluded if they:

- Contained no victim-side replies.
- Consisted of non-negotiation exchanges (e.g., only payment confirmation).
- Were in non-English languages that could not be reliably translated.

This selection process resulted in a final sample of 134 coded logs.

Procedure

This study was conducted in collaboration with another bachelor student. The analytical process followed three distinct stages:

Codebook Development Phase

A set of three diverse logs was individually coded to identify recurring behavioral patterns. This phase, lasting approximately 7–8 hours, involved iterative discussion and refinement, ultimately leading to a shared codebook with 11 distinct victim-side counter strategies. Special attention was paid to conceptual clarity, mutual exclusivity, and functional definitions.

Pilot Phase

Next, 24 logs were independently coded by both researchers using the developed codebook. Disagreements were discussed and resolved collaboratively, leading to minor refinements. Cohen's Kappa was calculated based on agreement/disagreement rates for the presence or absence of codes per log, yielding a value of 0.74, which reflects substantial interrater reliability (Landis & Koch, 1977).

Main Coding Phase

The remaining logs were then again independently coded. All 134 logs were then reviewed together to resolve disagreements and ensure consistent application of codes. Messages were coded based on their primary communicative function. In cases of overlap, the dominant tone or goal was used to determine the most appropriate code.

The full codebook is available in Appendix A. Table 1 shows a summary of the applied counter strategies.

Table 1

Victim Counter Strategies

Code	Description
Aggressive Response	Hostile or confrontational message
Assertive Pushback	Firm yet non-hostile rejection
Counteroffer from the Victim	Alternative payment or proposal
Demand of Quality	Request for proof or guarantees

Discount Request by Victim	Appeal to lower the ransom
Emotional Plea	Emotion-based appeal or moral claim
Financial Constrain Appeal	Citing inability to pay
Information Seeking	Questions to gain clarity
Readiness to Negotiate	Expression of willingness to talk
Strategic Delay	Attempts to stall or postpone
Strategic Politeness	Use of respectful or appeasing tone

Note. These codes describe observable communication patterns and tone, without presuming the presence of deliberate influence intent or emotional states.

An additional code, Initial Offer, was applied to label the attacker's first ransom demand. This code was not part of the victim strategy set but was used as a fixed anchor to define the response window.

Payment Status

Each transcript was also coded for payment outcome. If the log contained explicit confirmation of payment, such as file transfers, completion messages, or statements acknowledging resolution, it was marked as *paid*. If the exchange ended in rejection, silence, or indications that no agreement was reached, it was marked as *unpaid*. Transcripts with inconclusive outcomes were excluded from this part of the analysis. In total, 48 logs were coded as paid and 86 as unpaid.

Data Analysis

All transcripts were coded manually in ATLAS.ti (version 25) using the shared codebook. After export, quantitative analysis was conducted in RStudio (version 4.3.0). Absolute and relative frequencies were calculated for each counter strategy and communication style. A Chi-square test of independence was used to assess the association between communication style (Direct / Indirect / Mixed) and payment outcome (Paid vs. Unpaid). Statistical significance was assessed at the standard threshold of $p < .05$.

Results

Identified Counter Strategies and Their Frequencies

To explore the types of counter strategies used by victims in ransomware negotiations, a total of 1,535 coded victim responses were analyzed across 134 logs. The relative frequencies, rounded to one decimal, of each identified strategy are presented in Table 2.

Table 2

Relative Frequencies of Counter Strategy Codes

Code	Frequency	Relative Frequency
Assertive Pushback	294	19.2%
Counteroffer from Victim	206	13.4%
Information Seeking	223	14.5%
Strategic Politeness	205	13.4%
Emotional Plea	138	9.0%
Financial Constraint Appeal	113	7.4%
Demand of Quality	120	7.8%
Readiness to Negotiate	72	4.7%
Strategic Delay	60	3.9%
Discount Request by Victim	55	3.6%
Aggressive Response	49	3.2%

Note. Frequencies represent the total number of coded victim messages across all logs (N = 1,535). Relative frequencies indicate the percentage of total coded responses attributed to each strategy.

As shown in Table 2, the most frequently observed counter strategy was *Assertive Pushback* (19.2%), followed by *Information Seeking* (14.5%) and *Counteroffer from the Victim* (13.4%). These codes appeared more often than others, indicating that firm, inquisitive, or solution-oriented responses were commonly used in early victim reactions. Indirect strategies such as *Strategic Politeness* (13.4%) were also frequently observed. In contrast, codes like *Aggressive Response* (3.2%) and *Discount Request by Victim* (3.6%) occurred less frequently, representing a smaller portion of the response pool.

Communication Style Patterns

Each of the eleven counter strategy codes was grouped into one of two overarching communication styles – Direct or Indirect – based on their tone and function. The classification was designed to ensure mutual exclusivity and prevent overlap. The groupings are presented in Table 3.

Table 3

Grouping of Victim Counter Strategies by Communication Style

Direct Counter Strategies	Indirect Counter Strategies
Counteroffer	Readiness to Negotiate
Financial Constraint Appeal	Strategic Politeness
Discount Request	Information Seeking
Assertive Pushback	Strategic Delay
Demand of Quality	Emotional Plea
	Strategic Delay

	Aggressive Response

The following sections describe each of the strategy types, supported by representative quotes and short interpretations of how they were used in the early stage of the negotiation.

Direct Counter Strategies

Direct responses were characterized by a clear and prompt engagement with the attacker following the initial ransom demand. These messages typically reflected a task-oriented and explicit communication style, including efforts to clarify terms, discuss price-related issues, or express firm positions. While the emotional tone of these messages varied, from calm and professional to frustrated or defiant, their defining feature was the victim's overt and structured attempt to address the situation through concrete statements or proposals.

One of the most common direct strategies involved making a "Counteroffer". Victims often proposed an amount significantly lower than the original ransom. These messages were typically brief and to the point, reflecting an attempt to shift the negotiation onto more manageable terms.

"We can pay 750,000 USD."

"We will at position to pay you around 200k$."

"We are currently working on an additional source that could potentially give us around $256,000."

The counteroffer was sometimes supported with context or justification. In many cases, victims emphasized that their organization was unable to pay more, especially when they represented non-profit, educational, or public service institutions. These "Financial Constraint Appeals" frequently appeared alongside counteroffers or polite requests.

"We are already in loss… we are an educational cultural exchange program… since COVID we have not been able to send applicants out of the US."

"We are a charity run by [redacted]… we get money from public funds. We are not a

business."

"Because we don't have $100,000."

A related but slightly different tactic was to ask directly for a discount, without suggesting a specific number. These "Discount requests" were often polite and brief but clearly signaled the victim's financial limitations.

"What can we do about discount of the price? 500 is beyond our capacity to pay."

"At least provide some discount please."

"Brother, please give me a discount."

Beyond negotiation, some victims engaged in what could be called "Assertive Pushback". These responses stood out for their confident tone and clear resistance to intimidation. Rather than refusing to cooperate, the victim often framed themselves as rational and principled, sometimes even addressing the attacker as a "partner."

"Please refrain from hostile language. I will appreciate it. I was instructed to further elaborate with you a couple of points, obviously after I did my own research about you."

"Your 'client', my colleagues, are not a company… We are a charity fund… and our resources are all public."

"We are not joking and know that you are a serious organization… The global pandemic affected our ability to operate for an entire year causing us to lose most of our business."

Another frequently observed strategy was the emotional plea, which tended to appear when victims represented institutions responsible for people's wellbeing, such as hospitals or schools. These messages focused less on negotiation and more on the human impact of the attack.

"You are attacking a hospital, and many patients may die."

"Please stop this madness, we are just trying to survive."

"We are doing our best to save our jobs, our families, our company."

Finally, some victims expressed skepticism or caution regarding the attacker's promises. These messages questioned the reliability of the decryption tool or the attacker's ability to restore systems. While still direct, they tended to be framed as due diligence rather than confrontation which lead to the development of the code "Demand of Quality".

"We need to make sure the decryption tool will not damage our system."

"We cannot risk destroying critical files. Can you guarantee full recovery?"

"If you say you can decrypt everything, show us how the software works exactly."

Together, these direct responses suggest that many victims attempted to take an active role in the negotiation, even under pressure. Whether through offers, resistance, or appeals, their responses reflected an effort to regain control, protect their interests, or reduce the severity of the threat.

Indirect Counter Strategies

Not all victims responded with open or explicit engagement. Many initial responses were more cautious, vague, or relational in tone. These indirect strategies often involved polite phrasing, emotional appeals, or ambiguity, and tended to delay commitment or shift focus away from immediate demands. Although less assertive than direct responses, they still reflected structured ways of managing the situation without directly addressing the attacker's terms.

A common strategy here was to express a readiness to negotiate, without giving away a specific position. These messages helped victims appear cooperative while postponing the more difficult parts of the conversation.

"We would like to complete this unfortunate event as swift and clean as possible. Kindly let me know how should I call you and when should I expect your reply."

"What will be the discount if we pay quickly to you today?"

"I am ready to do payment, at least provide some discount please."

In other cases, victims used “Strategic Politeness” to manage the tone of the conversation. Messages were often overly formal or deferential, especially in the early stages of contact. While these messages may not have moved the negotiation forward, they seemed to help reduce tension and keep the attacker engaged.

“Dear Sir, thank you for your time and patience. We are working hard to find a solution.”

“We appreciate your support so far and hope to resolve this respectfully.”

“I hope you understand our difficult situation. Please consider this request kindly.”

Some victims used the opportunity to ask questions rather than make decisions. These “Information-Seeking” responses aimed to clarify what had happened, what was at stake, or what options existed. They also gave the victim time to prepare internally or verify the situation.

“How can we know this isn’t a scam?”

“What exactly happens if we don’t pay?”

“Can you explain which systems were affected?”

“Strategic Delay” was another clear theme. Victims often referred to internal decision-making processes, legal reviews, or the need to consult third parties. These responses were structured to justify postponement while avoiding confrontation.

“Our management board is still in discussion.”

“We are waiting for a response from the insurance company.”

“The legal team hasn’t approved anything yet, so we cannot proceed.”

Some victims used emotionally expressive language when responding to the ransom demand. These “Emotional Pleas” often highlighted hardship, urgency, or distress, without putting forward a specific proposal or objection. The messages tended to focus on the personal or organizational impact of the attack, using emotionally charged wording rather than task-oriented or assertive language.

“Please, we are a small business. We don’t have this kind of money.”

“We are desperate. Our employees’ data is at risk. We beg you to reconsider.”

“This is destroying us. Please lower the amount.”

A smaller number of victims responded with aggression or frustration. These “Aggressive Responses” were emotionally charged and often included sarcasm or outright rejection. While less common, they revealed the emotional strain of the situation and, in some cases, reflected a refusal to acknowledge the attacker’s authority. Their rarity however may reflect the risks of escalation or breakdown in communication under high-pressure conditions.

“You are disgusting, attacking innocent people for money.”

“We are not going to pay criminals. Period.”

“Do you think we’re stupid enough to believe you?”

Overall, indirect strategies allowed victims to remain present in the negotiation without overcommitting. These responses often reflected uncertainty, internal constraints, or a desire to delay decisions without closing off options entirely.

Finally, to understand the dominant communication style in each negotiation, logs were classified based on the relative proportion of direct and indirect strategies. A style was considered dominant if either direct or indirect codes made up at least 60% of a log’s total coded responses. Logs without a dominant style were labeled Mixed. Table 4 presents the distribution of communication styles across all 134 logs.

Table 4

Strategy Types by Communication Style

Strategy Type	Frequency	Percentage
Direct	78	58.2%

Indirect	15	11.2%
Mixed	41	30.6%

The majority of logs (58.2%) were classified as Direct, suggesting that most victims engaged with overt, task-oriented strategies in response to the first ransom demand. Indirect communication styles appeared less frequently (11.2%), while 30.6% of logs were classified as Mixed.

Communication Style and Payment Outcome

To explore whether communication style was associated with payment outcome, the logs were cross tabulated by style and payment status. Table 5 presents this distribution.

Table 5

Payment Outcome by Strategy Type

Strategy Type	Paid	Not Paid	Total
Direct	28	50	78
Indirect	9	6	15
Mixed	11	30	41

Table 5 displays the distribution of payment outcomes by communication style. Among the 78 Direct cases, 28 resulted in payment and 50 did not. Of the 15 Indirect cases, 9 were paid and 6 were unpaid. In the Mixed group, 11 cases were paid while 30 were not.

A Chi-square test of independence was conducted to assess the relationship between communication style (Direct, Indirect, Mixed) and payment status. The result was not statistically significant, $\chi^2(2, N = 134) = 5.26$, $p = .072$, indicating that no reliable association could be established between these variables in this sample.

To further explore the relationship while reducing potential ambiguity from the Mixed category, a follow-up analysis was conducted comparing only Direct and Indirect cases. This test also produced a non-significant result, $\chi^2(1, N = 93) = 1.87$, *p* = .17, suggesting no meaningful association between direct or indirect dominant communication and payment outcome.

Discussion

This study set out with the exploratory goal of developing a typology of counter strategies that victims employ when confronted with the first ransom demand in ransomware negotiations. The analysis of 134 negotiation transcripts revealed eleven distinct strategy types, ranging from clear offers and firm objections to emotionally charged or cautiously phrased replies. The frequency analysis showed that some strategies were used more prominently than others, with Assertive Pushback, Information Seeking, and Counteroffer from the Victim being the most commonly observed.

These findings offer valuable insights into the behavioral repertoire of victims during a critical moment in the negotiation process. Although ransomware victims operate under considerable pressure and uncertainty, their responses do not appear to be erratic or entirely passive. Instead, the evidence suggests that victims often engage in structured forms of communication that reflect both reactive and proactive tendencies. The observed counter strategies – many of which have analogs in crisis and hostage negotiation research (Giebels & Taylor, 2009; Rogan & Hammer, 1994) – suggest that even in these highly asymmetric encounters, victims display varying degrees of agency, attempting to reduce perceived threat, negotiate better terms, or simply buy time to gather information and resources.

The resulting typology contributes to the growing literature on negotiation behavior in cybercrime contexts by offering an empirically grounded vocabulary for classifying early-stage victim responses. These classifications can also inform future efforts to build automated systems for detecting, categorizing, or responding to ransomware messages.

As a second exploratory aim, this study investigated whether the eleven counter strategy types could be meaningfully grouped into two overarching communication styles: direct and indirect. This categorization drew from established theoretical distinctions in negotiation and intercultural communication literature (Olekalns & Smith, 2000; Mannix & Brett , 2002; Brown & Levinson, 1987). Direct strategies involved explicit proposals, firm positioning, and task-focused communication, while indirect strategies were characterized by emotional appeals, polite framing, or ambiguity.

The grouping process revealed that strategies could be classified without overlap, and the relative frequencies within each category suggested that direct strategies were generally more common than indirect ones. This distinction was not merely conceptual but allowed for a meaningful aggregation of communicative behavior across the logs, facilitating further analysis at the style level. It also helped clarify how victims navigate the initial pressure of ransomware demands: some through confrontation or problem-solving, others through caution or deflection.

While the grouping appears to hold conceptual clarity and descriptive value, it is important to acknowledge that some strategy types may still operate across a spectrum, and the interpretation of tone may not always be straightforward. Future studies could explore whether hybrid strategies or shifting styles within a negotiation occur more frequently than currently observed. Nonetheless, the direct/indirect distinction offers a useful lens for examining the broader tone and orientation of victim responses.

The third part of the study addressed the hypothesis that dominant communication style (direct vs. indirect) would be associated with payment outcome. Specifically, it was hypothesized that negotiations dominated by direct communication, characterized by clarity, assertiveness, and solution orientation, would be more likely to result in payment, while those dominated by indirect communication, characterized by politeness, ambiguity, or emotional appeals, would more likely end in non-payment.

This hypothesis was grounded in research highlighting the benefits of assertive communication under high-stakes negotiation conditions. For instance, Olekalns and Smith (2000) suggest that assertive strategies tend to enhance clarity and goal alignment, while Maddux and Kim (2020) point to the importance of tactical momentum in crisis negotiations. Similarly, Giebels and Taylor (2009) emphasize the role of structured, forward-driving communication in shaping outcomes in adversarial contexts.

However, the current results did not provide support for this hypothesis. A Chi-square test assessing the relationship between communication style and payment outcome revealed no statistically significant association, both when including the Mixed category and when limiting the analysis to clearly Direct and Indirect logs. While there was a numerical trend in the expected direction – more payments occurred in Direct cases – this pattern was not strong enough to reach significance.

Several interpretations are possible. One is that while communication style may play a role in shaping the tone and structure of a negotiation, other factors, such as organizational resources, attacker behavior, or external time constraints, may have a more decisive influence on outcome. Alternatively, it is possible that communication style interacts with other variables not captured in the current dataset, such as message timing, message content quality, or prior technical actions taken by the victim.

These findings align with prior research showing that negotiation outcomes are complex and often influenced by multi-level dynamics (Giebels & Taylor, 2010). The lack of significant results does not necessarily invalidate the theoretical reasoning behind the hypothesis, but it suggests that the relationship may be more conditional or context-dependent than initially assumed.

Theoretical and Practical Implications

This study contributes to the growing field of cyber extortion research by highlighting the communicative dimension of ransomware incidents. Most existing work focuses on technical vulnerabilities, economic incentives, or legal frameworks (e.g., Anderson et al., 2021; Holt & Bossler, 2015). By contrast, this research emphasizes the interactional behavior of victims, showing that even under severe pressure, victims employ patterned strategies that reflect tone management, delay, resistance, or tentative engagement.

Theoretically, the findings connect the field of cyber extortion with established research on crisis and hostage negotiations. In traditional hostage scenarios, maintaining communication, managing tone, and using delay or rapport-building strategies are considered critical for de-escalation and control (Taylor, 2002; Ireland & Vecchi, 2009). This study shows that similar communicative patterns can emerge in ransomware cases, where victims, despite not being physically held hostage, face high pressure, uncertainty, and limited options. The findings suggest that digital extortion may activate comparable behavioral responses, revealing a continuity between physical and virtual crisis communication dynamics.

Frameworks like the Table of Ten (Giebels & Taylor, 2010) focus on deliberate influence tactics, but the current study demonstrates that functionally distinct communication patterns may still emerge even when no influence is explicitly intended. These patterns may reflect familiar behavioral templates, such as politeness, assertion, or clarification that individuals fall back on when navigating risk and uncertainty (Lewicki et al., 2015). This aligns with findings from Meurs (2024), who argues that message framing and communicative stance during digital extortion play a key role in shaping perceived negotiation risk and outcome trajectories. Similarly, Ryan et al. (2021) found that subtle variations in early message tone can influence attacker response patterns, particularly in high-pressure or time-constrained scenarios.

Moreover, the findings echo earlier work by Schafer et al. (2021) on the importance of maintaining dialogue and psychological stability in crisis negotiations. Although ransomware victims are not physically detained, the institutional urgency and informational asymmetry they face can create a comparable sense of constraint commonly seen in crisis bargaining scenarios (Donohue et al., 1991). In this light, victim behavior is not merely reactive but may be guided by underlying principles of control management and escalation avoidance.

From a practical standpoint, the study offers useful insights for cybersecurity professionals, crisis managers, and organizational response teams. Recognizing the range of typical early-stage victim strategies, such as asking for proof, proposing discounts, or delaying for legal consultation can support better preparedness and communication planning.

These strategies could be incorporated into training simulations or decision-making protocols that help staff maintain clarity under duress. Furthermore, understanding these early responses may help incident handlers recognize the negotiation stage, assess institutional posture, or anticipate escalation risk more effectively. While this research refrains from assumptions about victim power or intent, it shows that their responses often follow structured patterns that influence the negotiation dynamic.

Limitations and Future Research

While this study contributes to a more detailed understanding of victim behavior in ransomware negotiations, several limitations must be acknowledged. First, the data source with leaked negotiation logs limits the generalizability of the findings. These logs may not represent all types of ransomware cases, particularly those involving highly confidential or undisclosed incidents. Furthermore, the context and stakes of each negotiation may vary considerably. Factors such as the victim organization’s sector, size, risk exposure, or available resources could influence both the perceived urgency of the situation and the strategic responses used. For instance, well-resourced organizations may be more likely to resist pressure or propose structured counteroffers, while smaller or more vulnerable entities might rely on emotional appeals or delay tactics. However, these contextual factors were not systematically examined in this study, which limits the ability to generalize findings across different organizational settings.

Second, the categorization of counter strategies into direct and indirect styles was based on a 60% classification threshold, which, although conceptually justified, remains a somewhat arbitrary cut-off. This decision may have affected how some mixed-style negotiations were grouped, and future research might explore alternative classification models or continuous scoring approaches to assess communication style more dynamically.

Another limitation concerns the exploratory nature of the first two research goals and the accompanying typology. While the codebook was developed through rigorous procedures with substantial interrater reliability, it remains interpretive and inductively constructed. Future work could test and refine this typology using additional datasets or experimental simulations, ideally integrating expert feedback or triangulating findings across different negotiation types.

Moreover, approximately 30% of the negotiation logs were classified as Mixed, indicating that neither direct nor indirect communication strategies were dominant. This proportion reflects variation in communication patterns, where multiple counter strategies were used without one prevailing. These mixed styles may reflect strategic flexibility in response to shifting situational demands. Future research could examine the structure and potential function of these mixed patterns in more detail, drawing on existing work on adaptability in negotiation contexts (Čehajić & Giebels, 2023).

The study's hypothesis – that direct communication would be associated with payment – was not supported by the statistical analysis. While this does not undermine the broader insights of the study, it suggests that the relationship between communication style and negotiation outcome may be more complex than anticipated. The hypothesis was grounded in prior work on negotiation clarity and momentum (Maddux & Kim, 2020; Giebels & Taylor, 2009), but such findings may not fully translate to the adversarial and morally charged context of ransomware. This mismatch should encourage more refined theorizing about how power asymmetries, threat legitimacy, and moral framing influence the role of communication in such settings.

Finally, the current study did not examine the attacker's behavior or reactions, which are likely to shape how victim strategies unfold and evolve. A more interactional approach could enrich our understanding of strategy sequencing, mutual influence, and escalation patterns – all of which remain unexplored in this single-sided analysis.

Conclusion

This study examined how victims respond to the first ransom demand in ransomware negotiations, using a mixed-methods approach to explore both the nature and implications of these early-stage counter strategies. Eleven distinct response types were identified and grouped into direct and indirect communication styles, offering a new framework for understanding victim behavior under pressure. While many victims engaged proactively, others responded more cautiously, often using delay, emotional appeals, or indirect questioning. However, the study did not find a significant link between communication style and payment outcome, challenging assumptions about the advantage of clarity or assertiveness in this context.

Taken together, the findings highlight the complexity of victim decision-making and the need to account for both psychological strain and strategic ambiguity in high-stakes cyber extortion. The typology developed here not only contributes to the theoretical literature on negotiation and crisis communication but also lays groundwork for future empirical work and practical intervention. As ransomware attacks continue to evolve, so must our understanding of the human choices embedded within them.

References

Anderson, R., Barton, C., Böhme, R., Clayton, R., van Eeten, M. J. G., Levi, M., Moore, T., & Savage, S. (2021). Measuring the cost of cybercrime. *Journal of Cybersecurity, 6*(1), 1–32. https://doi.org/10.1007/978-3-642-39498-0_12

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Casualtek. (n.d.). *Ransomchats* [Data repository]. GitHub. https://github.com/Casualtek/Ransomchats

Čehajić, A., Brockerhoff, A., & Giebels, E. (2023). Adaptiveness in conflict: The role of strategic flexibility in negotiation. *International Journal of Conflict Management*, 34(2), 193–216. https://doi.org/10.1108/IJCMA-02-2023-0028

Connelly, L., & Wall, D. S. (2019). The rise of crypto-ransomware in a changing cybercrime landscape: Taxonomising countermeasures. *Computers & Security, 87*, 101568. https://doi.org/10.1016/j.cose.2019.101568

Donohue, W. A., Kaufmann, G., Smith, R., & Ramesh, C. (1991). Crisis bargaining: A framework for understanding intense conflict. International Journal of Group Tensions, 21(2), 133-154.

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology, 81*(4), 657–669. https://doi.org/10.1037/0022-3514.81.4.657

Giebels, E., & Noelanders, S. (2004). Crisis negotiations: A multiparty perspective. *Vlaamse Psycholoog*, 60(4), 224–233.

Giebels, E., & Taylor, P. J. (2009). *Interaction patterns in crisis negotiations: Persuasive arguments and cultural differences.* Journal of Applied Psychology, 94(1), 5–19. https://doi.org/10.1037/a0012953

Giebels, E., & Taylor, P. J. (2010). *The Table of Ten: A tool for understanding and resolving negotiation deadlocks*. Centre for Critical Incident Research, University of Liverpool.

Holt, T. J., & Bossler, A. M. (2015). *Cybercrime in progress: Theory and prevention of technology-enabled offenses* (2nd ed.).

Ireland, C. A., & Vecchi, G. M. (2009). The Behavioral Influence Stairway Model (BISM): a framework for managing terrorist crisis situations? Behavioral Sciences of Terrorism and Political Aggression, 1(3), 203–218. https://doi.org/10.1080/19434470903017722

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lewicki, R. J., Barry, B., & Saunders, D. M. (2015). *Negotiation* (7th ed.). McGraw-Hill Education.

Mannix, Elizabeth & Brett, Jeanne. (2002). Negotiating Globally: How to Negotiate Deals, Resolve Disputes, and Make Decisions across Cultural Boundaries. Industrial and Labor Relations Review. 56. 193. https://doi.org/10.2307/3270665

Meurs, M. (2024). Negotiating under the shadow of encryption: Psychological risk in ransomware decisions.

Olekalns, M., & Smith, P. L. (2000). Understanding optimal outcomes: The role of strategy sequences in competitive negotiations. *Human Communication Research, 26*(4), 527–557. https://doi.org/10.1111/j.1468-2958.2000.tb00768.x

Rogan, R. G., & Hammer, M. R. (1994). Crisis negotiations: A preliminary investigation of facework in naturalistic conflict discourse. *Journal of Applied Communication Research, 22*(3), 216–231. https://doi.org/10.1080/00909889409365399

Ryan, P., Fokker, J., Healy, S., & Amann, A. (2021). Dynamics of targeted ransomware negotiation. https://doi.org/10.48550/arXiv.2110.00362

Schafer, J., Odle-Dusseau, H., & Levine, T. (2021). Communication-based crisis negotiation: Lessons from law enforcement hostage negotiation. *Security Journal, 34*(1), 23–39. https://doi.org/10.1057/s41284-020-00241-w

Taylor, P. J. (2002). A cylindrical model of communication behavior in crisis negotiations. Human Communication Research, 28(1), 7– 48. https://doi.org/10.1111/j.1468-2958.2002.tb00797.x

Ting-Toomey, S. (1999). Communicating across cultures. Guilford Press.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Wilner, A., Jeffery, A., Lalor, J., Matthews, K., Robinson, K., Rosolska, A., & Yorgoro, C. (2019). On the social science of ransomware: Technology, security, and society. *Comparative Strategy, 38*(4), 347–370. https://doi.org/10.1080/01495933.2019.1633187

Appendix: AI Statement

During the preparation of this work, the author, Melisa Imsak, used ChatGPT and Grammarly in order to improve structure, wording, and grammar. Additionally, parts of the text were originally written in German and translated into English using DeepL. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the work.

Appendix A: Codebook

Codes	Definitions
Initial Offer	The initial offer that the ransomware groups makes to the victim.
Assertive Pushback	The victim stands their ground or fights back.
Counteroffer by Victim	A counteroffer that the victim makes to the ransomware group.
Discount Request by Victim	The victim requests a discount for the initial price.
Financial Constraint	The victim tells the ransomware group that their company is struggling with money due to issues.
Emotional Plea	The victim uses strong emotional phrases to appeal to the ransomware group.
Information Seeking	The victim asks questions to understand what's happening or the process of ransomware.
Readiness to Negotiate	The victim shows readiness to negotiate with the ransomware through showing initiative in the dialog.

Demand of Quality	The victim demands to see proof of quality from products of the ransomware group.
Strategic Politeness	The victim is polite in order to achieve consensus with ransomware group.
Strategic Delay	The victim delays the covnersation in order to gain more time.
Aggressive Response	The victim aggressively responds to the ransomware group.