

THESIS APPROVAL FORM
NAZARBAYEV UNIVERSITY
SCHOOL OF SCIENCES AND HUMANITIES

FROM SIGNALS TO SILENCE: ANTICORRUPTION ENFORCEMENT AND CITIZEN
REPORTING IN KAZAKHSTAN

ОТ СИГНАЛОВ К МОЛЧАНИЮ: АНТИКОРРУПЦИОННОЕ ПРАВОПРИМЕНЕНИЕ И
СООБЩЕНИЯ ГРАЖДАН О КОРРУПЦИИ В КАЗАХСТАНЕ

СИГНАЛДАРДАН ҮНСІЗДІККЕ: ҚАЗАҚСТАНДАҒЫ СЫБАЙЛАС ЖЕМҚОРЛЫҚҚА
ҚАРСЫ ШАРАЛАР ЖӘНЕ АЗАМАТТАРДЫҢ ХАБАРЛАУЫ

BY

Tomiris Kabayeva

NU Student Number: 201959910

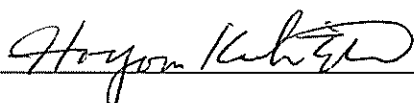
APPROVED

BY

Dr. Hoyoun Koh

ON

13 May 2026



Signature of Principal Thesis Adviser

In Agreement with Thesis Advisory Committee
Second Adviser: Dr. Dinara Pisareva

FROM SIGNALS TO SILENCE: ANTICORRUPTION ENFORCEMENT AND CITIZEN
REPORTING IN KAZAKHSTAN

ОТ СИГНАЛОВ К МОЛЧАНИЮ: АНТИКОРРУПЦИОННОЕ ПРАВОПРИМЕНЕНИЕ И
СООБЩЕНИЯ ГРАЖДАН О КОРРУПЦИИ В КАЗАХСТАНЕ

СИГНАЛДАРДАН ҮНСІЗДІККЕ: ҚАЗАҚСТАНДАҒЫ СЫБАЙЛАС ЖЕМҚОРЛЫҚҚА
ҚАРСЫ ШАРАЛАР ЖӘНЕ АЗАМАТТАРДЫҢ ХАБАРЛАУЫ

by

Tomiris Kabayeva

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts in

Political Science and International Relations

at

NAZARBAYEV UNIVERSITY -
SCHOOL OF SCIENCES AND HUMANITIES

2026

Abstract

This thesis examines how state-led anticorruption enforcement signals shape Kazakhstani citizens' perceptions of the sincerity of anticorruption efforts and their willingness to report corruption. Despite a centralized Anticorruption Agency, various forms of reporting channels such as the 1424 hotline and the e-Otinish portal, citizens remain reluctant to report corruption. I argue that the missing piece in current accounts of citizen reporting is the perceived sincerity of the state's anticorruption commitment — citizens' belief that enforcement reflects genuine efforts to combat corruption rather than politically selective enforcement. Building on Su and Ni's (2018) middle-range theory of whistleblowing, the collective action framework of Persson, Rothstein, and Teorell (2013), and signaling theory in the corruption literature (Barbabela et al. 2022), I hypothesize that consistent and impartial enforcement signals institutional sincerity, which in turn raises willingness to report.

I test these hypotheses through a 2×2 factorial survey experiment with Kazakhstani students at Nazarbayev University, complemented by nine semi-structured interviews with citizens who had encountered or had ever reported corruption. The experimental findings partially confirm the theoretical argument, as the hypothesized treatment effects are not statistically significant in the full sample, and perceived corruption prevalence emerges as the dominant predictor of perceived sincerity, eclipsing the effect of experimental signals.

In contrast to anticipated results, neither high nor low corruption tolerance moderated the effect of treatments on reporting, indicating that the cynicism equilibrium discourages both citizens who oppose corruption and those who tolerate it. Respondents from interviews dismissed anticorruption enforcement as insufficient on three grounds: invisible punishment without identifiable consequences read as cheap talk, sporadic prosecutions read as

performance rather than commitment, and high-profile cases read as factional politics rather than principled enforcement.

Thus, this thesis contributes to the literature on whistleblowing and collective action by specifying perceived sincerity as the mediating mechanism identified but left underdeveloped in prior research, and by specifying the scope conditions under which institutional anticorruption signals can encourage citizen whistleblowing in hybrid regimes, namely sustained signal accumulation and the absence of a politicized interpretive frame.

Table of Contents

Chapter 1. Introduction: Does the State Mean It? Institutional Sincerity and the Willingness to Report Corruption in Kazakhstan.....	1
1.1 The Paradox of Anticorruption Reporting in Kazakhstan.....	1
1.2 Research Question and Argument.....	3
1.3 Organization of the Thesis	5
Chapter 2. Literature Review: Anticorruption, Collective Action, and the Determinants of Citizen Reporting	7
2.1 Anticorruption as a Political Phenomenon	7
2.2 Theoretical Perspectives on Corruption and Citizen Action	10
2.3 Reporting Corruption: Definitions and Determinants	13
2.4 Institutional Anticorruption Efforts: State-led Strategies and Their Impact on Corruption Reporting.....	17
Chapter 3. Theoretical Framework: Sincerity Signals as the Missing Mechanism in the Reporting Decision.....	20
3.1 Collective Action Logic and Reporting Inaction.....	24
3.2 Anticorruption Efforts as Signals of Institutional Commitment.....	26
3.3 Theoretical Mechanism.....	28
Chapter 4. Research Design.....	31
4.1 Experimental design.....	33
4.2 Sample and recruitment	33
4.3 Treatment and Experimental Conditions.....	34
4.4 Outcome variables	37
4.5 Estimation Strategy.....	44
4.6 Qualitative Interviews	45
Chapter 5. Main Results	52
5.1 Data Collection and Sample Characteristics	52
5.2 Manipulation checks	54
5.3 Average Treatment Effects: Difference-in-Means	55
5.4 Regression Estimates of Average Treatment Effects	58
5.5 Perceived Sincerity as Mediator of Reporting Intention.....	60
5.6 Qualitative Findings.....	62
Chapter 6. Conclusion: From Signals to Silence	74
6.1 Null Effects on Reporting Intentions and Perceived Sincerity and the Limits of Institutional Signaling.....	74
6.2 Heterogeneity of treatment effects: subgroup analysis	76
6.3 Limitations.....	77
6.4 Conclusions	79

Bibliography.....	80
Appendix A. Survey Instrument.....	86
Appendix B. Interview Protocol.....	91
Appendix C. Supplementary Statistical Tables	95

Acknowledgments

I am deeply grateful to my primary supervisor, Dr. Hoyoun Koh, whose constant guidance shaped every stage of this thesis. Our weekly meetings not only helped with the research direction but also provided me with the confidence I needed as a researcher during the writing process. His course on experimental political science played an important role in shaping the methodological approach of this thesis, and many of the insights in the following chapters originated in those lectures. I am especially grateful for his patience while I revised, restructured, and sometimes completely rethought the entire project. Professor Koh remained engaged with every new version, exploring the theoretical framework with me and helping me see things I could not see on my own.

I would like to thank my secondary supervisor, Dr. Dinara Pisareva, for her help in the early stages of topic exploration. Her thoughtful guidance was very important in shaping the focus of this thesis.

I extend my sincere thanks to the nine interviewees who shared their time, experiences, and reflections with me. Their openness and engagement made the qualitative part of this study possible and made the research much more meaningful.

I thank my fellow MAPSIR students for their support and for listening to all my thoughts about the thesis. Discussing ideas with them made this journey much more engaging.

Finally, I am deeply grateful to my family, who patiently listened to all my doubts, challenges, and moments of excitement, and who shared in my successes. Their support made everything possible and gave it meaning.

List of Tables and Figures

Table 1. Experimental Conditions	35
Table 2. Manipulation Check Results	55
Table 3. OLS Regression Results	59
Table 4. Direct Effect Models for Mediation Analysis	60
Table 5. Bootstrap Mediation Results	61

Tables in Appendices

Table B 1. Profile of Interview Participants	94
Table C 1. Sample Characteristics by Treatment Condition	95
Table C 2. Randomization Balance Check	96
Table C 3. Mean Reporting Intention and Perceived Sincerity by Treatment Condition	97
Table C 4. Robustness Checks: Treatment Effects among Attention Check Passers	98
Table C 5. Per-Protocol Analysis	99
Table C 6. OLS Regression — Subgroup Analysis by Corruption Tolerance	100

Figures

Figure 1. Theoretical Mechanism: Enforcement Features, Perceived Sincerity, and Reporting Intention	28
Figure 2. Reporting Intention by Treatment Condition	56
Figure 3. Perceived Sincerity by Treatment Condition	57

Chapter 1. Introduction: Does the State Mean It? Institutional Sincerity and the Willingness to Report Corruption in Kazakhstan

Governments across the world increasingly recognize that corruption is a serious problem for political, economic, and social development. Fighting against corruption has become one of valence issues – the issue on which there is a high level of public consensus on the measures being taken to address the existing problem and achieve the desired outcome (Curini and Martelli 2015). The valence nature of anticorruption makes it politically advantageous for states to adopt anticorruption narratives and related policies. As a result, both democratic and authoritarian countries develop similar institutional architecture in the field of anticorruption: whistleblower protection, financial incentives for whistleblowers, anticorruption agencies, reporting channels, and platforms. Thus, the infrastructure of anticorruption is now a standard. The question this thesis takes up is why, in many settings, citizens decline to use it.

1.1 The Paradox of Anticorruption Reporting in Kazakhstan

Along with other countries that are committed to anticorruption norms and laws, Kazakhstan has a strong and extensive anticorruption infrastructure. The country has openly expressed its anticorruption stance recently. In 2019, President Tokayev introduced the "Listening State" initiative, whose aim is to restructure and strengthen feedback mechanisms between the state and citizens. Zhana Kazakhstan is also a key and major initiative aimed at institutional and political reset, including in the anticorruption system. This concept clearly recognizes the systemic threat of corruption to the state and places an open emphasis on investigations and the actual application of penalties. Following the January protests, also known as the "Bloody January" events, which took place in various regions of Kazakhstan, the

"Jana Kazakhstan" initiative actively draws on an anticorruption narrative, emphasizing equality before the law, de-oligarchization, the inevitability of punishment, and asset recovery.

Moreover, reporting corruption is legitimized and encouraged in the country. Reporting of corrupt practices is recognized as one of the anticorruption measures in the country. People can directly report corruption via the hotline number of the Anticorruption Agency 1424, messengers and chatbots (WhatsApp; Telegram), as well as written appeals to the anticorruption service and appeals through an "open dialogue" on the website of the anticorruption service.

Along with the fact that reporting corruption is legally recognized as one of the key anticorruption measures in Kazakhstan, the National Security Committee of Kazakhstan (Anticorruption agency previously) has introduced measures aimed at encouraging citizens to report cases of corruption. Monetary reward can be given when there is salient assistance in revealing corruption by revealing information regarding a corruption deal and/or a wanted person. The financial incentive is paid to a person who reported the information regarding corruption if a court decision or a guilty verdict is issued against the accused individual.

As the act of reporting is inherently sensitive, people are resistant to blowing the whistle on corruption. According to the Life in Transition survey (2016), only 6.71% of people in Kazakhstan believe that ordinary citizens can combat corruption by reporting it to officials. The belief that ordinary people can do something to help combat corruption in a country reflects a pessimistic outlook: approximately 45% of respondents assume that the fight against corruption lies beyond the capabilities of ordinary citizens.

According to the LiTS survey results, the most frequently cited reason why people abstain from reporting corruption was the concern about the possible repercussions, indicated by 26% of respondents. Around 14% of respondents believe that people are discouraged from whistleblowing on corruption because they risk exposing their own involvement in giving

bribes. The difficulty of proving actual corruption emerged as another deterrent to reporting corruption, with 13.2% indicating this particular reason.

With respect to the social acceptability of reporting corruption, there is a paucity of empirical data on how people perceive reporting. Findings from the LiTS survey demonstrate divergent perspectives on whether reporting corruption is socially acceptable in Kazakhstan. In broad terms, the distribution of responses is fragmented, yet the most common view among respondents (27.6 %) was disagreement with the fact that reporting corruption is socially acceptable in the country. A total of 37.7 % of respondents (aggregating responses of “disagree” and “strongly disagree”) opposed the idea that reporting is socially acceptable. Conversely, a total of 25.7% (aggregating responses of “agree” and “strongly agree”) of respondents regard it as socially acceptable.

The literature offers a useful diagnostic for why citizens may not engage in anticorruption activities. Existing scholarship emphasizes that what matters is not whether anticorruption mechanisms exist, but whether citizens read them as reflecting genuine institutional commitment rather than performative governance, intra-elite competition, or selective politics (Barbabela et al. 2022; Carothers 2020; Li and Manion 2022). In this sense, reporting corrupt behavior can be viewed as a classic collective action problem (Persson, Rothstein, and Teorell 2013). Citizens are more likely to report violations when they expect the state to act effectively and is committed to fighting corruption. When these expectations are not met, citizens become disengaged, not because they lack access to reporting channels, but because they do not believe that their opinions and actions will matter.

1.2 Research Question and Argument

This thesis argues that the missing piece in current accounts of citizen reporting is the perceived sincerity of state anticorruption commitment. Sincerity, in this study, refers to

citizens' belief that anticorruption efforts reflect genuine institutional motivations rather than selective, symbolic, or political ones. Building on signaling theory in the corruption literature (Barbabela et al. 2022; Barbabela 2025; Magalhães 2022), I argue that citizens infer sincerity from observable features of enforcement — specifically its consistency and its impartiality. Consistent enforcement, in which the government systematically responds to reported cases, and impartial enforcement, in which sanctions are imposed equally irrespective of the rank of the accused official, can constitute a signal of sincere institutional anticorruption commitment. When citizens perceive enforcement as genuine, the collective action equilibrium that underpins non-reporting can be disrupted. The central argument of this thesis is therefore that observable institutional behavior shapes individuals' perceptions of governmental anticorruption sincerity, which in turn influences their reporting intentions.

This thesis addresses the following research question: Do the consistency and impartiality of anticorruption enforcement increase citizens' perceived sincerity of institutional commitment, and does perceived sincerity, in turn, raise citizens' willingness to report?

To answer this question, the thesis develops and tests a theoretical argument in three parts. First, drawing on the collective action theory (Persson et al. 2013; Ostrom 1998), I argue that citizens' decisions to report are shaped not only by personal morality but by shared expectations regarding how the state addresses corruption. When citizens expect corruption to be systemic and institutions to be insincere and ineffective, abstaining from reporting is individually rational, even for those who oppose corruption. Second, drawing on signaling theory (Barbabela et al. 2022), I argue that institutional anticorruption efforts can update these shared expectations only when they emit signals that are costly to fake. Anticorruption rhetoric alone — what Barbabela et al. (2022) call "cheap talk" — cannot perform this function, because verbal opposition to corruption is universally available and politically cheap. Observable enforcement features, by contrast, can convey credible information about institutional

commitment when they are difficult to mimic. Consistent application of sanctions across cases and impartial application of sanctions across ranks are two such characteristics of sincere anticorruption commitment. Third, I argue that the cognitive bridge between these institutional signals and individual behavior is perceived sincerity — citizens' inferred belief that the state is genuinely committed to fighting corruption rather than taking a stand against corruption for performative purposes.

From this argument, three testable hypotheses follow. H1 predicts that exposure to impartial enforcement increases perceived sincerity relative to partial enforcement. H2 predicts that exposure to consistent enforcement increases perceived sincerity relative to inconsistent enforcement. H3 predicts that perceived sincerity is positively associated with willingness to report corruption.

The argument advances the existing literature in three ways. First, it specifies the mechanism that Su and Ni (2018) leave unspecified — namely, why some institutional anticorruption efforts crowd in whistleblowing behavior among citizens while others crowd out. The argument advanced here is that the relevant variation is not in enforcement volume but in the features of enforcement that citizens interpret as signals. Second, it extends the collective action framework by identifying perceived sincerity as a specific, measurable construct through which institutional signals can update equilibrium beliefs. Where Persson et al. (2013) identify shared expectations as the binding constraint on reform, this thesis identifies a particular kind of signal that may shift those expectations. Third, it integrates signaling theory with the whistleblowing literature that has not previously been developed in this form.

1.3 Organization of the Thesis

My thesis consists of seven chapters. In the next chapter, I critically review the literature on citizen reporting of corruption, focusing on three factors that scholars identify as central to

reporting behavior: individual characteristics such as corruption tolerance, situational conditions such as the perceived prevalence of corruption, and institutional anticorruption efforts. I also discuss two competing theoretical frameworks — the principal-agent approach and the collective action approach — and explain why the latter offers a more compelling account of citizen behavior in contexts of systemic corruption. Next, in Chapter 3, I present my theoretical framework, which builds on Su and Ni's (2018) middle-range theory of whistleblowing, the collective action logic developed by Persson, Rothstein, and Teorell (2013), and signaling theory as applied to anticorruption by Barbabela et al. (2022). I argue that institutional anticorruption enforcement shapes citizens' willingness to report corruption through perceived institutional sincerity — citizens' belief that the government is genuinely committed to fighting corruption rather than acting selectively or politically. Specifically, I argue that consistent and impartial enforcement signals such sincerity, which in turn raises citizens' intention to report. I discuss my research design, sample, and methodology — a 2×2 factorial survey experiment combined with semi-structured interviews — that I used to test my hypotheses about the effects of consistency and impartiality on perceived sincerity and reporting intentions in Chapter 4. In Chapter 5, I report, analyze, and discuss findings from the survey experiment, including the main regression results, mediation analysis, and robustness checks. In Chapter 6, I present the qualitative interview findings, which serve as triangulation and help to interpret the experimental results. Finally, the concluding chapter reflects on the significance of the research findings, acknowledges the limitations of the study, and provides recommendations for future research.

Chapter 2. Literature Review: Anticorruption, Collective Action, and the Determinants of Citizen Reporting

Corruption is a pressing issue that is strongly opposed and addressed by international actors and citizens around the world. Likewise, many states adopt a firm anticorruption position by demonstrating the outcomes of diverse anticorruption strategies (Barbabela et al. 2022). These initiatives are aimed at preventing, combating, and reducing the level of political and bureaucratic corruption. While governmental institutions monitor corruption through centralized enforcement, citizens tend to participate in anticorruption civic activism, which is usually legitimized and encouraged by the state itself.

One of the main objectives of anticorruption strategies is to motivate people to act against corruption and directly participate in public whistleblowing in order to investigate and punish cases of corruption. However, despite a variety of anticorruption measures implemented by governments, people remain hesitant to report corrupt behavior, especially in environments where corruption is normalized (Peiffer and Walton 2022). The existing literature points to the need to explain why anticorruption efforts do not always translate into greater citizen engagement in activities such as whistleblowing.

2.1 Anticorruption as a Political Phenomenon

Corruption is widely recognized as a major constraint on progressive transformation, democratic development, and social justice across all political systems. It induces a variety of detrimental consequences for governance, social trust, and state legitimacy in both developing and developed states (Heywood 1997, Johnson 2004). Most significantly, corruption is not a static phenomenon; it can "adapt to different contexts and changing circumstances" (Transparency International). Despite its obvious deleterious consequences, corruption persists across different settings and sectors because of its evolving nature.

At its core, anticorruption deals with the problem of information about corrupt practices. Corruption dealings occur in a hidden and covert manner, which frequently obscures the disclosure of corruption cases. Traditional methods of law enforcement rely on police officers, prosecutors, and investigative agencies to uncover illegal or unethical behavior, such as corruption. Research shows that while reliance on traditional top-down monitoring by formal authorities remains a conventional approach to combat corruption at different scales, its implementation is frequently resource-intensive and limited by informational asymmetries (Abbink and Serra 2012; Givati 2016). Consequently, scholars have emphasized persistent challenges in detecting and controlling corrupt behavior through state-led enforcement alone.

Beyond state-led anticorruption strategies, a growing body of research conceptualizes anticorruption as involving the prevention, detection, and control of corruption through various forms of civic engagement. Citizen involvement in combating corruption is commonly referred to as anticorruption civic activism or bottom-up monitoring (Peiffer and Alvarez 2016; Su 2020). Related research further highlights that reporting behavior constitutes a core component of bottom-up anticorruption efforts and is widely viewed as an integral part of anticorruption governance (Su and Ni 2018). Citizen engagement is often embedded within institutional anticorruption policies, which render such participation legitimate and compatible with top-down anticorruption strategies (Su and Ni 2018, Su 2020). Governments frequently rely on citizens' monitoring activities to obtain information directly from individuals who possess knowledge or evidence of corrupt behavior (Near and Miceli 1995; Lavena 2016). Moreover, existing studies suggest that disclosure-based approaches involving citizen participation can help overcome key constraints of top-down monitoring such as limited access to information and high enforcement costs (Abbink and Serra 2012; Givati 2016). Consequently, existing research recognizes the role of citizens in anticorruption efforts and commonly refers to such involvement as bottom-up civic engagement.

Beyond its administrative dimension, anticorruption is employed as a political strategy to gain electoral support and respond to the public's demands. Opposing corruption as a normative stance has been broadly endorsed by the public and rarely contested in principle, which makes it one of the valence issues. This creates strong strategic incentives for political actors to emphasize corruption as a campaign tool (Curini and Martelli 2015). Building on this logic, Barbabela et al. (2022) argue that anticorruption rhetoric risks being perceived as skeptical by people, especially in contexts characterized by systemic corruption and low institutional trust. Drawing on signaling theory, the authors argue that only costly and observable acts that reveal both impartiality and efficiency in combating corruption can signal institutional commitment to anticorruption. As a result, the research shows that anticorruption serves as a political strategy for institutional actors, thereby facing a credibility deficit in commitment that must be credibly addressed.

Corruption is a social problem, and efforts to combat it can be more effective by diagnosing corruption and incorporating both public and collective initiatives. People tend to judge corruption based on broader moral beliefs by referring to society's capacity to provide an integral and just system. The public tends to view corruption in terms of favoritism and unjust privilege granted to certain groups, which ultimately undermines merit-based advancement in the governing process (Mungiu-Pippidi, 2015, 4-5). Mungiu-Pippidi (2006) suggests that identifying the systemic mode of governance in which corruption exists is a significant step before addressing anticorruption initiatives. The author argues that in developing and post-communist countries, corruption is embedded in particularism – the mode of governance where public goods are distributed unfairly, relying on personal networks. In contrast, under ethical universalism, the laws and standards apply to everyone in society regardless of their group affiliation, making corrupt behavior an apparent deviation from the universalist norm (Mungiu-Pippidi 2006, 88). Mungiu-Pippidi (2006) thus claims that corruption itself functions as a norm

in contexts where particularism prevails, institutionalizing unequal treatment based on an individual's social status and personal connections. From this perspective, it is significant to “diagnose” corruption – determine whether corruption is an exception or the norm – before designing anticorruption strategies. Consequently, anticorruption policy should consider the nature of the governance mode itself to achieve effective control of corruption.

2.2 Theoretical Perspectives on Corruption and Citizen Action

Much of the scholarly literature has analyzed both corruption and anticorruption efforts within the logic of the principal–agent framework (Klitgaard 1988; Rose-Ackerman 1978). According to this theoretical perspective, corruption involves agency–principal relations, where principals usually expect agents to act on their behalf since the latter delegates the power and authority to the agents. Political corruption occurs when agents or officials follow their interests instead of considering the demands of principals — population (Rose-Ackerman 1999, 91). In bureaucratic corruption, the roles of agents and principals differ from those in political corruption: bureaucrats are typically the agents, and rulers are the principals. Importantly, across both types of corruption — bureaucratic and political — agents possess more information than principals, which results in information asymmetry between the two types of actors (Rose-Ackerman 1978).

The principal–agent framework suggests that the problem of corruption is typically attributed to the agent, while the principal is expected to exercise control and monitoring of corruption (Mungiu-Pippidi 2006). Within this framework, anticorruption efforts are expected to be effective when principals are able to prevent, punish, and negatively influence the agent's motivation or opportunity to engage in corruption. Ideally, principals must actively monitor corrupt practices among agents. Institutional reforms often rely on anticorruption strategies derived from the principal–agent logic, typically focusing on reducing agents' discretion,

limiting their monopoly power, and strengthening accountability in public office (Persson et al. 2013, 452). Persson et al. (2013) argue that anticorruption policies grounded in this framework are “analytically useless” in the case when the principals themselves are corrupt, and the responsibility for controlling corruption is not undertaken by those actors. Implementing anticorruption reforms based on the principal–agent framework poses a significant practical challenge, particularly in developing countries where corruption tends to be systemic and embedded within institutions (Persson et al. 2013).

Recognizing the limitations of the principal–agent framework in explaining systemic corruption, scholars have increasingly advanced the collective action approach. A collective action dilemma emerges when individually beneficial choices made in isolation lead to an outcome that is worse for the group as a whole. In the context of corruption, while society would benefit from universal honesty, individuals might find personal advantages by deviating from cooperative behavior and participating in corruption (Carson and Prado 2016, 58). In a collective action scenario, individuals’ choices are conditioned by shared expectations about how others will act (Ostrom 1998). If corrupt behavior is the expected norm, most of the people, including both groups of actors – agents and principals – are expected to act in a corrupt way (Persson et al. 2013, 456-457). Moreover, in a thoroughly corrupt environment, when principals want to act honestly, the costs of principled behavior appear to be high compared to the possible individual benefits. The costs of acting with the principles of integrity in highly corrupt settings may involve being socially excluded or stigmatized, and in some cases, facing retaliation such as job loss or even threats to one’s life from other members of the community (Carson and Prado 2016, 58). Along with personal costs associated with principled behavior, acting against corruption is seen as something that would not change “the rules of the game”. Thus, people choose corrupt acts over non-corrupt ones in environments where this behavior is expected (Persson et al. 2013).

When corruption is an expected norm in society, implementing anticorruption reforms can present a “second-order” collective action dilemma (Ostrom 1998). Anticorruption initiatives require cooperation among actors who personally benefit from corruption and therefore tend to resist such cooperation. Since corrupt actors themselves often participate in enforcing anticorruption policies, these initiatives are vulnerable to alteration and manipulation aimed at preserving the existing corrupt system (Carson and Prado 2016, 58). Thus, constructing effective anticorruption institutions and reforms is itself a collective action problem, which is particularly challenging to overcome in societies dominated by corrupt actors.

Under the collective action problem framework, individuals are often reluctant to engage in anticorruption activities because their individual action is viewed as insignificant. People are unlikely to support anticorruption reform if they acknowledge that “the rules of the game” remain unchanged and society maintains its existing norms regarding corruption (Carson and Prado 2016, 56). In turn, Bauhr (2017) contends that individuals’ engagement in anticorruption activities depends on the type of corruption they experience as well. Bauhr (2017) argues that experiencing need corruption—a type of corruption that citizens practice to receive public services or gain access to “fair treatment”—increases the willingness for civic engagement against corruption because people feel as victims of a corrupt system and are motivated to impose sanctions to secure fair treatment. In contrast, when people try to gain special illicit advantages, they engage in greed corruption. Acting against this type of corruption is less advantageous for those who participate in it because the consequences of reducing corruption can have tangible drawbacks (e.g., losing some privilege), whereas the advantages of establishing a corruption-free society are less tangible and take longer to achieve. The collective action framework suggests that the effectiveness of anticorruption initiatives can be

enhanced by integrating and transforming informal rules and norms (Carson and Prado 2016, 58-59).

2.3 Reporting Corruption: Definitions and Determinants

Literature examines the impact of distinct messages about corruption on reporting corruption. Peiffer & Walton (2022) conducted a survey experiment to explore how different anticorruption messages affect individuals' willingness to report corruption in Papua New Guinea (PNG). The results of the study showed that messages focused on the detrimental effect of corruption on local communities increased citizens' willingness to report it (Peiffer & Walton, 2022). Importantly, messages emphasizing immorality, illegality, and the prevalence of corruption had no significant effect on the citizens' willingness to report corruption.

Recent studies have argued that individual-level beliefs produce an effect on whether a person would be inclined to report corruption. According to Amini et al. (2020), having individualistic values is associated with a higher likelihood of being willing to report a corrupt act. Individualism plays a key role in reporting behavior because it encourages people to act in accordance with personal responsibility and accountability, thus increasing the willingness to report corruption. In contrast, societies where collectivist values prevail tend to conform to group norms and are more likely to conceal reporting (Amini et al. 2020). In accordance with this argument, Lagos and Zhang (2025) empirically supported the view that individuals who prioritize collectivist values are less likely to act against corruption in the setting where corruption is systemic.

Previous studies aimed to examine the relationship between knowledge of anticorruption campaigns and reporting corruption. Using cross-sectional data from a household drop-off survey in Mauritius, Jaunky et al. (2020) found that media consumption and knowledge about anticorruption legislation increase the probability of reporting an act of

corruption. Interestingly, higher political interest was found to reduce the probability of whistleblowing, highlighting that political engagement does not always lead to anticorruption civic action (Jaunky et al. 2020). Yet some scholars dispute the view that educational factors produce a direct effect on civic willingness to expose corruption. For instance, Lagos and Zhang (2025) indicated that educational anticorruption programs promoted through social media did not increase willingness to report directly. Nevertheless, the authors demonstrate that anticorruption educational programs produce a moderating effect on reporting intentions: such programs enhance the effect of perceived corruption control on reporting inclination (Lagos & Zhang 2025, 10).

Research on citizens' willingness to report corruption sought to explain how individuals' previous experience in reporting and prior engagement in corrupt practices influence their willingness to act against corruption. For example, Lagos and Zhang (2025) found that an individual's prior reporting experience increases his or her future willingness to expose corrupt behavior. Regarding personal experiences with bribery, Oboro-Offerie et al. (2025) found that individuals who have personally experienced corruption are more likely to fear reporting it due to concerns about negative consequences.

The literature emphasizes that citizens' anticorruption activism depends on how they perceive the corruption environment. Research examines two pivotal cognitive dimensions that can affect civic anticorruption activism, including reporting corruption: perceived prevalence of corruption and perceived effectiveness of state control over it (Bauhr 2012; Peiffer and Alvarez 2016).

Research consistently shows that perceiving corruption as widespread tends to discourage citizens' engagement against it. A well-established finding is that awareness of a systemic corruption environment creates a collective action problem in which individuals believe that their efforts are likely to alter established practices, while the potential benefits of

a cleaner system would be shared collectively. In their interview study conducted in Kenya and Uganda, Persson et al. (2013) revealed that individuals were reluctant to report corruption in their countries because it would not change the system anyway (458). Similarly, Bauhr and Grimes (2014) argue that in contexts of endemic corruption, transparency, which exposes more information about corrupt practices, may produce resignation rather than indignation, further discouraging citizens from engaging in accountability-seeking activities. By using data from over 50 countries, the authors found that increased transparency about corruption is associated with lower levels of political involvement and political interest among citizens. The recent research of Oboro-Offerie et al. (2025) found that citizens who perceive corruption as prevalent on a state level tend to exhibit more fear of reporting corruption. Thus, high perceived corruption is associated with corruption fatigue, withdrawal from civic action, and decreased reporting behavior.

At the same time, perceptions of state effectiveness in combating and controlling corruption might exert more ambiguous effects on the willingness to engage in anticorruption and reporting. A growing body of evidence suggests that perceiving that the government is capable of managing corruption, in some contexts, stimulates citizens to engage in anticorruption actions, as it raises the perceived utility of their participation. Peiffer and Alvarez's (2016) research shows that citizens are more willing to engage in anticorruption activism when they view corruption as serious and believe that the government's efforts to control it are effective. Similarly, analysis of survey data gathered from corruption witnesses in Chile showed that higher perceived effectiveness of corruption control is positively and significantly associated with the willingness to report corruption (Lagos and Zhang 2025). Cross-national hybrid analysis confirms that macro-level effectiveness captured by the Control of Corruption Index negatively affects fear of reporting across African populations. On the other hand, Bauhr (2012) argues that people can free-ride when they are confident in the

government's efficacy in managing corruption. Existing research thus suggests that perceptions of effective governmental control over corruption are associated with contested pathways of anticorruption action.

Indeed, intention to act against corruption is not equitable to the actual behavior of acting against corruption and, importantly, does not always lead to the anticorruption activism, particularly, reporting. Building on Rest's (1983) ethical decision model, Ponemon (in Zhang et al. 2009) and later Culiberg and Mihelič (2017, 793) distinguish four sequential components of whistleblowing process: awareness of wrongdoing, ethical judgement, whistleblowing intention, actual whistleblowing behavior. Studies emphasize a substantial "distance" that separate a whistleblowing intention from the overt act (Bagozzi 1992, Miller and Grush 1988, Mesmer-Magnus and Viswesvaran 2005). Mesmer-Magnus and Viswesvaran (2005) argue that intention and action "are separated by psychological, motivational, and implementation processes" such as planning, psychological commitment, situational appraisal, and effort (292). For this reason, this thesis treats willingness to report as conceptually separate from actual reporting, acknowledging that anticorruption efforts may shape attitudinal readiness differently from concrete reporting behaviors.

The study conducted by Su and Ni (2018) highlights that intentions to blow the whistle and actual whistleblowing behavior can differ significantly. The researchers used objective and self-reported data gathered from residents in Guangdong province, China, and examined the influence of individual, situational, and institutional factors on whistleblowing intention, actual whistleblowing act, and anonymity decisions. The findings of their study indicate that high whistleblowing intentions are predominantly explained by tolerance of corruption: individuals who exhibit a low tolerance for corruption demonstrate a 160% higher likelihood of intending to report corrupt behavior (Su and Ni 2018, 413). Intention to report corruption can reduce when the severity of government corruption increases, meaning that people tend to express their

willingness to report less when they acknowledge a harsh situation with corruption in their country. In contrast, the authors revealed that actual whistleblowing acts appear to increase when people know that the situation with corruption is severe. This suggests that citizens are pressured by the high prevalence of corruption and want to personally engage in anticorruption efforts in reality, even though they tend to be less expressive about reporting when referring only to their willingness.

2.4 Institutional Anticorruption Efforts: State-led Strategies and Their Impact on Corruption Reporting

States demonstrate their commitment to fighting corruption by implementing a range of anticorruption policies in practice. Institutional anticorruption practices constitute a key element of good governance. The literature explains controlling corruption in different ways; however, it does not explicitly delineate a solid definition of this phenomenon. Wedeman (2005) states that government anticorruption campaigns are usually aimed at controlling corruption rather than eliminating it, which is nearly impossible (94). Scholarly works admit that institutional anticorruption efforts are state-led actions designed to deal with government corruption through a variety of actions, such as detection, punishment of corrupt behavior, education, and generally deterring the rate of corruption. These efforts differ in their key objectives, continuity, and intensity.

Existing research distinguishes between top-down and bottom-up approaches to controlling corruption. While top-down institutional efforts refer to strict monitoring and punishment by centralized anticorruption institutions, bottom-up anticorruption efforts rely on citizen participation, such as public monitoring and reporting (Gong and Xiao 2017, 178-179; Su 2020). Furthermore, state-led efforts are usually categorized by the intensity of state enforcement into routine bureaucratic anticorruption work and high-profile campaign-style

enforcement (Ko and Weng 2012). Anticorruption campaigns usually represent episodic and centralized crackdowns through an intense burst of enforcement coordinated by government anticorruption bodies. They are typically concentrated on targeted investigations rather than embodying the universal application of rules. (Wedeman 2005). The research highlights that anticorruption campaigns are frequently associated with revealing and reporting a vast amount of anticorruption information, including the cases of investigation, trials, and prosecution of corrupt officials (Sun et al. 2022). Compared to campaign-style efforts, routine anticorruption work is characterized by permanent, rule-based, routinized practices embedded in the state's legal and bureaucratic architecture. Such efforts incorporate a spectrum of actions, including routine policing and day-to-day monitoring (Su and Ni 2018; Sun et al. 2022; Wedeman 2005).

The literature evaluates government anticorruption efforts in many aspects: their effectiveness in reducing corruption incidence, and their impact on state legitimacy, public trust, and civic participation (Treisman 2007; Bauhr 2017). The success of anticorruption strategies depends on public support and the extent to which they are able to encourage average citizens to report corruption. Otherwise, anticorruption agencies can be severely constrained by access to information in their pursuit of investigations (Gong and Xiao 2017). Scholars have addressed the specific effects of different institutional efforts on reporting corruption.

The authors evaluate institutional anticorruption mechanisms as securing reporting actions for people who can rely on institutional channels. Institutional factors can provide “legitimate venues” for individuals who are willing to contribute to combating corruption (Heyes and Kapur 2009, cited in Su and Ni 2018). Increasing investigations and punishment can raise the expected cost of corruption for officials, thus deterring misconduct and making whistleblowing an efficacious act (Abbink and Serra 2012). Scholars state that, regardless of individual motives, institutional anticorruption as an extrinsic factor can “change the calculus of whistleblowing decisions” (Su and Ni 2022, 409; Gong and Xiao 2017).

Theoretically, robust institutional actions are expected to stimulate reporting through complementary or substitutional effects. Su and Ni (2018) argue that institutional anticorruption efforts do not affect whistleblowing in one direction—either encouraging it or discouraging it. Instead, they suggest that state-led enforcement can either crowd in or crowd out citizen participation. Institutional efforts, such as convictions against corrupt officials, can function as a complementary mechanism to citizens' active role in reporting by incentivizing them to join the fight against corruption. In contrast, the substitute effects of anticorruption efforts are reflected in the government's management of corruption and citizens' free-riding.

Su (2020), in turn, showed a more nuanced mechanism of how institutional anticorruption shapes public whistleblowing. Her study found that institutional anticorruption practices produce direct and indirect effects on public whistleblowing, which may produce conflicting outcomes. While punishment of senior officials (top-down institutional anticorruption) directly discourages public reporting of corruption due to the free-riding problem, it can indirectly positively affect whistleblowing by signaling more corruption incidents. In turn, more citywide public whistleblowing (bottom-up institutional anticorruption) stimulates further acts through social learning, yet indirectly reduces the probability of reporting corruption by demonstrating reduced corruption occurrence to the public.

Chapter 3. Theoretical Framework: Sincerity Signals as the Missing Mechanism in the Reporting Decision

This study builds on previous research emphasizing how institutional anticorruption efforts shape citizens' behavioral responses by altering evaluations of public institutions and perspective of corruption as a collective action (Su and Ni 2018; Su 2020; Persson et al. 2013). While the literature suggests that state-led anticorruption actions matter for citizens' mobilization against corruption, it remains less explicit why some anticorruption efforts lead to more citizens' engagement while others discourage people from reporting. This study argues that the sincerity of governmental anticorruption efforts affects citizens' willingness to report corruption through institutional signals of impartiality and consistency. In the context where corruption constitutes a collective action problem, people infer the sincerity of anticorruption commitment from visible enforcement actions. Sincerity of anticorruption practices can represent a strong cue that combating corruption is genuine and implemented in a predictable manner. Belief that government is genuinely and consistently tackling corruption, in turn, motivates people to contribute to anticorruption personally. In contrast, distrust in anticorruption measures can render skepticism toward the efficacy of corruption controlling by authorities (Barbabela et al. 2022). Thus, this study argues that enforcements implemented through inconsistent and impartial punishments of corrupt officials signal insincere institutional commitment to combat corruption and discourage citizens from engaging in whistleblowing.

Existing research recognizes that the decision to report wrongdoings is multi-determined. Whistleblowing literature identifies three broad categories of factors that influence citizens' proclivity and behavior regarding corruption reporting. These factors can be broadly divided into individual-level and institutional-level or intrinsic and extrinsic factors (Gong and Xiao 2017, 179). Su and Ni (2018) propose a middle-range theory of public whistleblowing that conceptualizes citizens' reporting decisions as functions of three main clusters of factors:

individual, situational, and institutional. The suggested framework highlights that intention to report is typically driven by moral individual-level values such as tolerance of corruption, while the act of reporting refers to a rational behavior that is predominantly shaped by situational and institutional forces.

The previous research recognizes that it “remains puzzling why some institutional factors crowd-out whistleblowing acts and others crowd in” (Su and Ni 2018). Both Su and Ni (2018) and Su (2020) consider top-down anticorruption enforcements as substitutive to citizen engagement and reporting. The researchers suggest that the intensity and the actual amount of state-led anticorruption efforts can result in decreased reporting behavior since people might delegate fighting against corruption to the government. At the same time, more anticorruption efforts have complementary effects on reporting behavior by providing legitimate channels for whistleblowing and reducing potential costs. However, the authors do not develop a theoretical account of why certain forms of top-down anticorruption enforcement encourage citizens to act against corruption while others do not. Thus, the prior research provides an incomplete picture of why citizens respond heterogeneously to anticorruption measures and under what conditions they choose to report corruption.

In this study, I follow the collective action perspective and build on Su and Ni’s (2018) middle-range theory of whistleblowing, which posits that institutional factors can exert an independent effect on reporting behavior regardless of intrinsic factors. As mentioned before, Su and Ni (2018) suggest that institutional anticorruption practices can either incentivize reporting by legitimizing it or, on the contrary, substitute reporting by demonstrating that the government handles corruption control itself, making personal contribution of citizens less necessary. Collective action theorists, in turn, argue that individuals’ decisions and willingness to act against corruption critically depends on shared expected behaviors of others and evaluations of public institutions (Persson et al. 2013; Peiffer and Alvarez 2016). In this study,

I follow the collective action-perspective and extend Su and Ni's logic by arguing that individuals' willingness to engage in anticorruption is contingent on the government's ability to signal sincere institutional commitment to anticorruption.

Key Concepts

The theoretical explanation in this study employs a range of concepts, including reporting intention, sincere institutional commitment to anticorruption, consistency of anticorruption punishments, and impartiality of anticorruption punishments.

Reporting corruption

In this study, reporting refers to the form of citizen behavior that is associated with “disclosing illegal, immoral and illegitimate practices, particularly related to government corruption, wrongdoing and misconduct in pursuit of public values” (Su 2020; 228). Following Su (2020), reporting corruption is a pro-social behavior aimed at combating corruption through the disclosure of information about a corrupt incident to relevant authorities.

Whistleblowing intention captures an individual's self-reported propensity or readiness to report witnessed wrongdoing to authorities under hypothetical conditions. Willingness to report is understood as an attitudinal orientation toward exposing corruption and as a form of prosocial motive (Su and Ni 2018, 416). Whistleblowing acts, in turn, are conceptualized as a decision to reveal information about corrupt behavior, which is associated with personal risks and “rational decision calculus”. Following the literature, this thesis treats willingness to report as conceptually distinct from actual reporting, acknowledging that anticorruption efforts may shape attitudinal readiness differently from concrete reporting behaviors.

Sincerity of anticorruption commitment

Sincerity of anticorruption institutional commitment refers to the extent to which governmental anticorruption measures reflect genuine and principled motivations rather than

selective and biased intentions. Although previous research acknowledges the fact that internal motives of political institutions regarding anticorruption cannot be observed directly by citizens, it can be assessed through institutional signals or heuristics (Barbabela et al. 2022; Barbabela 2025). Observable institutional actions and patterns of behavior usually demonstrate how sincere or insincere anticorruption efforts are.

Consistent with prior research, sincerity is distinct from the effectiveness of outcomes of anticorruption enforcements. Barbabela (2025) suggests that the sincerity of anticorruption commitment is derived from the consistent performance of political institutions and their ability to have “a shared understanding of principles of justice and integrity”. The consistency of reaching similar conclusions – non-conflicting verdicts — regarding corruption cases is viewed as a strong cue of certainty and sincerity in institutional commitment (Barbabela 2025). In this thesis, the sincerity of anticorruption commitment is conceptualized as citizens’ perceptions of whether governmental actions aimed at combating corruption are carried out consistently and impartially by formal institutions.

Consistency of anticorruption efforts

Consistency of anticorruption efforts refers to the predictable and regular application of anticorruption sanctions at a systemic level. Enforcements are consistent when accusations are followed by punishment in the vast majority of cases and carried out systematically. Previous studies examined the consistency of institutional behavior in the context of judicial disagreement and argued that conflicting verdicts on the same corruption case serve as a cue of insincere commitment, which in turn, reduces citizens’ willingness to report corruption (Barbabela 2025). Similarly, Magalhães (2022) examines the absence of punishment in the consequence of anticorruption investigations, as the inability to emit a costly signal, decreasing the public trust in the authorities. By extending the logic of intra-case judicial consistency (Barbabela 2025; Magalhães 2022), this study conceptualizes consistency of anticorruption

efforts as the systematic application of sanctions following accusations across corruption cases. In contrast, enforcement inconsistency refers to the pattern-level failure of administering punishments following accusations regarding corruption-related behavior.

Impartiality of anticorruption efforts

Impartiality is commonly embedded in broader concepts such as government effectiveness (Uslaner 2008) or the rule of law (The World Justice Project Rule of Law Index 2020). Following Rothstein and Teorell (2008), impartiality implies that, when implementing laws and policies, government officials do not take into consideration any characteristics of the citizen or the case that are not stipulated in advance by law or policy (170). Prior research focused on the impartiality of courts and suggested that judicial impartiality denotes the capacity of authorities to exact punishments free from external political pressure and without selective targeting of defendants. Such punishments of corrupt behavior serve as credible evidence that institutions are genuinely committed to anticorruption (Barbabela et al. 2022). In this study, impartiality of anticorruption efforts refers to the non-selective application of sanctions in response to the corrupt offences, whereby corruption accusations are punished on the basis of the same legal criteria, regardless of the rank or political status of the accused.

3.1 Collective Action Logic and Reporting Inaction

Reporting corruption constitutes a pragmatic collective action problem. Within the collective action framework, people choose to abstain from reporting due to unequal distribution of costs and benefits associated with the act of whistleblowing in expectation that others will engage in anticorruption actions. While the costs associated with whistleblowing predominantly lie on the informants, the benefits of acting against corruption are shared collectively. In particular, benefits accruing from reporting — reduced corruption incidence, greater accountability, stronger governance, and a cleaner system- take the form of public good.

These benefits are non-rivalrous, non-excludable, and shared with all the members of society regardless of who chose to contribute to them. In contrast, the costs associated with reporting – including negative consequences, time, efforts, and psychological burden — fall on the individual who chose to report (Culiberg and Mihelic 2017; Miceli and Near 1985; Rothschild, 2013). Acknowledging the fact that reporting corruption is an individual act that is beneficial for society but costly for the informant, individuals have incentives to refrain from action (Persson et al. 2013).

According to the collective action approach, the decision to act against corruption depends on the shared expectations of how others will behave. The core problem is that the decision to refuse and combat corruption is connected to how citizens and public institutions in society are expected to be corrupt. In the context of systemic corruption, the shared belief that “everyone else” will continue to act corruptly can negatively influence reporting behavior, making it seem futile (Persson et al. 2013). Even when corruption is morally condemned, people can choose to stay silent to not bear the costs of the action. Persson et al. (2013) state that in such settings, acting honestly might be considered as being a “sucker” – paying a high personal price with less noticeable benefits to society at large. According to Peiffer and Alvarez (2016), citizens’ willingness to engage in anticorruption activism might also depend on the perceived effectiveness of political institutions and government in general. Taken together, non-reporting reflects a rational response to a collective dilemma in which individuals expect others to act because of high personalized costs and believe that individual actions will be futile. As a result, the collective action problem surrounding corruption reporting is sustained because of a combination of factors, including perceived personal risks, perceived prevalence of others’ corrupt behavior, and beliefs regarding the effectiveness of individual action.

Alongside perceptions of the prevalence of corruption in society, individuals also evaluate the government’s capacity to address it. A collective action problem arises and is

reinforced when individuals hold certain perceptions about a corrupt environment and evaluations of public institutions in this setting. However, beliefs about the prevalence of corruption and the government's capacity to control it can be updated as a consequence of institutional signaling. In this regard, anticorruption performance may shift beliefs about the government's commitment to combat corruption, thereby shaping citizens' willingness to engage in whistleblowing.

3.2 Anticorruption Efforts as Signals of Institutional Commitment

The sincerity of anticorruption commitment is not directly observable to average citizens. The literature emphasizes that the idea of combating corruption became a valence issue, which makes it universally supported across political systems (Curini and Martelli 2015). In this context, governments increasingly demonstrate their strong opposition to corruption of any type or scale. According to Barbabela et al. (2022), intensified anticorruption rhetoric can be perceived as a “cheap talk” by the public, who, in turn, might struggle to differentiate between sincere or symbolic anticorruption motives of governments. Importantly, the problem of sincerity is maintained even when institutions publicly undertake actions against corruption. Scholars highlight that anticorruption efforts are frequently used as a political tool in intra-elite competition and a tactic to purge rivals (Li and Manion 2022; Carothers 2020; Zhu and Zhang 2017). So, skepticism toward anticorruption commitment among citizens can still persist in response to both public statements and observable government actions against corruption.

As anticorruption practices are demonstrated for different political purposes, citizens struggle to recognize non-biased actions from more symbolic and selective enforcements. More specifically, people may lack confidence in anticorruption actions not only because of their politicized nature, but also due to the fact that formal institutions lack the capacity to punish corruption impartially and consistently (Barbabela et al. 2022; Barbabela 2025; Magalhães

2022). It should be noted that when institutional commitment to fighting corruption is difficult to observe, citizens rely on observable and costly actions to form their assessments. In this regard, when public institutions are able to exact punishments in accordance with impartiality and consistency, people perceive it as a signal of sincere commitment to combating corruption (Barbabela et al. 2022; Barbabela 2025). Whereas prior research treats inconsistent punishment as indicative of insincere commitment, this study expands the conceptualization of sincerity by incorporating impartiality. It contends that perceived sincerity is multidimensional and that both consistency and impartiality in punishment signals sincere institutional commitment and thereby influence citizens' willingness to report corruption. Building on Su and Ni's (2018) theory of whistleblowing, this thesis argues that reporting intentions depend not only on the volume of anticorruption efforts but also on the signals that governments emit when such efforts are implemented through punishment. Consistent enforcement and impartial punishment of corruption are difficult to mimic and thereby convey information about institutional commitment.

Taking into consideration the fact that shared expectations about corruption prevalence and the government's capacity to control it shape the collective action problem, the role of institutional signals becomes specifically relevant. When corruption is punished with limited discrimination toward corrupt offenders and when accusations of corrupt behavior are consistently translated into sanctions at a pattern level, it serves as a signal of sincere anticorruption commitment (Barbabela 2025; Magalhães 2022). Informational cues generated by anticorruption efforts enable more specific evaluation of corruption control and, more importantly, the assessment of the genuineness of these policies. Rather than directly resolving the collective action problem, such observable actions alter the informational context in which citizens assess the institutional commitment of government. In her study, Barbabela (2025) argues that judicial inconsistency provides a strong informational cue that the court and public

institutions as a whole lack the capacity to respond to corruption in a predictable way for citizens. Thus, signals emphasizing observable punishment of corruption can uncover sincere or insincere governmental motives that are otherwise unobservable to citizens.

3.3 Theoretical Mechanism

As discussed above, anticorruption actions provide clearer expectations of how the government responds to wrongdoings. Sincere government's commitment to fighting corruption becomes salient with demonstrative actions that are characterized by certain costly institutional signals (Barbabela et al. 2022). In this study, sincere institutional commitment is signaled through two features of government performance: consistency and impartiality of punishment. The government is consistent when accusations are predictably and reliably translated into sanctions. Consistency is reflected in patterned state-led behavior in which accusations are systematically screened according to predictable rules and not ignored when sanctions are warranted. Government is impartial if the punishments do not discriminate among defendants based on their political status and rank (Barbabela et al. 2022). Such features of government performance constitute a visible signal, making them salient points of reference in settings where citizens might not differentiate between committed and not committed actors. Exposure to state-led actions in the form of impartial and consistent punishment of corrupt officials can represent a salient signal indicating the extent to which the government is sincere in its fight against political and bureaucratic corruption.



Figure 1. Theoretical Mechanism: Enforcement Features, Perceived Sincerity, and Reporting Intention

Once anticorruption actions send institutional signals publicly, individuals draw inferences about the sincerity of anticorruption commitment. People interpret impartial and consistent punishment of corrupt politicians as a sincere motive to fight corruption within the state (see Figure 1). Specifically, impartiality is signaled when corruption cases are addressed and punished equally, regardless of the accused's rank and political position. For example, when punishments are addressed predominantly to lower ranks of officials, it might be interpreted as a symbolic policy that does not target the powerful political actors.

H1: Exposure to impartial anticorruption enforcement increases beliefs that anticorruption efforts are sincere, compared to selective enforcement.

In turn, when the government punishes reported cases in line with their calls for civic engagement, such actions signal another dimension of sincere institutional commitment. In contrast, systematic and patterned failure to impose sanctions can be interpreted as weak commitment and limited capacity to control government corruption. Magalhães (2022) argues that high-scale corruption investigations that end in failure significantly decrease public trust in courts and political institutions. Existing empirical research suggests that citizens are attentive not only to the mere existence of anticorruption enforcement, but also to its regularity. For instance, Zhang and Kim (2018) show that increases in corruption convictions are associated with higher trust in political institutions, whereas declining conviction rates are interpreted as insufficient enforcement. These findings indicate that patterns in the quantity of punishment shape citizens' perceptions of institutional performance, lending plausibility to the argument that enforcement consistency can function as a signal of sincere anticorruption commitment:

H2: Exposure to consistent anticorruption enforcement increases beliefs that anticorruption efforts are sincere, compared to inconsistent enforcement.

Beliefs in the sincerity of anticorruption efforts encourage individual reporting by signaling that institutions are genuinely combating corruption and by increasing the perceived predictability and meaningfulness of reporting as a course of action. In authoritarian regimes, for example, punishment of corrupt politicians is a costly signal: it signals government capacity and moral commitments (Tsai et al. 2022). Pavão (2019) demonstrates that the perception that the judicial system is impartial, trustworthy, and effective is associated with a stronger belief that individuals can personally contribute to anticorruption.

H3: Stronger beliefs in the sincerity of anticorruption efforts are associated with a higher intention to report corruption.

Chapter 4. Research Design

This chapter describes the research design employed to explain how institutional signals shape citizens' willingness to report corruption. The empirical strategy aims to examine the proposed hypotheses in Chapter 3. As previously discussed, this study aims to identify the causal effect of institutional anticorruption commitment, as signaled by governmental impartiality and consistency, on reporting intentions. To test whether these signals shape individual perceptions about the state's commitment to fighting corruption and intentions to report corruption personally, this study employs a randomized survey experiment and interviews with citizens who have encountered corruption and reported it to authorities or attempted to report corruption incidents.

Before explaining my research strategy, it is important to briefly rationalize the choice of experimental design. The experimental approach is used to provide a rigorous strategy for establishing a causal relationship between exogenous institutional signals and endogenous behavioral intentions (Mutz 2011; Druckman et al. 2011). Observational studies of corruption perceptions and corruption reporting are frequently subject to endogeneity concerns. In observational data, individuals' perceptions of institutional performance and intentions to report corruption can be driven by unobserved confounding factors such as civic culture or trust in government, making it difficult to disentangle causal relationships. Moreover, the relationship between independent and dependent variables can reflect reverse causality rather than a clear causal link. By employing random assignment, this study introduces exogenous variation in the independent variable and ensures comparability across treatment groups in terms of both observed and unobserved pre-treatment characteristics. Because treatment assignment precedes the measurement of the outcome variable, the design minimizes the concern about reverse causality. As a result of randomization, any systematic differences in reporting intentions can

be attributed to the manipulated information about institutional anticorruption efforts rather than pre-existing differences among respondents.

Experimental approaches have been widely applied to the studies that examine corrupt behavior and anticorruption individual actions. Experimental research on corruption scrutinized the individual decisions to engage in corrupt dealings with the elements of monitoring and punishments in designed laboratory bribery games (Abbink et al. 2002; Barr and Serra 2009). Similar types of corruption games were used to study the changes in corrupt behavior after the priming effects of social norms related to corruption (Köbis et al. 2015; Köbis et al. 2022). Survey experiments examined the effect of priming messages emphasizing moral views and general tendencies of corruption behavior in society on citizens' intentions to participate in bribery themselves (Agerberg 2021; Corbacho et al. 2016).

Recent experimental research investigates drivers of reporting corrupt wrongdoing and related behaviors, including moral apposition to corruption or support for anticorruption efforts. Survey experiments have been used to analyze how anticorruption messages influence reporting intentions (Peiffer and Walton 2016). In addition, the survey experiments examined how institutional anticorruption performance affects citizens' attitudes towards bribery, anticorruption spending and perceptions of corruption prevalence (Barbabela et al. 2022; Sun et al. 2022). To investigate how citizens respond to the government's anticorruption enforcements and whether these enforcements influence reporting inclinations, this study uses a survey experiment with randomized vignettes. By experimentally varying informational conditions about the different dimensions of government anticorruption performance, the design introduces exogenous variation in perceptions of government anticorruption efforts and allows the study to reflect the theoretical logic of institutional factors' influence.

4.1 Experimental design

This study employs a vignette-based survey experiment to test whether and how institutional signals about anticorruption commitment influence citizens' intentions to act against corruption personally. The vignette-based approach allows modeling and administering realistic scenarios of anticorruption policy performance while maintaining control over experimental stimuli. The experiment is based on a between-subjects design with a randomization of participants across experimental groups. Randomization was implemented at the individual level using the online survey platform immediately after pre-treatment questions were administered.

After being assigned to their respective experimental conditions, participants were asked to answer a set of questions measuring their perceptions of the sincerity of anticorruption efforts and the government's responsiveness to citizens' reports. Following these questions, respondents answered the main dependent variable question—intention to report the incidence of corruption.

4.2 Sample and recruitment

The experiment was administered to students at Nazarbayev University in Astana. Participants were recruited through the university's official student email distribution list. An invitation email containing the survey link and the overview of the study was sent to all currently enrolled undergraduate and graduate students in March 2026. The data collection took place between 20 March and 5 April, 2026.

To align with the study's target population and ethical requirements, participation in this study was restricted to students who met two eligibility criteria: 1) being at least 18 years of age and 2) being a Kazakhstani citizen. These criteria were stated in the invitation email and reconfirmed through self-reported screening questions at the beginning of the survey. Thus, the

final sample consisted of students at Nazarbayev University who are Kazakhstani citizens aged 18 and above.

The convenience student sample was used for several reasons. First, recruiting participants from a single university provided a relatively homogeneous population in terms of age, educational background, social status, and digital literacy, which can reduce extraneous variance and strengthen internal validity (Druckman and Kam 2011). Second, in the context of this study, the student sample represents a theoretically relevant subpopulation for testing the mechanisms underlying citizens' motivations to report corrupt behavior. While student samples can limit the external validity of the present study, they provide suitable settings for examining citizens' responses to anticorruption policies under controlled conditions.

A priori power analysis was conducted using G*Power to determine the minimum required sample size. According to the analysis, a total sample size of 351 respondents would be needed to detect a small-to-moderate effect size ($f = 0.15$) with $\alpha = 0.05$ and statistical power of 0.80 for a between-subjects design with four experimental groups. This suggests that approximately 88 respondents are required in each experimental group.

4.3 Treatment and Experimental Conditions

The study utilizes a 2×2 factorial design, in which two distinct institutional features of anticorruption enforcement are experimentally varied: the consistency of prosecutions and the impartial imposition of sanctions. These characteristics of governance reflect the sincerity of institutional anticorruption commitment and demonstrate credible signals of fighting against governmental corruption (Barbabela et al. 2022; Tsai et al. 2022). The first factor manipulates the consistency of anticorruption enforcement, distinguishing between corruption reports that are systematically followed by investigation and sanctioning and corruption cases that were not effectively pursued by authorities. The second factor manipulates the impartiality of

enforcement. In this respect, impartiality of enforcement is operationalized as the extent to which anticorruption enforcement selectively targets specific groups of public officials based on their rank and political position. This factor distinguishes between enforcement that was applied equally across officials and punishment that was selectively directed at lower-level officials, with higher-level politicians remaining largely unaffected by the policy. Table 1 summarizes four treatment conditions that were generated by a 2×2 factorial design. This design allows the study to estimate both the independent effects of each institutional signal and their interaction effect on the individuals' propensity to report corruption while controlling for the informational environment and minimizing potential confounding factors (Auspurg and Hinz 2015; Mutz 2011).

Table 1. Experimental Conditions

	Impartial enforcement	Selective enforcement
High consistency	Condition 1	Condition 2
Low consistency	Condition 3	Condition 4

Respondents were randomly assigned to one of four experimental conditions and were presented with a short vignette describing the general information about anticorruption agency functioning and its latest policy outcomes.

All four treatments started with the same introductory information:

“In Kazakhstan, corruption by public officials is prohibited by law and may result in legal penalties. Citizens who encounter suspected corruption can submit reports through official channels operated by the Anticorruption Agency. The agency is responsible for receiving complaints, reviewing submitted information, and deciding whether to take further action in accordance with procedures established by law.”

After being exposed to introductory information about the Anticorruption Agency's general responsibilities and the existence of reporting channels through which it carries out its work, respondents were presented with information about authorities' responses to reported corruption cases. The experimental manipulation was embedded in vignettes and varied information about the performance of a newly introduced anticorruption policy, signaling different levels of institutional consistency and impartiality.

In the sincere institutional signal condition, anticorruption enforcement is described as both consistent and impartial: the government regularly and non-selectively follows corruption cases through to punishment. In the fully "sincere" condition, the respondents read the following text: "Since the government introduced a new anticorruption system, citizens have submitted an average of 128.5 reports per month. The *government met its initial target* and has punished *more than 80%* of all registered cases. *Approximately equal proportions* of high-ranking and low-ranking officials were punished."

In the second condition, the enforcement is presented as consistent but selective, emphasizing a successful achievement of anticorruption punishment but targeting mostly low-level officials. The manipulated text stated the following: "Since the government introduced a new anticorruption system, citizens have submitted an average of 128.5 reports per month. The *government met its initial target* and has punished *more than 80%* of all registered cases. A larger share of sanctions was imposed on lower-ranking officials, while high-ranking officials *were rarely sanctioned*."

In the third condition, respondents read the scenario of enforcement being implemented inconsistently but non-selectively: authorities act irregularly and ineffectively against corruption, targeting both low-level and high-level politicians. The "inconsistent and impartial" manipulation text stated the following: "Since the government introduced a new anticorruption system, citizens have submitted an average of 128.5 reports per month. The government *did not*

meet its initial target and has punished *only 15%* of all registered cases. Approximately *equal proportions* of high-ranking and low-ranking officials were punished.”

Finally, in the fourth condition, enforcement is presented as both inconsistent and selective, with authorities rarely applying sanctions in corruption cases and directing them at low-level officials in most cases. The “least sincere” scenarios indicated the following: “Since the government introduced a new anticorruption system, citizens have submitted an average of 128.5 reports per month. The government *did not meet its initial target* and has punished *only 15%* of all registered cases. A larger share of sanctions was imposed on lower-ranking officials, while high-ranking officials were *rarely sanctioned*.”

The vignettes were identical in terms of length, narrative structure, tone, background details, and brief instructions before manipulation of texts across all four conditions. The vignettes only manipulate two main independent variables, thereby varying the share of successful investigations and sanctions and emphasizing whether the measures were applied equally or selectively targeted certain groups of officials. All vignettes were designed to reflect realistic anticorruption practices and punishment outcomes to enhance the mundane realism of the scenarios for the student sample. From an ethical standpoint, at the end of the survey, respondents were informed that the quantitative indicators, evaluative measures, and indicators of institutional performance that were used in the text do not reflect real official statistics or specific cases from life.

4.4 Outcome variables

After reading the vignettes, all participants across four experimental groups answered secondary outcome and primary outcome questions. Two secondary outcome questions were asked immediately after manipulation checks to test the proposed causal mechanisms and attitudinal effects of the experimental manipulations. Secondary outcome questions measured

perceived sincerity of anticorruption policies and perceived expected responsiveness of authorities.

Perceived government sincerity in fighting corruption was operationalized as respondents' evaluations of whether the government is actually willing to punish corruption. Barbabela (2025) theorizes that judicial inconsistency can be a sign of insincere institutional commitment and reduce the likelihood of whistleblowing among citizens. At the same time, the author acknowledges the impossibility to measure the perceived genuineness of anticorruption commitment directly using the available data and instead relies on a proxy measure such as public trust in the judicial courts (178). Given the fact that perceived government sincerity is the central theoretical mechanism proposed by Barbabela (2025), this study contributes to filling a gap in the measurement of perceived sincerity of institutional anticorruption efforts.

Belief in the sincerity of anticorruption commitment was measured using a single-item 5-point Likert-type scale with explicit behavioral anchors at each point. Given the fact that response categories can be the subject to different interpretations by the respondents, the scale included concrete behavioral anchors in order to: 1) enhance construct validity by reducing potential interpretive ambiguity, 2) increase ecological validity by reflecting real-world cues and examples of government anticorruption performance and 3) reduce social desirability bias by referring to explicit behavioral descriptions of state-led actions rather than abstract evaluative terms. The item was placed immediately after the experimental vignette and before any manipulation or attention checks to test the treatment effect while minimizing potential post-treatment deliberation and priming effect from subsequent survey questions (Imai et al. 2011). The exact wording of the survey question and response scale was as follows: "How willing do you think the government is to actually punish corruption when it is reported?" (1 = Not at all willing — the government only pretends to fight corruption 2 = Slightly willing — efforts are mostly symbolic 3 = Moderately willing — the government sometimes acts seriously

4 = Very willing — the government is seriously committed to punishment 5 = Extremely willing — the government consistently and effectively punishes corruption)

The primary dependent variable measures individuals' intention to report corruption incidence. Consistent with previous research on whistleblowing and collective action literature, this construct is operationalized as a behavioral intention to engage in whistleblowing (Su and Ni 2018; Zhang et al. 2025).

The outcome question first presented the hypothetical corruption scenario describing a local public employee who is asking citizens for unofficial payment in exchange for faster government procedure. Self-reported items were adopted from whistleblowing and corruption research (Su and Ni 2018; Su 2020; Zhang et al. 2025). Willingness to report corruption was measured on a 5-point Likert scale, where 1 = “Not at all willing” and 5 = “Extremely willing”, where higher values refer to respondents' stronger reporting intentions. The exact question wording used in the survey was: “If you had reliable evidence of this misconduct, how likely would you be to report it to the relevant authorities?”

The primary outcome question was placed after two secondary outcome question and before manipulation checks and attention checks to capture the causal effect of the treatment and the perceived genuineness of institutional commitment and minimize post-treatment deliberation or priming effect. In addition, phrases such as “reliable evidence” and “unofficial payment” were used to reduce the potential social desirability bias among respondents and to emphasize a more realistic decision-making situation rather than a vague moral obligation.

Pre-treatment covariates

A set of pre-treatment variables was included in a survey to measure respondents' background characteristics and baseline attitudes and perceptions. Pre-treatment variables measuring respondents' corruption tolerance, familiarity with corruption reporting channels, perceptions of corruption prevalence in the country, personal corruption experience and socio-

demographic characteristics were collected after screening questions and before randomization and vignette exposure.

Previous studies emphasize that broader socio-demographic characteristics may influence citizens' willingness to report corruption. Amini et al. (2022) find that women and respondents with a larger household income are more willing to report corrupt incidences. Older individuals tend to be less likely to report corruption or participate in anticorruption efforts (Amini et al. 2022; Peiffer and Alvarez 2016). Higher education level may be associated with greater willingness to engage in whistleblowing (Lagos and Zhang 2025). Scholars explain the variation in corruption perception on the individual level, emphasizing that people assess corruption based on their personal characteristics as well (Melgar et al. 2010; Tverdova 2011). These studies indicate that demographic and socio-economic factors such as age, gender and educational level can be predictors of both corruption perceptions as well as willingness to report corruption.

Major individual socio-demographic characteristics were collected in the first part of the survey. They were collected to provide a descriptive overview of the sample characteristics and to conduct inter-group balance checks. These covariates included respondents' age, gender, level of education, year of study, faculty affiliation, household income, respondents' type of settlement (urban vs. rural) and region of residence prior to the age of 18.

Gender was coded as a binary variable (1 = female; 0 = male). The age of respondents was measured in years and treated as a continuous variable. The level of study of respondents was classified into three distinct categories: Bachelor's, Master's, Phd and was included as a categorical variable. Participants' year of study was included as a categorical variable with five response categories (1 = first year, 2 = second year, 3 = third year, 4 = fourth year, 5 = fifth year). Faculty affiliation was recorded as a categorical variable capturing the respondent's school within the university. Household income was coded as an ordinal variable using self-

reported income categories. A dummy variable was introduced to capture respondent's type of settlement (1 = urban; 0 = rural). Finally, respondents' region of residence was included as a categorical variable and incorporated into the model as a set of binary indicators.

Furthermore, several theoretically relevant predispositional measures were included as pre-treatment covariates. First, respondents' corruption tolerance level was assessed using four survey items on a 1-5 scale (1 = completely unacceptable, 5 = completely acceptable) addressing conflicts of interest, nepotism, gift-giving, and facilitation payments. These scenario-based items were adapted from established measures used in prior whistleblowing research and modified to fit the context of this study (Su and Ni 2018; Lagos and Zhang 2025). The items are: "A public official uses their position to help a relative get a job", "A person gives a gift to a public employee as a thank-you after receiving a service", "A public official accepts a dinner invitation from a business that they oversee or supervise" and "A person gives extra money to a public employee to receive a public service faster (e.g., to obtain a medical certificate more quickly)". The extent to which respondents consider the described behavior of corrupt officials or citizens as acceptable serves as a proxy of their tolerance of corruption. Consequently, an index of corruption tolerance was constructed by averaging four items, with higher values indicating a greater tolerance for corruption and lower values indicating stronger disapproval of corrupt behavior.

As pre-treatment covariates, respondents' perception of corruption prevalence in the country, prior familiarity with reporting procedures, prior exposure to corruption. To avoid post-treatment bias, these questions were asked before the vignette exposure and outcome questions. Perceived corruption prevalence was measured on a 5-point scale, where respondents assessed how common corruption is among public officials. Respondents were asked: "In your opinion, how common is corruption among public officials in Kazakhstan?". Familiarity with reporting procedures was gauged using a single item that asks about the extent to which respondents

know where and how to submit a report. Respondents indicated their knowledge of the procedures for reporting corruption on a scale of 1 to 5, where 1 means “not at all familiar” and 5 means “extremely familiar”. Prior experience with corruption was measured as a binary variable, capturing whether respondents had been asked for an unofficial payment or gift in exchange for certain public service in the past 12 months (0 = No, 1 = Yes).

As some pre-treatment variables may involve sensitive information, careful consideration was given to their placement within the survey instrument. d’Urso et al. (2025) examine whether the placement of sensitive questions that measure prejudice affects the measurement of prejudice itself or experimental outcomes. The study builds on two competing methodological arguments that were the focus of the prior studies. Scholars such as Montgomery et al. (2018) and Coppock (2019) argue that post-treatment variables should not be used as covariates because the experimental treatment itself can influence the measurement of sensitive items. Moreover, conditioning the analysis on a post-treatment variable might undermine the randomization: responses on the covariates are influenced by the treatment, the treated and control groups are no longer comparable on that variable (3). Other scholars like Klar et al. (2020) suggest that asking sensitive questions pre-treatment might artificially “prime” participants. Such kind of priming on the pre-treatment stage can increase the identity salience and distort how respondents interact with the treatment in ways that can alter experimental outcomes. This methodological concern highlights that pre-treatment measurement of sensitive items results in affecting the internal validity of the experimental study.

To avoid pre-treatment bias, questions related to prior reporting experience were placed after the manipulation text and outcome questions. Prior experience reporting corruption is also binary, coded 1 if respondents indicated having previously reported a corruption incidence to the authorities, and 0 otherwise.

Manipulation checks

To validate that experimental treatments produced the intended variation in respondents' perceptions of governmental anticorruption actions across treatment groups, manipulation checks were included in the survey. Following best practices in survey-experimental methodology (Bhattacharjee, 2012; Coleman & Montgomery, 1993), the checks were employed to capture respondents' perceptions of the governmental consistency and impartiality in handling corruption cases presented in the vignette. Consequently, respondents were presented with two distinct perceptual manipulation checks. The manipulation checks were administered before the primary and secondary outcome measures.

Participants were required to indicate the extent to which they agreed with the statements describing the characteristics of the anticorruption policy response. The manipulation checks consisted of the following statements that evaluated the perceived consistency and impartiality of government in controlling corruption: 1) "To what extent do you agree with the following statement: 'The authorities responded consistently to corruption cases'?"; 2) "To what extent do you agree with the following statement: 'The authorities applied punishment equally to high-ranking and low-ranking officials'?". Responses to these items were recorded on a five-point Likert-type scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Higher values indicate stronger perceptions of the informational signals embedded in the experimental treatments.

Following best practices in experimental political science, perceptual manipulation checks were employed only for diagnostic purposes – to evaluate the strength of the experimental manipulation and to conduct robustness checks. Respondents who failed or provided non-expected responses on the manipulation check were not excluded from the main analysis. This was made in order to preserve the integrity of random assignment and to avoid

post-treatment selection bias (Montgomery et al., 2018; Aronov et al., 2019). The manipulation checks show the general validation of the treatment.

Attention checks

The survey also included one instruction attention check to identify inattentive respondents prior to the experimental treatment and to enhance data quality in an online survey environment (Berinsky et al., 2014; Oppenheimer et al., 2009). The attention check was placed immediately after the pre-treatment questions and before the vignette and random assignment to the experimental conditions. Respondents were instructed to select a specific response option (“Somewhat agree”) presented among other Likert-type statements. The item was placed pre-treatment to ensure that subjects who did not pay attention or answered randomly were identified before reading the vignette, thereby maintaining the integrity of the randomization procedure.

Furthermore, respondents who did not comply with the requirements and failed the attention check were not automatically excluded from the main analysis to avoid post-treatment selection bias (Berinsky et al., 2014; Aronov et al., 2019). Instead, robustness analysis was conducted excluding inattentive respondents.

4.5 Estimation Strategy

To test the proposed hypotheses, I employ ordinary least squares (OLS) regression models. Given a random assignment, the treatment is independent of respondents’ characteristics, allowing for the identification of causal effects from the estimated coefficients. The dependent variable – willingness to report corruption – is measured on a continuous five-point scale. Accordingly, OLS regression is employed, as it yields unbiased and easily interpretable estimates of average treatment effects (ATE) (Druckman et al., 2011). The main model is specified as:

$$Y_i = \beta_0 + \beta_1 \text{Impartiality}_i + \beta_2 \text{Consistency}_i + \gamma' X_i + \epsilon_i$$

where Y_i is either respondent i 's perceived sincerity or intention to report corruption, Impartiality_i and Consistency_i are binary indicators of the experimental conditions (high/low levels), and X_i is a vector of covariates (when included). Coefficients β_1 and β_2 capture the main effects of impartiality and consistency (Hypotheses H1 and H2), respectively. To test H3 — whether sincerity mediates the relationship between enforcement signals and reporting intention — a separate mediation analysis is conducted.

Before analyzing the effects on the main outcome variable, manipulation checks are conducted to verify the success of the experimental manipulations. Moreover, a mediation model is estimated (Imai et al., 2010), in which the perceived sincerity of institutional commitment mediates the effect of experimental conditions on individual reporting intentions. Furthermore, the models are estimated via OLS with robust standard errors.

4.6 Qualitative Interviews

In addition to the primary vignette-based survey experiment, the present study includes supplementary semi-structured in-depth interviews in order to triangulate findings and strengthen the credibility of the obtained results. The qualitative component of the study allows the investigation of micro-level mechanisms underlying reporting behavior. Interviews with people who had already reported corrupt cases to authorities and who had witnessed corruption but encountered challenges in whistleblowing illuminate factors influencing individuals' decisions to report and the context-specific challenges that characterize the Kazakhstani setting. First, qualitative evidence was used to scrutinize how people interpret institutional signals that they notice in their daily life and how the perceived information about governmental actions influences their decisions regarding reporting. Second, supplementary interviews were conducted to explore the unmeasured pathway from perceived anticorruption commitment to

individual efforts to counter corruption identified by Barbabela (2025). Third, this component of this study examines context-specific factors that can moderate treatment effects. Thus, the study adopts a mixed-methods sequential explanatory design that complements and strengthens causal inference.

Sampling and recruitment

For the interview component, a purposive sampling strategy is employed to recruit participants who have prior experience with reporting or who had witnessed corruption but opted not to report. Participants were recruited from survey respondents willing to participate in an in-depth interview and through online outreach in messaging platforms. It means that a broader pool of respondents was reached beyond the student population. The interview sample consists of 8–10 individuals with relevant experience who are aged 18 and above and are citizens of Kazakhstan.

Interview procedure

Interviews followed a semi-structured format and consisted of an introductory statement, four blocks of questions, and a final section. Each interview lasted approximately 30-60 minutes. They were conducted either face-to-face on campus or via online meeting platform (Zoom, WhatsApp call), depending on participant preference. Interviews were carried out in Russian or Kazakh, depending on respondents' level of fluency and their comfort level in speaking. The qualitative interviews were conducted between March and April 2026.

The interview guide consisted of open-ended questions organized around four thematic blocks derived from the study's theoretical framework: 1) initial perceptions and trust in public institutions; 2) the decision-making process; 3) perceived sincerity of institutional commitment; 4) experience of reporting and post-reporting perceptions and views. During the interview,

probing questions were used to allow the respondent to elaborate on themes and examples that emerged spontaneously or remained unexplained.

Researcher positionality

It is also important to state explicitly how the author's positionality, such as gender, age, education institution, and other individual experiences and personal background, could influence the research process and interpretation of the qualitative findings. The researcher is a female master's student enrolled in Nazarbayev University, recognized for its strong research focus. Even though the research focused on reporting and personal views regarding corruption in Kazakhstan, which can be considered a highly sensitive topic to discuss, respondents during interviews felt less threatened during the conversation with a young female researcher.

Another constituting factor of positionality was ethnicity and citizenship. Respondents disclosed more information and felt comfortable expressing their views openly, expecting that the researcher was familiar with and understood patterns of corruption in the Kazakhstani context. When respondents understood that I am local, they elaborated on the topics by drawing on specific instances and events that occurred in the Kazakhstani context. It also should be noted that affiliation with Nazarbayev University could also influence respondents' answers. Some respondents who also studied in the university were already acknowledged research policies and procedures and internal rules of the university, so they trusted the researcher more.

Analytical approach

Given the supplementary nature of interviews, deductive thematic analysis was employed to uncover respondents' views and decision-making processes regarding corruption reporting. Deductive approach was used to test the proposed theoretical mechanisms and identify recurring patterns in respondents' motivations for reporting, with particular attention to how perceived sincerity of governmental anticorruption actions shape incentives to act

against corruption personally. The deductive strategy was chosen because the interviews focused on the effect of anticorruption policies on whistleblowing and tested whether these mechanisms operate in real-world decision-making specific to corruption environment and institutional context of Kazakhstan.

Analysis proceeded in five stages and began with familiarization with the data. The interview transcripts were read at least two times to understand the general perspectives and reasonings of respondents. At this stage, initial patterns of relevant to theoretical model were taken into account. After reading all transcripts, the deductive coding was conducted using pre-defined theoretical constructs. Each transcript was coded line-by-line. Codes were assigned to any meaningful segment referring to respondent's perceptions of institutional sincerity, impartiality and consistency of state-led measures and the act of whistleblowing. After completing coding the segments of text, the codes were grouped into higher-order themes corresponding to the core theoretical constructs to elucidate recurring patterns. Then, themes were reviewed in order to ensure a high level of internal coherence within as well as clear differentiation between them. Even though the analysis relied primarily on pre-defined codes, emergent themes that extended beyond the initial theoretical expectations were also considered. Segments that captured mainly emergent patterns in respondents' perceptions of corruption environment and institutional behavior were examined to provide a context for the quantitative data. Once the themes were developed from the coded data, each theme was analyzed as confirmatory, disconfirming, or elaborative evidence regarding the proposed theoretical mechanisms and experimental findings.

A detailed coding table, including deductive codes, emergent themes, definitions, illustrative quotes, and explicit theoretical linkages, is presented in Appendix.

Limitations

Several limitations associated with the qualitative supplementary aspect of research design should be acknowledged. The purposive sample of 9 reporters and non-reporters is limited in scope and may therefore not be representative of the broader population in Kazakhstan. At the same time, it should be noted that this phase of the research was conducted to illuminate the cognitive and perceptual mechanisms rather than to achieve statistical representativeness. Specifically, the interviews aimed to uncover cognitive pathways through which citizens interpret visible efforts of government and translate these perceptions into reporting decisions. Nonetheless, the small sample, which consisted mostly of students and young people, limits generalizability of the insights derived from the study and may constrain the range of perspectives captured.

Furthermore, participants – both reporters and non-reporters – were self-selected: these respondents showed greater interest in corruption-related issues and agreed to discuss sensitive experiences connected to corruption and share their views related to institutional behavior. For example, individuals with more negative experience of reporting, fear of retaliation or deep institutional distrust may have been less willing to participate initially, thus skewing the sample toward those with more reflexive attitudes. While reliance on convenience sample was necessary to identify respondents with relevant experiences, it could generate participation bias. It is impossible to fully eliminate such form of participation bias in studies involving sensitive topics such as corruption, which may affect respondents' willingness to participate and provide more honest responses.

Quantitative and qualitative data were collected concurrently rather than one after another. Interview guide was developed prior to receiving quantitative results and therefore could not be revised or changed according to revealed experimental findings. The interview guide was developed based on the theoretical framework and research question of this study.

After qualitative data and quantitative data were collected, they were compared to evaluate the convergence, divergence and explanatory depth. Thus, one of the key limitations of this study is inability to construct interview guide that would explain emerged experimental patterns.

Another methodological concern is retrospective recall bias: all of the respondents that were interviewed for qualitative phase of the research discussed their perceptions of anticorruption commitment and barriers to report corruption after their personal experiences of reporting or failures to report corruption. This suggests that post-reporting experiences or interactions with authorities could reshape how respondents remember their initial expectations and attitudes toward reporting, performance of authorities and many other aspects concerning corruption. Regarding interviewing non-reporters, the interpretive challenge is somewhat different – this group of participants were asked to recall challenges or barriers to reporting at the time they encountered concrete situation involving corruption. It is important to emphasize that this type of counterfactual recall is prone to post-hoc rationalization because respondents tend to describe plausible reasons for not reporting rather than pointing to actual worries and decisions based on the circumstances at the time. These risks were partially addressed by semi-structured probing to ground responses in specific episodes and lived experiences.

Even though respondents had a right to participate online in interviews and not reveal their identities, turn on the cameras during the discussions, interview responses may be influenced by social desirability bias, particularly given the sensitive nature of corruption-related topics. Participants who encountered corruption but didn't report it could frame their inaction as distrust in institutions rather than attributing their behavior to personal choice. Although semi-structured interviewing with probing follow-up questions were employed to encourage respondents to elaborate on topics related to personal experience, the possibility that answers might be shaped by impression management and tendency to conform to socially accepted norms and expectations.

Research ethics

The entire study, encompassing both the primary vignette-based survey experiment and the supplementary semi-structured interviews, was approved by Nazarbayev University Institutional Research Ethics Committee on March 28, 2026. The approval covered all procedures, including random assignment, data collection, data analysis, and the storage of the information.

Informed consent was obtained prior to data collection. For participants in the online survey experiment, a consent form was presented on the first page of the Qualtrics questionnaire before proceeding with the next questions. The survey was automatically finished for those who declined to provide informed consent. Face-to-face and online interviews with participants who witnessed corruption or reported such cases were conducted only after receiving oral consent.

Special attention was paid to the sensitive nature of the topic (corruption and trust in government). In survey experiment and interviews, participants were reminded that the answers will remain anonymous, no identified or specific information will be recorded, and that there are no “correct” or “incorrect” answers. In addition, no deception was used in either component of the study. All participants were debriefed at the end of the survey via an information page and after each interview.

Chapter 5. Main Results

This chapter discusses the key findings of an experimental analysis testing hypotheses H1-H3. The experiment is implemented in a 2 x 2 factorial design to examine the main effects of each factor – the consistency and impartiality of anticorruption actions on perceived sincerity (secondary outcome variable) and readiness to report corruption (primary outcome variable). The analysis first examines how two independent factors affect perceived sincerity of anticorruption efforts and then how perceived sincerity influences willingness to report corruption.

This section proceeds as follows. The quantitative analysis first presents the characteristics of the sample and the results of randomization verification. The following subsection discusses the results of manipulation checks, which assess whether the experimental conditions were perceived as intended by the respondents. The main part of the section tests hypotheses H1-H3. A factorial regression model is presented first, in which perceived sincerity is the dependent variable and consistency and impartiality of anticorruption punishment are independent variables. The coefficients for the variables of impartiality and consistency are interpreted as estimates of the main effects corresponding to the hypotheses H1 and H2. In order to test hypothesis H3, a mediation analysis is conducted to determine if the relationship between experimental condition and reporting intention is mediated by perceived sincerity. The results of this study also include means of both reporting intention and perceived sincerity based on all possible combinations of experimental conditions. Additionally, some additional robustness checks and per-protocol analyses are provided in the final subsection.

5.1 Data Collection and Sample Characteristics

The survey experiment was conducted after SSH IREC approval from 30 March to 12 April 2026. Participants were recruited through university mailing lists and social media interest

group channels for NU students. A total of 264 respondents took part in the study, and 229 met two eligibility criteria and completed the survey, providing complete responses in all the survey sections. So, the final analytical sample of $N = 229$ consists of respondents who provided informed consent in the beginning, passed two screening questions confirming that they are at least 18 years old and are Kazakhstani citizens, answered both outcome questions, and completed the survey in full.

The average respondent age is 22.1 years with a standard deviation of 4.4. The sample consists primarily of female students (58.9 %) and is also heavily comprised of individuals from urban locations (92.6 %). Most of the survey respondents are enrolled at the undergraduate level of study (bachelor's level), as 72.1 % of the total sample fall into this category. Regarding respondents' pre-existing attitudes toward corruption and personal corruption experience, the sample reports high perceived corruption prevalence ($M = 4.44$ on a 1-5 scale), with 91.7% participants evaluating corruption prevalence at higher scales 4 or 5. While direct personal experiences with corrupt activities are somewhat infrequent (13.7%), very few students have prior experience with formally reporting instances of corrupt behavior (3.2%). Additionally, familiarity with reporting procedures is relatively low, with 37.9% students indicating they are entirely unfamiliar with where and how to report corruption incidents. Sample characteristics by treatment condition are reported in Appendix C (Table C1).

Randomization check

Respondents were randomly assigned to one of four experimental conditions: (1) inconsistent and partial enforcement group ($n = 56$), (2) inconsistent and impartial enforcement group ($n = 61$), (3) consistent and partial enforcement group ($n = 53$), and (4) consistent and impartial enforcement group ($n = 59$). The first group serves as the baseline condition in each regression analysis. This reflects an experimentally created "worst case" enforcement context — both low in consistency and low in fairness.

Balance checks were conducted across these groups using one-way ANOVA for continuous variables and chi-square tests for categorical variables. The results show no statistically significant differences across conditions in age ($F(3,225) = 0.81, p = .489$) or corruption tolerance ($F(3,225) = 1.05, p = .370$). Additionally, chi-squared tests reveal no statistically significant imbalance in terms of gender, regional location, academic degree level, experiences with corruption, or familiarity with reporting corruption. The only potential threat to internal validity is the marginal statistical difference found in the urban variable ($\chi^2(3) = 8.59, p = .035$). However, it can be argued that this difference is due to the small sample size. Bonferroni-corrected pairwise t-tests found no statistically significant differences at $\alpha < .05$ between any two treatment groups relative to either respondent age or respondent levels of tolerance for corruption. Detailed balance statistics are reported in Table C2 in Appendix C.

5.2 Manipulation checks

It is essential to assess whether the experimental manipulations were perceived as intended by the respondents before testing the main hypotheses. Table 2 presents mean scores on manipulation check items (perceived consistency and perceived impartiality) by treatment conditions, along with a set of OLS regression model estimates in which the inconsistent and partial condition was used as the reference category.

These results indicated that the experimental design was effective at separating the perceptions of consistency and impartiality across conditions. First, regarding the perception of consistency, the means of the consistent and impartial treatment condition differed from those of the baseline condition ($\beta = .40$, standard error = .18, p-value = .026), but the means of the inconsistent and impartial treatment condition did not ($\beta = -.10$, p-value = .591). Therefore, there is no indication that the partiality/impartiality manipulation contaminated the perception of consistency. Second, concerning the perception of impartiality, the means of the consistent

and impartial treatment condition differed significantly from those of the baseline condition ($\beta = .48$, $p = .029$), while the inconsistent and impartial condition showed a marginal but non-significant difference ($\beta = .40$, $p = .064$). Importantly, when I examine two treatment conditions that share each core dimension (i.e., consistency or impartiality) separately, their respective means are statistically equivalent, demonstrating that the vignettes effectively isolated these two dimensions. Finally, the overall models are significant for both manipulation checks (consistency: $F(3, 225)=3.06$, $p=.029$; impartiality: $F(3, 225)=3.18$, $p=.025$).

Table 2. Manipulation Check Results

Group	Treatment condition	Perceived consistency (mean)	Perceived impartiality (mean)
1 (<i>control</i>)	Inconsistent & Partial enforcement	2.38	1.88
2	Inconsistent & Impartial enforcement	2.28	2.28†
3	Consistent & Partial enforcement	2.55	1.81
4	Consistent & Impartial enforcement	2.78*	2.36*

Note. † $p < .10$; * $p < .05$ relative to the inconsistent and partial baseline condition.

5.3 Average Treatment Effects: Difference-in-Means

Figures 2 and 3 provide the means for the two outcome measures — reporting intent and perceived sincerity — by treatment conditions, along with their respective 95% CI. A dotted line in each figure indicates the mean value for the baseline group ($n = 56$), where there was both inconsistent and partial enforcement. Table C3 in Appendix C displays mean reporting intentions and perceived sincerity across treatment conditions.

Reporting Intention Outcomes (See Figure 2): The means reported in Figure 1 do not support the theoretically predicted relationship. Specifically, contrary to expectations based on theory, those receiving the most favorable enforcement signal (consistently and impartially enforced) reported the lowest levels of intended reporting ($M = 3.05$), and the least favorably signaled group (partial and inconsistent enforcement) reported the highest level of intended reporting ($M = 3.32$). Mean reporting intentions are also less under the conditions of consistent but partial punishments ($M=3.08$), or inconsistent and impartially applied ($M=2.95$) sanctions. All four conditions had substantial overlap in their 95% CIs. Therefore, there is no statistically significant evidence that treatment had an effect on willingness to report.

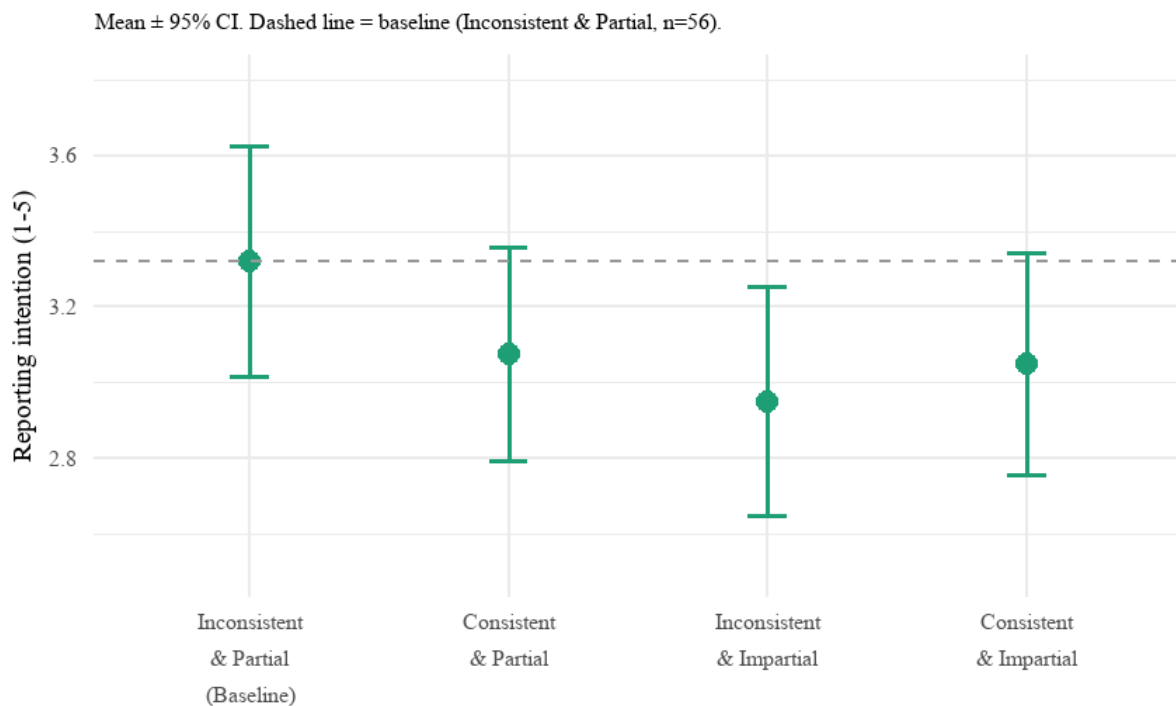


Figure 2. Reporting Intention by Treatment Condition

Perceived Sincerity Outcome (See Figure 3): In contrast to reporting intent outcomes, the means for perceived sincerity by treatment conditions exhibit a theoretical relationship. That is, those who received the most favorably signaling message (consistently and impartially enforced) perceived institutions to have been sincere at the greatest degree ($M = 2.32$); whereas, those who received messaged describing inconsistent but impartially enforced anticorruption measures perceived institutions to have been sincere to the least degree ($M = 1.98$), or 0.34 points less than those receiving the most favorably signaling message. Those receiving either the most unfavorable signaling message (inconsistent and partial, $M = 2.23$) or the message describing consistent but partial enforcement ($M = 2.15$) fell somewhere in between these extreme values. Confidence intervals around the estimates for each condition do overlap, however, and none of the estimated differences meet conventional levels of statistical significance.

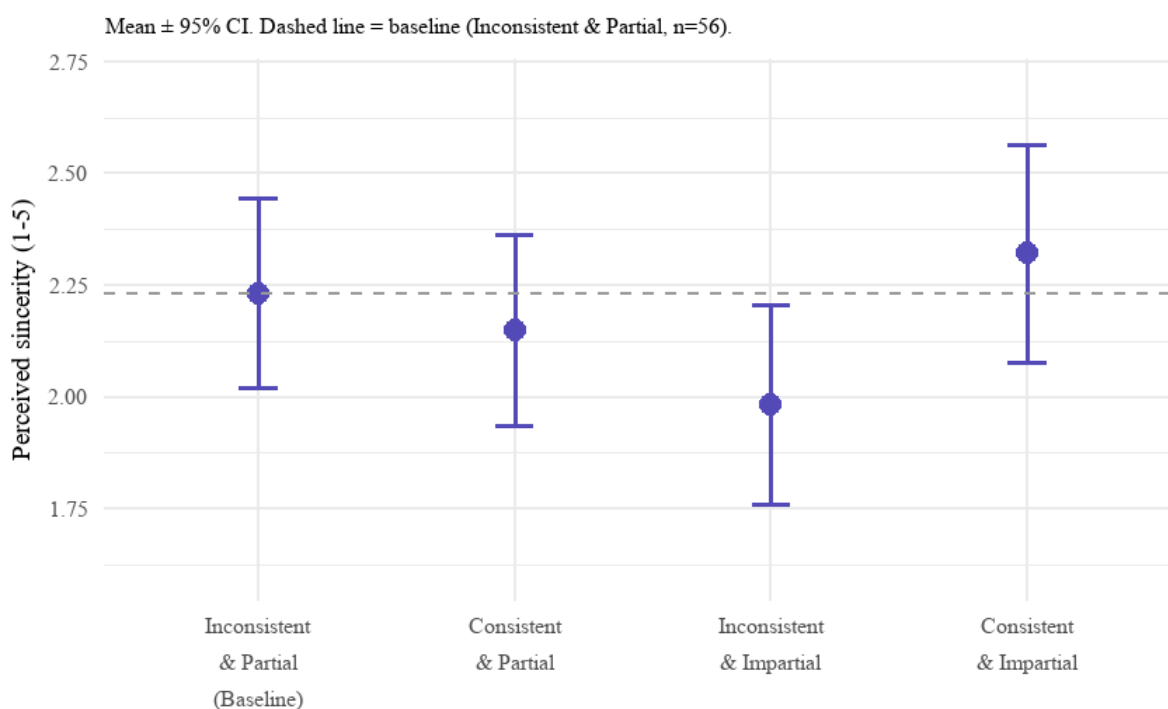


Figure 3. Perceived Sincerity by Treatment Condition

5.4 Regression Estimates of Average Treatment Effects

Table 3 shows ordinary least squares (OLS) estimates of the effects of consistency and impartiality in anticorruption enforcement on reporting intentions and perceived sincerity. A 2 x 2 factorial design was used to analyze the data from this study. Inconsistent and partial enforcement serves as the base category for all the analyses.

The regression models (Table 3) include two binary treatment indicators. CONSISTENT equals 1 for respondents assigned to consistent enforcement conditions and 0 otherwise. The second indicator, IMPARTIAL, equals 1 for those assigned to impartial enforcement and 0 otherwise. Respondents in the baseline condition receive values of 0 on both indicators. Three treatment groups are coded (0,1), (1,0), and (1,1) respectively. Thus, the regression coefficients for CONSISTENT and IMPARTIAL reflect the separate contribution of each enforcement dimension to the outcome, holding the other constant.

In Table 3, Models (1) and (3) provide bivariate effects of the treatment variables (consistent and impartial), while Models (2) and (4) contain control variables, namely corruption tolerance, perceived prevalence of corruption, sex, and age. All four models included the entire sample ($N = 229$). Corruption tolerance that is included in the model is measured as respondents' self-reported willingness to accept certain corrupt behavior, captured via a multi-item scale (higher values indicate greater tolerance for corruption). It is measured as respondents' perceived frequency of corruption among public officials in Kazakhstan, rated on a 1–5 scale, where higher values indicate a belief that corruption is more widespread.

Effects on Reporting Intention. Neither consistent ($\beta = -0.065$, $SE = 0.152$, $p = .670$) nor impartial ($\beta = -0.201$, $SE = 0.152$, $p = .186$) significantly predicts reporting intention in the additive model. Results did not differ when the data set was limited to those who passed the attention check question ($n = 221$; results remained stable; see Appendix C, Table C4).

Effects on Perceived Sincerity. Neither consistent ($\beta = +0.139$, $p = .229$) nor impartial ($\beta = -0.043$, $p = .707$) significantly predicts perceived sincerity in the bivariate model. After controlling for covariates, the treatment effects remain non-significant (consistent: $\beta = +0.161$, $p = .129$; impartial: $\beta = -0.035$, $p = .743$). Corruption prevalence was identified as a strong negative predictor of perceived institutional sincerity ($\beta = -0.501$, $p < .001$), greatly improving model fit ($R^2 = .184$).

Table 3. OLS Regression Results

Variable	(1) Reporting intention	(2) Reporting intention + controls	(3) Perceived sincerity	(4) Perceived sincerity + controls
CONSISTENT	-0.065 (0.152)	-0.104 (0.151)	+0.139 (0.115)	+0.161 (0.106)
IMPARTIAL	-0.201 (0.152)	-0.219 (0.151)	-0.043 (0.115)	-0.035 (0.106)
Corruption tolerance	—	-0.248** (0.092)	—	0.122† (0.064)
Corruption prevalence	—	0.075 (0.106)	—	-0.501*** (0.074)
Gender	—	0.039 (0.154)	—	-0.015 (0.108)
Age	—	0.000 (0.017)	—	0.002 (0.012)
Constant	3.233*** (0.132)	3.497*** (0.679)	2.125*** (0.100)	4.006*** (0.475)
R ²	0.009	0.043	0.007	0.184
F-statistic	0.975 (p = .379)	1.650 (p = .135)	0.796 (p = .453)	8.341*** (p < .001)
N	229	229	229	229

Notes. Standard errors in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Baseline category: Inconsistent & Partial condition. Models (2) and (4) include control variables for corruption tolerance, corruption prevalence, gender, and age.

5.5 Perceived Sincerity as Mediator of Reporting Intention

H3 examines the correlation between perceived sincerity and reporting intention. There are three steps to mediation analysis (Tables 4, 5). In line with H3, each component model is reported before showing the full mediation results.

Step 1 (Model of the Mediator). I estimated a regression of perceived sincerity on two treatment indicators. I found that neither consistent enforcement ($\beta = +0.139$, $p = .229$) nor impartial enforcement ($\beta = -0.043$, $p = .707$) significantly predicts sincerity. This suggests that the experimental treatments did not strongly activate the proposed mediator ($R^2 = .007$).

Step 2 (Model of the outcome). When I regress reporting intention on sincerity alone, I found no significant relationship between these two variables ($\beta = +0.048$, $p = .585$, $R^2 = .001$). This means that perceived sincerity does not independently shape individuals' willingness to report corruption.

Table 4. Direct Effect Models for Mediation Analysis

	Model 1 Sincerity ~ Treatments	Model 2 Reporting Intent ~ Sincerity
CONSISTENT	+0.139 (0.115) $p = .229$	—
IMPARTIAL	−0.043 (0.115) $p = .707$	—
Sincerity	—	+0.048 (0.087) $p = .585$
Constant	2.125*** (0.100)	2.992*** (0.204)
R ²	.007	.001
N	229	229

Note. Unstandardized coefficients with standard errors in parentheses. *** $p < .001$.

Step 3 (Full mediation model): In the full model (see Table 5), a regression of reporting intention on sincerity and both treatment variables simultaneously shows that sincerity remains a significant predictor. ($\beta = +0.048$, $p = .582$). Also, neither treatment coefficient reaches significance. The findings from all three steps indicate that H3 has been rejected.

Table 5. Bootstrap Mediation Results

	ACME	95% CI	ADE	p
Panel A: CONSISTENT pathway				
Indirect effect (sincerity)	0.007	[-0.027, 0.046]	-0.071	.666
Panel B: IMPARTIAL pathway				
Indirect effect (sincerity)	-0.002	[-0.029, 0.024]	-0.199	.954

Notes. ACME = Average Causal Mediation Effect (indirect effect via perceived sincerity). ADE = Average Direct Effect. Percentile bootstrap confidence intervals.

5.6 Qualitative Findings

The qualitative component of the study does not constitute a parallel investigation but fulfills a specific and limited function to explain the results of quantitative data analysis in this research project. The results of the survey experiment showed that there were no statistically significant differences across treatment groups with regard to the respondents' self-reported behavior on whistleblowing intentions. The quantitative part of the research also identified that the perceived sincerity of anticorruption efforts was marginally affected by manipulations. Additionally, a strong relationship was observed between levels of reported corruption tolerance and intention to report corruption personally. The qualitative component of the research process seeks to uncover mechanisms, beliefs, and contextual factors that quantitative measures can only partially capture.

Characteristics of the Participants

Qualitative phase of the research included interviews with 9 participants – 8 non-reporters who witnessed corruption but, for some reasons and barriers, did not report it, and 1 reporter who filed a report to several official bodies. Respondents were recruited through purposive sampling via various online communities, including university student and alumni groups, as well as thematic Telegram channels. Only one respondent who participated in the online survey expressed interest in a follow-up interview and was interviewed as well. The interviews were conducted online via Zoom or WhatsApp call and on the university campus, scheduled according to participants' preferences and availability. Table B1 (Appendix B) presents the demographic profile of individuals who participated in the qualitative phase of the study.

As was mentioned previously in the Research design chapter, the study initially aimed to conduct follow-up interviews with a subsample of survey respondents; however, such an

approach was not realized due to recruitment challenges, particularly, survey respondents' low willingness to participate in interviews. Even though it was specified that the interviews would be anonymized and would not collect personally identifiable information in both recruitment messages and a section of the survey that included an interview invitation, recruitment proved challenging due to the sensitive nature of the topic.

Participants represented a quite diverse demographic profile: age ranged from 18 to 50 years, among the participants were both men and women, residents of both capital and large and small cities, and employees of both government and private sectors. Non-reporters constituted the majority of the sample. In addition, most participants were female, with the majority also being young adults aged 18-29 working in private companies.

Only minimal demographic information – specifically, information on respondents' gender, age group, place of residence, and employment sector – was collected during the interview. This was done to ensure anonymity and reduce social desirability bias when discussing personal experiences of reporting corruption and encountering signs of corruption. No additional personal details were recorded. This approach is consistent with methodological and ethical best practices for qualitative research specifically on sensitive political issues (Braun & Clarke, 2022).

Although the final sample size is relatively small, in most cases, interviewees referred to multiple instances of corruption or reporting. During discussions, most of the participants recalled corruption-related instances that occurred in their own lives as well as in those of their friends, relatives, and others. It is important to note that during interview recruitment, several individuals responded that they witnessed corruption but declined to participate in interviews. The most frequently cited reason for not participating in discussions was employment in public institutions or state-affiliated organizations – such connections with official bodies created concerns about future career and potential professional repercussions. This pattern of self-

inclusion further illustrates the barriers to revealing corruption cases in society, in particular, the pervasive fear of retaliation and the barriers to whistleblowing, especially among those who work within official governmental bodies.

The thematic analysis involved two steps. As was previously mentioned, the main purpose of the qualitative phase is to complement the experiment's findings, so the first step of the qualitative research was deductive: analyzing how participants assessed institutional consistency, impartiality, and institutional sincerity when explaining their decision-making regarding reporting or abstaining from reporting. The second step was inductive – any new or unanticipated patterns that were not identified within the theoretical model and could not be detected by the survey experiment. A total of three analytical themes appeared in the inductive phase. These themes are each examined separately and are discussed in dialogue with the experimental findings.

Before delving into the thematic analysis, I would like to outline how the interview data relate to the central argument of the study. Through an experiment, I tested how information about the consistency and impartiality of the authorities affects the perception of the sincerity of the anticorruption measures taken. Additionally, I evaluated the extent to which sincerity influences reporting based on the theoretical model. The results of the experiment revealed a different picture, indicating that the mediating channel between institutional signals and reporting intention is not activated in the data. The qualitative phase of the study explains and complements the quantitative findings by showing where the causal mechanism begins to unravel and which factors instead come to the forefront in actual reporting behavior.

The qualitative analysis of the interviews aims to provide deeper explanations for why the proposed mechanism works or does not operate in real-world reporting decisions. On the one hand, the effect of informational signals on the perceived sincerity of anticorruption measures was confirmed. Respondents often mentioned that visible punishments of corrupt

individuals that are implemented consistently rather than sporadically are important for them when assessing governmental performance. On the other hand, in contexts of systemic corruption, the interview results suggest that citizens rely on complex and multidimensional heuristics when evaluating institutional signals, and that the standard toolkit of vignette experiments underestimates this complexity. For example, the same signs of sincerity can have the opposite effect of being perceived as fake. In other words, even if punishment may signal that the government is attempting to fight corruption, the punishment of high-ranking officials can sometimes be interpreted as another political purge or a struggle between elite clans rather than a genuine anticorruption effort.

In the interviews, it becomes clear that these same respondents, who are able to articulate criteria for sincerity, still do not report corruption. When asked why, they do not refer to institutional insincerity but instead provide different explanations. Participant 6 says that the system protects itself and that reporting would backfire on the reporter; Participant 7 says that a single report would not make a difference because there are too many corrupt individuals; and Participant 2 says that the corrupt individual was a friend and that he did not want to harm them. In none of these cases is the perception of the state as insincere the main obstacle. Instead, the obstacles are the asymmetry of power, the perceived hopelessness of individual action, and social proximity to the corrupt actor.

This explains why sincerity did not function as a mediator in the quantitative part of the research. The mediating variable — perceived sincerity — in the experiment may not be central to how individuals make reporting decisions in practice. Respondents may perceive the government as either sincere or insincere, yet their decisions are shaped by other considerations that were not affected by the experimental vignettes. These factors are explored further in Themes 2 and 3.

Only one participant had prior experience reporting corruption, and his behavior does not fully support the logic of the original argument. Participant 1 did not report because he believed in the sincerity or effectiveness of government agencies. Moreover, he continued to act even after technical failures through the electronic platform, as well as many redirects from one agency to another. It can be argued that Participant 1's experience is more about expressing moral resistance. The motivation for reporting was a personal intolerance towards corruption, and the expectation of institutional responsiveness was not even considered in the decision-making process.

From these observations, three analytical themes emerge. The first shows how citizens interpret the sincerity of anticorruption policies, thereby delineating the substantive boundaries of experimental manipulations. The second identifies three types of non-reporting behavior, of which only one—disengagement—corresponds to the theoretical model; the other two expand it and point to the limitations of the binary “report/don't report” framework. The third, the “Paradox of the Reported,” focuses on Participant 1's case and discusses institutional failures after the decision to report has been made.

The themes are described individually as follows. The quotes from respondents' comments are translated into English and are interwoven with the analysis to substantiate key interpretations. Each respondent is identified by a number and labeled as either a reporter or a non-reporter during the analysis.

Theme 1: Institutional Signals as Cheap Talk

Each respondent was asked how they differentiate between sincere and insincere anticorruption enforcement, and most of the interviewees repeatedly emphasized certain characteristics and indicators of serious government action against corruption. Even though the question was asked directly and framed the discussion in terms of the sincerity of state efforts, each participant was allowed and encouraged to articulate their criteria for distinguishing

sincere from performative anticorruption. Moreover, despite differences in regions, age, and employment sector, the responses and views were relatively consistent.

Three clear diagnostic criteria emerge from the responses. The first is the visible imposition of consequences on identifiable individuals. Participant 4 from Astana stated: “The real fight against corruption is when I see in the news that someone is really being punished or imprisoned.” Participant 8, also from Astana, expressed a similar idea: “The real fight (against corruption) is when there are real punishments and real faces.” According to Participant 8, it is especially important to see the real punishment of public officials (judicial conviction) when the media and public already have enough evidence that these officials have engaged in corruption. When corruption cases are publicly known but no actions are taken to investigate them, people only reinforce the perception that corruption will go unpunished. But the opposite outcome does not provide credibility too: when a high-ranking official is prosecuted, citizens do not automatically reconsider their skepticism. Participant 8 reinterprets such cases not as principled anticorruption efforts, but as a “clan warfare”. According to her view, politicians are punished not always because of efficient anticorruption system but because their political position becomes inconvenient to those who hold power. Thus, even though prosecutions are implemented by authorities, citizens might not interpret them as sincere but rather perceive them as another political game.

Importantly, participants were not prompted to refer to named individuals or specific consequences; they were simply asked to distinguish between real and fake anticorruption measures. Yet both emphasized the importance of visible, tangible punishment directed at identifiable individuals.

Second, respondents emphasized the importance of regular and sustained anticorruption enforcement, as well as other practices aimed at raising anticorruption awareness over time. Participant 8 further noted that public trust develops when punishments occur regularly rather

than sporadically for “show”: “Trust develops when punishments occur regularly, rather than occasionally for publicity purposes.” Participant 4 expressed a similar idea in the context of university anticorruption education programs, which are mandatory in most public universities in Kazakhstan. In her view, these activities often take the form of symbolic rituals—posters, mandatory sessions, and formal meetings—but lack continuity and commitment to promoting anticorruption awareness: “the curator (advisor) comes, takes photos like he has explained everything to everyone and leaves.” From her perspective, genuine anticorruption, whether it is university anticorruption courses or other government strategies, requires ongoing practice that is implemented well, regardless of its visibility and documentation. This criterion corresponds closely to the consistency dimension: institutional sincerity is associated with regular, sustained enforcement rather than episodic or symbolic action.

Third, two participants articulated a more epistemologically oriented distinction. Participant 2 emphasized transparency and verifiability by citizens as indicators of genuine anticorruption: “For show, all things happen behind closed doors. If a case is open and accessible for fact-checking, then it is something you can rely upon.” In contrast, Participant 5 argued that extensive amplification of anticorruption successes through state-controlled media can actually generate suspicion: “If this story appears frequently in government-controlled media outlets, then questions arise.” Taken together, these perspectives reflect a shared epistemic logic: people perceive government as more sincere when they are able to verify the information regarding governmental corruption themselves, whereas more performative anticorruption relies on state-controlled communication that cannot be independently checked.

Theme 1 shows that respondents refer to sets of criteria for sincerity that overlap with each other. It is important that these criteria are substantively consistent with the consistency and impartiality measures used in the experimental design. This suggests that the conceptual framework is valid at the level of perception, as respondents recognize the kinds of institutional

behavior that, according to the theory, shape perceptions of sincerity. However, confirming the cognitive part of the mechanism does not explain the persistence of behavioral silence in the face of corruption. Therefore, the qualitative analysis suggests that identifying the reasons for non-reporting requires going beyond the original argument and examining how decisions not to report are actually structured. Theme 2 advances this step by showing that non-reporting is not a uniform refusal to act, but rather involves three distinct logics, only one of which fits the theoretical model developed in this work.

Theme 2: Three Distinct Types of Non-Reporting

Most of the research examines non-reporting behavior as a single behavioral outcome that can be explained by personal attitudes, institutional trust, the fear of consequences, and many other factors. The qualitative data from this study distinguishes three main reasons for non-reporting behavior or three main types of non-reporting behavior.

Type I: Disengagement Non-Reporting. This is the type of reporting that the theoretical model predicts: citizens assess that reporting will produce no meaningful outcome and disengage. Participants 6 and 7 exemplify it most clearly.

Participant 6, a non-reporter from Pavlodar who paid a bribe to secure her child's kindergarten placement, describes the institutional logic and barriers that occurred at the time corruption was required and reporting was precluded: "It starts at the top," "it's all connected — everyone covers for each other," "They protect their own." These are not expressions of generalized cynicism toward government. These expressions highlight a perceived asymmetry of power, especially when acting against corrupt individuals: the institutional hierarchy protects its own, and reporting a corrupt actor connected to that hierarchy is seen as more likely to result in retaliation against the reporter than punishment of the offender. Reporting is therefore not merely unlikely to succeed; it is likely to harm the reporter in some cases. It can be concluded

that this is a rational response to a correctly perceived power asymmetry, not an irrational failure of civic engagement.

Participant 7, a non-reporter from Zhezkazgan, reaches the same conclusion through a different epistemic route: " Even if a complaint had been filed, it would not have led to significant consequences, since there are other people involved in corrupt practices, and nothing happens to them." According to him, one reporting act might not be effective in an environment where corruption is pervasive, and most importantly, the single report itself can result in no success.

Type II: Relational Non-Reporting. The binding constraint on reporting is not institutional distrust or any other institutional factor, but social proximity to the corrupt actor — the fact that reporting would harm people the citizen knows personally. Participant 2 can be the paradigm case. He knew existing reporting channels, described his institutional trust as improving over time, and even indicated increasing willingness to report in the future — yet he chose not to report because the corruption he witnessed involved people he knew personally: "I didn't want to tell about this case because they are acquaintances... There is no desire to harm them in any way, even if they are not the main actors."

Type III: Navigational Non-Reporting. Some citizens neither report nor passively comply but make decisions that align with their moral principles and can be made through informal negotiation. Participant 5 was coerced into signing a fraudulent promotion order but negotiated to ensure a deserving colleague was also promoted alongside the "corrupt" one. The corruption was not eliminated and was not reported through official channels, but its damage was partially decreased. This type reveals that anticorruption behavior is not contingent on institutional trust, and that the binary report/don't report framework assumed by both the experiment and the broader literature misses a meaningful category of civic action.

The interview data do not support the assumption that citizens' non-response is homogeneous or mainly determined by institutional conditions. Su and Ni acknowledge that it remains puzzling why some institutional factors discourage whistleblowing while others encourage it. This study offers a partial answer: institutional signals, by definition, cannot influence relational and navigational non-response, since these are not fundamentally responses to the institutional environment.

The Reporter's Paradox: P1 and Institutional Failure After Engagement

Participant 1 shared his experience on how he submitted a formal report through several official channels – official bodies and the online platform e-Otinish. Participant first investigated the corruption case himself through being exposed to details of government procurement in his work. His decision to report and investigate corruption cases was predominantly influenced by his intolerance of corruption and personal values such as honesty. From the perspective of the study's initial argument, this behavior requires explanation: Participant 1 had no reason to perceive the institutions as sincere, and nevertheless decided to act. In terms of Hirschman's framework (1970), his choice represents voice in a situation where most respondents in the sample default to exit or loyalty—either avoiding interaction with the state or tacitly accepting the existing state of affairs. At the same time, voice here is not driven by trust in institutions, but instead expresses a moral stance rooted in internalized normative commitments. This is consistent with the quantitative findings: in the models of Chapter 5.1, tolerance of corruption, rather than perceived sincerity, emerged as a significant predictor of the intention to report.

However, his personal experience of reporting worsened his perception of anticorruption institutions. The e-Otinish system rejected his complaint for technical reasons — he registered and submitted it as an individual, not a legal entity, after which he was unable to successfully report through the online platform. Then, the participant contacted government

agencies through cell phone and directly went to public organizations which are responsible for corruption investigations in government procurement. After contacting government agencies, he was redirected between several institutions without clear guidance on what to do next. As a result, the case did not move forward. Importantly, the case itself was also difficult to discuss, as it involved irregularities in public procurement rather than straightforward bribery evidence. This made the case less obvious to others, less easily understood, and difficult for institutions to act upon in the absence of concrete evidence.

This case differs from the others in the sample. The first participant's experience uncovers an additional potential barrier. Even when an individual decides to report, the public institutions may be unable to process the report efficiently and may even refer the case to other agencies, making the process itself more complicated for citizens. In other words, barriers arise not only before reporting when many factors play a role in decision-making, but also during the report processing. This pattern reproduces a characteristic asymmetry long discussed in the whistleblowing literature. Previous studies suggest that whistleblowers often act alone and make decisions based on internal moral conviction, while the institutions receiving the report are resistant to processing such information (Miceli, Near, and Dworkin 2008; Alford 2001). Participant 1's case demonstrates that in the Kazakhstani context, a voice channel technically exists—there is e-Otinish, and there are relevant agencies—but in practice, the process does not lead to resolution. Instead, it dissipates the individual's efforts through repeated redirects between authorities.

What is also crucial is that this experience does not remain confined to Participant 1. When the person who faced challenges in cooperating with the state in the past shares his experience with others, he generates social information about institutional reliability. This information diffusion about real reporting experience may shape others' perceptions of institutional competence and their willingness to report in the future. One frustrated reporter

within a social network may reduce the willingness of many potential reporters around him. This points to post-reporting social diffusion as a theoretically important mechanism that future research should take into account.

Chapter 6. Conclusion: From Signals to Silence

This study examines how institutional signals of consistency and impartiality within anticorruption enforcement affect citizens' willingness to report corruption through their perceptions of government commitment. Building upon the collective action theory (Persson et al. 2013) and Su and Ni's (2018) middle-range theory of whistleblowing, it was predicted that citizens who perceive anticorruption punishment as both impartial and consistent would be more likely to view the government as sincere in its commitment to fight corruption, and that this belief would positively affect reporting intentions. However, the proposed hypotheses in this study were rejected. The results obtained, in turn, allow us to clarify the conditions under which the signaling mechanism can operate and identify the limitations of existing theoretical models. This section builds this argument sequentially, integrating quantitative experimental data with the results of qualitative interviews.

6.1 Null Effects on Reporting Intentions and Perceived Sincerity and the Limits of Institutional Signaling

The null results can be theoretically informative and may speak to broader corruption literature. They tell us that using vignettes to create an enforcement "signal" is not enough for people to modify their intended behavior or assessments of whether institutions are serious and committed enough to fight corruption. In further sections, I explain how interview data explains the disbelief in information about governmental anticorruption policies and their implementation outcomes.

Neither consistency nor impartiality of anticorruption enforcement produced statistically significant differences in reporting intention within the full sample. A similar pattern was observed for perceived institutional sincerity: both conditions also failed to reach conventional statistical significance in the overall model as separate factors. The results can

contradict the theoretical mechanism proposed by Barbabela et al. (2022), according to whom impartial and effective punishments for corruption can send a credible signal to citizens about the genuine anticorruption commitment of institutions. However, per-protocol analysis that was limited to participants who correctly learned experimental manipulations ($n = 92$), reveals other results of treatment effects on two dependent variables. Among these participants, consistent law enforcement significantly increased perceived sincerity ($\beta = 0.463$, $p = .003$), and impartial enforcement also significantly increased respondents' perceived sincerity ($\beta = 0.379$, $p = .019$). That is, this study shows that there is a theoretically predicted mechanism that exists in the data, but it is limited to a narrow group of respondents who perceived manipulation as intended by the study.

Perceived corruption prevalence was the strongest negative predictor for perceived sincerity across the entire sample, also explaining most of the variance in the sincerity model. Regardless of what enforcement signal citizens receive, citizens who view corruption as pervasive in their society were significantly less likely to have faith in the sincerity of their government in its efforts to combat it. Roughly 92 percent ($n = 151$) of all respondents ($N = 229$) rated the prevalence of corruption among government officials as 4 or 5 out of 5. In this context, the experimental vignette, which consisted of one article reporting on anticorruption punishment, was likely viewed by respondents as an abstract and uncharacteristic example and therefore, is unlikely to have altered people's long-held views about the nature of government and the state's commitment to fighting corruption. This finding relates to Persson et al.'s (2013) argument about the self-reinforcing nature of the collective action problem. This implies that when citizens expect corruption to be widespread and believe that institutions are incapable of effectively addressing it, individual signals of commitment are discounted. It is possible that this equilibrium was too strong for the experimental manipulation to overcome.

The treatment does not influence sincerity; the findings suggest that perceptions of corruption prevalence shape personal beliefs about institutional sincerity, irrespective of specific state action(s) taken in anticorruption. Therefore, it can be argued that the credibility problem faced by anticorruption agencies is likely rooted in the broader information environment in which these actions are evaluated, rather than in the effectiveness or quality of separate enforcement actions.

The qualitative phase of the study was conducted to explain and supplement the argument. The interviews show that anticorruption measures are indeed perceived and differentiated into sincere and insincere measures. Moreover, respondents used the criteria of consistency and impartiality, which were also embedded in the experimental vignettes. In other words, the theory captures two real criteria citizens use to evaluate the state's anticorruption efforts. However, in the actual decision-making process of reporting, the perceived sincerity of institutions is not a decisive factor, but rather a parallel consideration. In respondents' own explanations, other factors come to the forefront: the expected effectiveness of reporting in an asymmetric power situation, social proximity to the corrupt actor, and the respondent's own normative attitudes. This explains the null result of the experiment in a meaningful way. The variable manipulated in the experimental design exists in citizens' perceptions, but in their own explanations, it is not the central factor they cite when describing their decision to report or not report corruption.

6.2 Heterogeneity of treatment effects: subgroup analysis

To examine heterogeneous treatment effects, the sample was split into low-tolerance respondents ($n = 123$), who reject corruption, and high-tolerance respondents ($n = 106$), who are more accepting of it.

Neither the low-tolerance group ($n = 123$) nor the high-tolerance group ($n = 106$) showed statistically significant effects of CONSISTENT or IMPARTIAL on either reporting intention or perceived sincerity (see Appendix C, Table C6). These results suggest that the null effects in the main analysis do not conceal offsetting subgroup effects. Rather, the treatment effect appears uniformly null regardless of citizens' moral stance toward corruption.

This result is substantively important. Su and Ni's (2018) theory suggests that individual tolerance for corruption is a key predictor of reporting intention and decision to report, and this study confirms this in the main analysis ($\beta = -0.248$, $p < .01$). However, tolerance does not appear to moderate responsiveness to institutional signals. In other words, citizens who are most morally opposed to corruption are no more responsive to enforcement signals than citizens with a higher tolerance for corruption. This suggests that the divide between moral stance and institutional trust may run deeper than existing theory implies, and that the corrupt equilibrium described by Persson et al. (2013) may be equally stable regardless of individual moral values.

Notably, however, perceived corruption prevalence predicts perceived sincerity significantly more strongly in the low-tolerance group ($\beta = -0.665$, $p < .001$) than in the high-tolerance group ($\beta = -0.271$, $p < .05$). Citizens who most strongly oppose corruption appear most vulnerable to background pessimism about its prevalence. This is a paradox that directly follows from the logic of Persson et al.: it is those who morally expect more from institutions who are most affected by the gap between normative expectations and perceived reality, and it is these individuals whose prior beliefs most strongly block the perception of any institutional signals.

6.3 Limitations

This study has several limitations that suggest directions for further work.

The first limitation is that the present study has limited external validity. The experimental sample consisted of students enrolled at Nazarbayev University. While this group is theoretically relevant, the majority of students are politically aware and likely to engage with corruption-related issues — it does not reflect the diversity of Kazakhstani citizens more broadly.

A second limitation concerns the treatment itself. The vignette-based design included an exposure of respondents to a single informational vignette, while the literature posits that perceived sincerity is updated through verifiable evidence that is costly to mimic. This design choice was necessary for causal identification, but it also means that the present study cannot directly test how more costly signals, such as live events (e.g., judicial convictions of real politicians), would affect perceived sincerity. For this reason, this study points to the value of natural experiments for future research to see if live institutional signals can actually exert an effect on perceived sincerity and reporting decisions. Also, the study could not test whether longitudinal exposure to consistent and impartial enforcement would produce the belief change that a single vignette could not induce. A natural next step is a longitudinal or panel-based design that tracks how citizens' sincerity perceptions evolve in response to actual enforcement events over an extended period.

Third, the qualitative component of this thesis was modest in scope, with nine interviews drawn from a purposive sample. While these interviews were sufficient to identify theoretically meaningful patterns, they cannot establish the prevalence of relational barriers, informal civic strategies, or politicized interpretations of enforcement in the broader Kazakhstani population. Larger qualitative samples or structured survey items targeting these specific cognitive patterns would allow future research to estimate how widespread each of these mechanisms is and how they interact with enforcement signals at the population level.

6.4 Conclusions

Three main conclusions follow from this study.

First, there is evidence of a mechanism linking institutional signals to perceived sincerity — as shown in the per-protocol analysis — but this mechanism appears to require repeated or accumulated signals rather than a single exposure. In context of systemic corruption, the level of trust required for such signals to be credible may not be reached.

Second, perceived corruption prevalence appears to be a stronger determinant of institutional sincerity than the quality of individual enforcement signals. This might suggest that anticorruption strategies focused on isolated successes may have limited effectiveness, as citizens are likely to interpret them within a broader and stronger perception of systemic corruption.

Third, the typology of non-response developed from the interviews partially extends Su and Ni (2018). It suggests that a reporting intention and behavior are shaped by institutional, situational, or individual factors and are therefore not responsive to enforcement signals, regardless of their quality. This raises broader questions about the limits of institutional explanations in studies of whistleblowing behavior.

Bibliography

- Abbink, Klaus, and Danila Serra. 2011. "Anti-Corruption Policies: Lessons from the Lab." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1971779>.
- Abbink, Klaus, Bernd Irlenbusch, and Elke Renner. 2002. "An Experimental Bribery Game." *Journal of Law, Economics, and Organization* 18 (2): 428–454. <https://doi.org/10.1093/jleo/18.2.428>.
- Agerberg, Mattias. 2022. "Messaging about Corruption: The Power of Social Norms." *Governance* 35 (3): 929–950. <https://doi.org/10.1111/gove.12633>.
- Alford, C. Fred. 2001. *Whistleblowers: Broken Lives and Organizational Power*. Ithaca, NY: Cornell University Press.
- Amini, Chiara, Elodie Douarin, and Tim Hinks. 2022. "Individualism and Attitudes towards Reporting Corruption: Evidence from Post-Communist Economies." *Journal of Institutional Economics* 18 (1): 85–100.
- Aronow, Peter M., Jonathon Baron, and Lauren Pinson. 2019. "A Note on Dropping Experimental Subjects Who Fail a Manipulation Check." *Political Analysis* 27 (4): 572–589. <https://doi.org/10.1017/pan.2019.5>.
- Auspurg, Katrin, and Thomas Hinz. 2015. *Factorial Survey Experiments*. Quantitative Applications in the Social Sciences, Vol. 175. Thousand Oaks, CA: SAGE Publications. <https://doi.org/10.4135/9781483398075>.
- Bagozzi, Richard P. 1992. "The Self-Regulation of Attitudes, Intentions, and Behavior." *Social Psychology Quarterly* 55 (2): 178–204. <https://doi.org/10.2307/2786945>.
- Barbabela, Letícia. 2025. "Judicial Inconsistency and Citizen Anti-Corruption Demobilization: Evidence from Brazil." *Government and Opposition* 60 (1): 168–187. <https://doi.org/10.1017/gov.2023.36>.
- Barbabela, Letícia, Miquel Pellicer, and Eva Wegner. 2022. "Court Performance and Citizen Attitudes toward Fighting Corruption." *Governance* 35 (3): 717–735. <https://doi.org/10.1111/gove.12604>.
- Barr, Abigail, and Danila Serra. 2009. "The Effects of Externalities and Framing on Bribery in a Petty Corruption Experiment." *Experimental Economics* 12 (4): 488–503.
- Bauhr, Monika. 2012. "Need or Greed Corruption?" In *Good Government*. Cheltenham, UK: Edward Elgar Publishing.
- Bauhr, Monika. 2017. "Need or Greed? Conditions for Collective Action against Corruption." *Governance* 30 (4): 561–581. <https://doi.org/10.1111/gove.12232>.
- Bauhr, Monika, and Marcia Grimes. 2014. "Indignation or Resignation: The Implications of Transparency for Societal Accountability." *Governance* 27 (2): 291–320. <https://doi.org/10.1111/gove.12033>.

- Berinsky, Adam J., Michele F. Margolis, and Michael W. Sances. 2014. "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys." *American Journal of Political Science* 58 (3): 739–753. <https://doi.org/10.1111/ajps.12081>.
- Bhattacharjee, Anol. 2012. *Social Science Research: Principles, Methods, and Practices*. 2nd ed. Tampa, FL: University of South Florida, Tampa Bay Open Access Textbooks Collection. https://digitalcommons.usf.edu/oa_textbooks/3.
- Braun, Virginia, and Victoria Clarke. 2022. *Thematic Analysis: A Practical Guide*. Thousand Oaks, CA: SAGE Publications.
- Carothers, Christopher. 2022. "Taking Authoritarian Anti-Corruption Reform Seriously." *Perspectives on Politics* 20 (1): 69–85. <https://doi.org/10.1017/S1537592720001371>.
- Carson, Lindsey D., and Mariana Mota Prado. 2016. "Using Institutional Multiplicity to Address Corruption as a Collective Action Problem: Lessons from the Brazilian Case." *The Quarterly Review of Economics and Finance* 62: 56–65.
- Coleman, David E., and Douglas C. Montgomery. 1993. "A Systematic Approach to Planning for a Designed Industrial Experiment." *Technometrics* 35 (1): 1–12. <https://doi.org/10.1080/00401706.1993.10484984>.
- Coppock, Alexander. 2019. "Avoiding Post-Treatment Bias in Audit Experiments." *Journal of Experimental Political Science* 6 (1): 1–4. <https://doi.org/10.1017/XPS.2018.9>.
- Corbacho, Ana, Daniel W. Gingerich, Virginia Oliveros, and Mauricio Ruiz-Vega. 2016. "Corruption as a Self-Fulfilling Prophecy: Evidence from a Survey Experiment in Costa Rica." *American Journal of Political Science* 60 (4): 1077–1092. <https://doi.org/10.1111/ajps.12244>.
- Culiberg, Barbara, and Katarina Katja Mihelič. 2017. "The Evolution of Whistleblowing Studies: A Critical Review and Research Agenda." *Journal of Business Ethics* 146 (4): 787–803.
- Curini, Luigi, and Paolo Martelli. 2015. "A Case of Valence Competition in Elections: Parties' Emphasis on Corruption in Electoral Manifestos." *Party Politics* 21 (5): 686–698. <https://doi.org/10.1177/1354068813491540>.
- d'Urso, Amanda Sahar, Tabitha Bonilla, and Gabriela Bogdanowicz. 2025. "Where to Place Sensitive Questions? Experiments on Survey Response Order and Measures of Discriminatory Attitudes." *Political Science Research and Methods*: 1–22.
- Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia, eds. 2011. *Cambridge Handbook of Experimental Political Science*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511921452>.
- Druckman, James N., and Cindy D. Kam. 2011. "Students as Experimental Participants: A Defense of the 'Narrow Data Base.'" In *Cambridge Handbook of Experimental Political Science*, ed. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 41–57. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511921452>.

- European Bank for Reconstruction and Development. 2016. *Life in Transition: A Decade of Measuring Transition (LiTS III)*. London: EBRD. <https://litsonline-ebrd.com/>.
- Givati, Yehonatan. 2016. "A Theory of Whistleblower Rewards." *The Journal of Legal Studies* 45 (1): 43–72. <https://doi.org/10.1086/684617>.
- Gong, Ting, and Hanyu Xiao. 2017. "Socially Embedded Anti-Corruption Governance: Evidence from Hong Kong." *Public Administration and Development* 37 (3): 176–190. <https://doi.org/10.1002/pad.1798>.
- Heyes, Anthony, and Sandeep Kapur. 2009. "An Economic Model of Whistleblower Policy." *Journal of Law, Economics, and Organization* 25 (1): 157–182.
- Heywood, Paul. 1997. "Political Corruption: Problems and Perspectives." *Political Studies* 45 (3): 417–439.
- Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15 (4): 309–334. <https://doi.org/10.1037/a0020761>.
- Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105 (4): 765–789. <https://doi.org/10.1017/S0003055411000414>.
- Jaunky, Vishal Chandr, Jamiil Jeetoo, Caitanyadass Bajah, and Vani Ramesh. 2020. "The Importance of Understanding the Anti-Corruption Legislation to Promote Corruption Reporting: Lessons from Mauritius." *International Journal of Public Administration* 43 (15): 1282–1292. <https://doi.org/10.1080/01900692.2019.1669049>.
- Johnson, Ronald. 2004. *The Struggle Against Corruption: A Comparative Study*. New York: Springer.
- Kane, John V., and Jason Barabas. 2019. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments." *American Journal of Political Science* 63 (1): 234–249. <https://doi.org/10.1111/ajps.12396>.
- Klar, Samara, Thomas J. Leeper, and Joshua Robison. 2020. "Studying Identities with Experiments: Weighing the Risk of Post-Treatment Bias against Priming Effects." *Journal of Experimental Political Science* 7 (1): 56–60. <https://doi.org/10.1017/XPS.2019.26>.
- Klitgaard, Robert. 1988. *Controlling Corruption*. Berkeley: University of California Press.
- Ko, Kilkon, and Cuifen Weng. 2012. "Structural Changes in Chinese Corruption." *The China Quarterly* 211: 718–740. <https://doi.org/10.1017/S0305741012000793>.

- Köbis, Nils C., Marleen Troost, Cyril O. Brandt, and Ivan Soraperra. 2022. "Social Norms of Corruption in the Field: Social Nudges on Posters Can Help to Reduce Bribery." *Behavioural Public Policy* 6 (4): 597–624.
- Köbis, Nils C., Jan-Willem van Prooijen, Francesca Righetti, and Paul A. M. Van Lange. 2015. "'Who Doesn't?' — The Impact of Descriptive Norms on Corruption." *PLoS ONE* 10 (6): e0131830. <https://doi.org/10.1371/journal.pone.0131830>.
- Lagos, Nicolás, and Yahong Zhang. 2025. "Reporting in a Systemic Corruption Context: Voices of Citizen Witnesses." *Public Integrity*. <https://doi.org/10.1080/10999922.2025.2472925>.
- Lavena, Cecilia Florencia. 2016. "Whistle-Blowing: Individual and Organizational Determinants of the Decision to Report Wrongdoing in the Federal Government." *The American Review of Public Administration* 46 (1): 113–136.
- Li, Zeren, and Melanie Manion. 2023. "The Decline of Factions: The Impact of a Broad Purge on Political Decision Making in China." *British Journal of Political Science* 53 (3): 815–834. <https://doi.org/10.1017/S000712342200062X>.
- Magalhães, Pedro C. 2024. "When Corruption Investigations Come to Nothing: A Natural Experiment on Trust in Courts." *Governance* 37 (1): 99–117. <https://doi.org/10.1111/gove.12754>.
- Melgar, Natalia, Máximo Rossi, and Tom W. Smith. 2010. "The Perception of Corruption." *International Journal of Public Opinion Research* 22 (1): 120–131.
- Mesmer-Magnus, Jessica R., and Chockalingam Viswesvaran. 2005. "Whistleblowing in Organizations: An Examination of Correlates of Whistleblowing Intentions, Actions, and Retaliation." *Journal of Business Ethics* 62 (3): 277–297. <https://doi.org/10.1007/s10551-005-0849-1>.
- Miceli, Marcia P., and Janet P. Near. 1985. "Characteristics of Organizational Climate and Perceived Wrongdoing Associated with Whistle-Blowing Decisions." *Personnel Psychology* 38 (3): 525–544. <https://doi.org/10.1111/j.1744-6570.1985.tb00558.x>.
- Miceli, Marcia P., Janet P. Near, and Terry Morehead Dworkin. 2008. *Whistle-Blowing in Organizations*. New York: Routledge. <https://doi.org/10.4324/9780203809495>.
- Miller, Leona E., and James E. Grush. 1988. "Improving Predictions in Expectancy Theory Research: Effects of Personality, Expectancies, and Norms." *Academy of Management Journal* 31 (1): 107–122.
- Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." *American Journal of Political Science* 62 (3): 760–775. <https://doi.org/10.1111/ajps.12357>.
- Mungiu-Pippidi, Alina. 2006. "Corruption: Diagnosis and Treatment." *Journal of Democracy* 17 (3): 86–99.

- Mungiu-Pippidi, Alina. 2015. "Understanding Control of Corruption." In *The Quest for Good Governance: How Societies Develop Control of Corruption*, 1–26. Cambridge: Cambridge University Press.
- Mutz, Diana C. 2011. *Population-Based Survey Experiments*. Princeton, NJ: Princeton University Press.
- Near, Janet P., and Marcia P. Miceli. 1995. "Effective Whistle-Blowing." *Academy of Management Review* 20 (3): 679–708. <https://doi.org/10.2307/258791>.
- Oboro-Offerie, Ruby Amanda, Sewordor Toklo, and Precious Wedaga Allor. 2025. "Why Are Citizens Afraid to Blow the Whistle on Corruption? Understanding the Impact of Personal Bribery Experiences on Corruption Reporting in Africa." *Journal of Asian and African Studies*. <https://doi.org/10.1177/00219096251376663>.
- Oppenheimer, Daniel M., Tom Meyvis, and Nicolas Davidenko. 2009. "Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power." *Journal of Experimental Social Psychology* 45 (4): 867–872. <https://doi.org/10.1016/j.jesp.2009.03.009>.
- Ostrom, Elinor. 1998. "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997." *American Political Science Review* 92 (1): 1–22. <https://www.jstor.org/stable/2585925>.
- Pavão, Nara. 2019. "Corruption, Courts, and Public Opinion in Brazil." In *Corruption in Latin America: How Politicians and Corporations Steal from Citizens*, ed. Robert I. Rotberg, 93–107. Cham: Springer.
- Peiffer, Caryn, and Linda Alvarez. 2016. "Who Will Be the 'Principled-Principals'? Perceptions of Corruption and Willingness to Engage in Anticorruption Activism." *Governance* 29 (3): 351–369.
- Peiffer, Caryn, and Grant W. Walton. 2022. "Getting the (Right) Message across: How to Encourage Citizens to Report Corruption." *Development Policy Review* 40 (5): e12621. <https://doi.org/10.1111/dpr.12621>.
- Persson, Anna, Bo Rothstein, and Jan Teorell. 2013. "Why Anticorruption Reforms Fail: Systemic Corruption as a Collective Action Problem." *Governance* 26 (3): 449–471.
- Rose-Ackerman, Susan. 1978. *Corruption: A Study in Political Economy*. New York: Academic Press.
- Rose-Ackerman, Susan. 1999. "Bribes, Patronage, and Gift Giving." In *Corruption and Government: Causes, Consequences, and Reform*, 91–110. Cambridge: Cambridge University Press.
- Rothschild, Joyce. 2013. "The Fate of Whistleblowers in Nonprofit Organizations." *Nonprofit and Voluntary Sector Quarterly* 42 (5): 886–901.
- Rothstein, Bo, and Jan Teorell. 2008. "What Is Quality of Government? A Theory of Impartial Government Institutions." *Governance* 21 (2): 165–190.

- Su, Su. 2020. "Why So Few Acting Whistleblowers? Impacts of Institutional Anticorruption." *Governance* 33 (2): 227–247.
- Su, Xuhong, and Xing Ni. 2018. "Citizens on Patrol: Understanding Public Whistleblowing against Government Corruption." *Journal of Public Administration Research and Theory* 28 (3): 406–422.
- Sun, Zongfeng, Lin Zhu, and Xing Ni. 2022. "How Does Anti-Corruption Information Affect Public Perceptions of Corruption in China?" *China Review* 22 (2): 113–143.
- Transparency International. n.d. "What Is Corruption?" Accessed February 10, 2025. <https://www.transparency.org/en/what-is-corruption>.
- Treisman, Daniel. 2007. "What Have We Learned About the Causes of Corruption from Ten Years of Cross-National Empirical Research?" *Annual Review of Political Science* 10: 211–244. <https://doi.org/10.1146/annurev.polisci.10.081205.095418>.
- Tsai, Lily L., Minh Trinh, and Shiyao Liu. 2022. "What Makes Anticorruption Punishment Popular? Individual-Level Evidence from China." *The Journal of Politics* 84 (1): 602–606.
- Tverdova, Yuliya V. 2011. "See No Evil: Heterogeneity in Public Perceptions of Corruption." *Canadian Journal of Political Science/Revue canadienne de science politique* 44 (1): 1–25.
- Uslaner, Eric M. 2008. *Corruption, Inequality, and the Rule of Law*. Cambridge: Cambridge University Press.
- World Justice Project. 2020. *Rule of Law Index 2020*. Washington, DC: World Justice Project.
- Zhang, Julia, Randy Chiu, and Liqun Wei. 2009. "Decision-Making Process of Internal Whistleblowing Behavior in China: Empirical Evidence and Implications." *Journal of Business Ethics* 88 (Suppl. 1): 25–41.
- Zhang, Yahong, and Min-Hyu Kim. 2018. "Do Public Corruption Convictions Influence Citizens' Trust in Government? The Answer Might Not Be a Simple Yes or No." *The American Review of Public Administration* 48 (7): 685–698.
- Zhang, Yahong, Wen-Jhan Hung, Hyun-Kyu Dong, and Don-Yun Chen. 2025. "Public Employees' Whistleblowing Intention: Explanation by an Adapted Theory of Planned Behavior." *Public Administration* 103 (1): 115–135. <https://doi.org/10.1111/padm.13009>.
- Zhu, Jiangnan, and Dong Zhang. 2017. "Weapons of the Powerful: Authoritarian Elite Competition and Politicized Anticorruption in China." *Comparative Political Studies* 50 (9): 1186–1220. <https://doi.org/10.1177/0010414016672234>.

Appendix A. Survey Instrument

This appendix presents the full text of the survey instrument administered to participants through the Qualtrics platform between 20 March and 5 April 2026. The survey was approved by the Nazarbayev University School of Sciences and Humanities Institutional Research Ethics Committee (SSH IREC). All questions are reproduced as they appeared to respondents. Branching logic and randomization rules are noted where applicable.

A.1. Informed Consent

Q1. Do you agree to participate?

☐ Yes

☐ No

Routing: Respondents selecting "No" exited the survey.

A.2. Eligibility Screening

Q2. Are you 18 years or older?

☐ Yes

☐ No

Q3. Are you a citizen of Kazakhstan?

☐ Yes

☐ No

Routing: Respondents answering "No" to either screening question were excluded from the study.

A.3. Demographic Background

Instruction: Before proceeding, we would like to ask a few brief questions about your background. Your responses will remain anonymous.

Q4. What is your age? (numbers only): ____

Q5. What is your gender?

☐ Male

☐ Female

Q6. What is your current level of study?

☐ Bachelor's degree

☐ Master's degree

☐ PhD student

Q7. What year of study are you currently in?

- ☐ 1st year
- ☐ 2nd year
- ☐ 3rd year
- ☐ 4th year
- ☐ 5th year or higher

Q8. Which school are you currently studying in?

- ☐ School of Sciences and Humanities (SSH)
- ☐ School of Engineering and Digital Sciences (SEDS)
- ☐ School of Medicine (NUSOM)
- ☐ School of Mining and Geosciences (SMG)
- ☐ Graduate School of Business (GSB)
- ☐ Graduate School of Public Policy (GSPP)
- ☐ Graduate School of Education (GSE)
- ☐ Prefer not to answer

Q9. What is your household income?

- ☐ Up to 300,000 KZT
- ☐ 300,000 – 500,000 KZT
- ☐ 500,000 – 700,000 KZT
- ☐ 700,000 – 900,000 KZT
- ☐ 900,000 KZT or more
- ☐ Prefer not to answer

Q10. Before the age of 18, did you primarily live in:

- ☐ A city or town
- ☐ A rural area or village
- ☐ Prefer not to answer

Q11. In which region of Kazakhstan did you primarily live before the age of 18 (i.e., where you spent most of your childhood and adolescence)?

[Drop-down list of all regions of Kazakhstan]

A.4. Pre-Treatment Predispositions

Instruction: Next, we would like to ask a few questions about your views and experiences related to corruption. Please note that your responses will remain anonymous.

Corruption Tolerance

Instruction: Please indicate how acceptable you find the following behavior. Response scale: 1 = Completely unacceptable, 2 = Somewhat unacceptable, 3 = Neither acceptable nor unacceptable, 4 = Somewhat acceptable, 5 = Completely acceptable.

Q12. A public official uses their position to help a relative get a job.

Q13. A person gives a gift to a public employee as a thank-you after receiving a service.

Q14. A public official accepts a dinner invitation from a business that they oversee or supervise.

Q15. A person gives extra money to a public employee to receive a public service faster (e.g., to obtain a medical certificate more quickly).

Note: An index of corruption tolerance was constructed by averaging Q12–Q15.

Perceived Corruption Prevalence

Q16. In your opinion, how common is corruption among public officials in Kazakhstan?

- ☐ 1 = Very rare
- ☐ 2 = Rare
- ☐ 3 = Neither rare nor common
- ☐ 4 = Common
- ☐ 5 = Very common

Personal Experience and Familiarity

Q17. In the past 12 months, have you been asked to make an unofficial payment or give a gift in exchange for public services?

- ☐ Yes
- ☐ No
- ☐ Prefer not to answer

Q18. Before participating in this survey, how familiar were you with the official procedures for reporting corruption (for example, where and how to submit a report)?

- ☐ 1 = Not at all familiar
- ☐ 2 = Slightly familiar
- ☐ 3 = Moderately familiar
- ☐ 4 = Very familiar
- ☐ 5 = Extremely familiar

A.5. Attention Check

Q19. [Embedded attention-check item placed before the experimental treatment. Respondents were instructed to select "Somewhat agree" among Likert-style response options.]

A.6. Experimental Treatment

Following the attention check, respondents were randomly assigned to one of four experimental conditions using Qualtrics' built-in randomization tool. Each respondent saw only one vignette.

Common preamble (shown to all conditions):

Before answering the following questions, please carefully read the information below describing how government authorities handled corruption cases.

In Kazakhstan, corruption by public officials is prohibited by law and may result in legal penalties. Citizens who encounter suspected corruption can submit reports through official channels operated by the Anti-Corruption Agency. The agency is responsible for receiving complaints, reviewing submitted information, and deciding whether to take further action in accordance with procedures established by law.

Condition 1: Consistent and Impartial (Best Signal)

Since the government introduced a new anti-corruption system, citizens have submitted an average of 128.5 reports per month. The government met its initial target and has punished more than 80% of all registered cases. Approximately equal proportions of high-ranking and low-ranking officials were punished.

Condition 2: Consistent and Partial

Since the government introduced a new anti-corruption system, citizens have submitted an average of 128.5 reports per month. The government met its initial target and has punished more than 80% of all registered cases. A larger share of sanctions was imposed on lower-ranking officials, while high-ranking officials were rarely sanctioned.

Condition 3: Inconsistent and Impartial

Since the government introduced a new anti-corruption system, citizens have submitted an average of 128.5 reports per month. The government did not meet its initial target and has punished only 15% of all registered cases. Approximately equal proportions of high-ranking and low-ranking officials were punished.

Condition 4: Inconsistent and Partial (Baseline / Control)

Since the government introduced a new anti-corruption system, citizens have submitted an average of 128.5 reports per month. The government did not meet its initial target and has punished only 15% of all registered cases. A larger share of sanctions was imposed on lower-ranking officials, while high-ranking officials were rarely sanctioned.

A.7. Outcome Measures

Primary Outcome: Perceived Sincerity (Mediator)

Q20. To what extent do you believe the government is sincere in its commitment to combating corruption?

- ☐ 1 = Not sincere at all — efforts are purely symbolic
- ☐ 2 = Slightly sincere — commitment is mostly symbolic
- ☐ 3 = Moderately sincere — some genuine commitment
- ☐ 4 = Very sincere — largely genuine commitment
- ☐ 5 = Extremely sincere — fully genuine and consistent commitment

Primary Outcome: Reporting Intention

Q21. Imagine that you discover that a local public service employee is asking citizens for unofficial payments in exchange for faster processing of documents. If you had reliable evidence of this misconduct, how likely would you be to report it to the relevant authorities?

- ☐ 1 = Extremely unlikely
- ☐ 2 = Somewhat unlikely
- ☐ 3 = Neither likely nor unlikely
- ☐ 4 = Somewhat likely
- ☐ 5 = Extremely likely

A.8. Manipulation Checks

Q22. To what extent do you agree with the following statement: "The authorities responded consistently to corruption cases."

- ☐ 1 = Strongly disagree
- ☐ 2 = Somewhat disagree
- ☐ 3 = Neither agree nor disagree
- ☐ 4 = Somewhat agree
- ☐ 5 = Strongly agree

Q23. To what extent do you agree with the following statement: "The authorities applied punishment equally to high-ranking and low-ranking officials."

- ☐ 1 = Strongly disagree
- ☐ 2 = Somewhat disagree
- ☐ 3 = Neither agree nor disagree
- ☐ 4 = Somewhat agree
- ☐ 5 = Strongly agree

A.9. Post-Treatment Reporting Experience

Q24. Have you ever reported a case of corruption to the authorities?

- ☐ Yes
- ☐ No
- ☐ Prefer not to answer

A.10. Invitation to Follow-Up Interview

Routing: Displayed only to respondents who answered "Yes" to Q17 (corruption experience) or Q24 (prior reporting).

We invite participants who have prior experience reporting corruption or attempting to report corruption to participate in a confidential follow-up interview (30–60 minutes).

Q25. If you are willing to share your experience in more detail, you may leave your email address below. (Optional)

A.11. Debriefing Statement

A.12. End of Survey

Appendix B. Interview Protocol

This appendix presents the semi-structured interview protocols used in the qualitative component of this study. Nine interviews were conducted between 1 April and 15 April 2026 with Kazakhstani citizens recruited through the experimental survey (see Appendix A.10) and through purposive sampling. Eight interviewees had encountered corruption but did not report it; one interviewee had submitted a formal report. Two distinct protocols were used to reflect these different experiences. All interviews were audio-recorded with informed consent, transcribed, and thematically analyzed. Pseudonyms (P1–P9) are used throughout the thesis to protect respondent identity. The interviews lasted between 30 and 60 minutes.

B.1. Introductory Script and Informed Consent

The following script was read to all participants prior to the interview.

Thank you for participating in this study. My name is Tomiris Kabayeva, and I am researching how citizens perceive government anti-corruption efforts and how those perceptions shape the decision to report corruption.

Before we begin, I would like to ensure you are comfortable with the process:

Voluntary participation: You may skip any question or stop the interview at any time without consequence.

Confidentiality: Your identity will remain strictly anonymous. A pseudonym will be used in all written reports and academic publications.

Data storage: With your permission, I will record this session for transcription purposes only. The audio file will be stored on a password-protected device and destroyed upon completion of the study.

Goal of the conversation: I am interested in your honest personal reflections. There are no right or wrong answers.

Do you have any questions before we begin? Do I have your consent to start the interview and to record our conversation?

B.2. Protocol 1 — For Participants Who Encountered Corruption but Did Not Report

This protocol was used with eight participants (P1–P8 excluding the one reporter) who had observed or experienced corruption but did not file a formal report through official channels.

Part A. Initial Perceptions and Expectations Prior to the Encounter

1. What was your general view of the government's anti-corruption efforts at the moment when you encountered corruption directly?
2. If you had decided to report, to what extent did you believe the authorities would actually take up your case and act on it? What shaped that belief — or that doubt?

3. How important was it for you that punishment would be applied equally — to a high-ranking official as well as to an ordinary person? Did this consideration influence your decision not to report?

Part B. The Decision Not to Report

4. Thinking back to the moment when you encountered corruption, what specifically stopped you from reporting it?
5. Did you have any hesitations or moments of doubt? Was there a point at which you almost decided to report — and what ultimately held you back?
6. Were you aware that you could submit a report through the chatbot or the hotline? If you knew about these channels, why did you not use them?

Part C. Perceptions of Institutional Sincerity

7. How do you distinguish a genuine fight against corruption from a performative one — when actions are taken only for appearances?
8. At the moment you were deciding whether to report, how strongly did you believe that the government was sincerely committed to fighting corruption? What influenced this perception — news reports, personal experiences, stories from people you know?
9. Did you consider that the outcome of a case might depend on who was involved — an ordinary person versus an influential official? Did this consideration affect your decision?

Part D. Looking Back

10. Looking back now, how do you evaluate your decision not to report? Would you act the same way again?
11. How has this experience affected your overall trust in the authorities — has it increased or decreased?
12. If you encountered corruption again tomorrow, would you report it? What would need to change for you to be willing to do so?

B.3. Protocol 2 — For Participants Who Had Reported Corruption

This protocol was used with one participant (P1) who had submitted a formal corruption report through official channels.

Part A. Initial Perceptions and System Trust (Pre-Reporting)

1. How would you describe your general view of the government's anti-corruption efforts before you decided to submit a report?
2. Before reporting, to what extent did you expect authorities to follow through consistently — that is, to process the case and apply sanctions in a predictable and regular manner? What shaped that expectation?
3. How important was your belief that sanctions would be applied equally regardless of the rank or status of the person involved? Did perceptions of possible selectivity influence your decision?

Part B. The Decision-Making Process

4. Reflecting on the moment you decided to report, what was the primary factor that convinced you it was the right step to take?

5. What specific concerns or doubts, if any, did you weigh before finalizing your decision to come forward?
6. You chose to use [specific channel — e.g., chatbot, hotline, e-Otinish portal, in-person submission]. Why did you select this particular method? Did you perceive it as more reliable or secure than other available options?

Part C. Perceived Sincerity of Institutional Commitment

7. From your perspective, what is the main difference between a genuine effort to fight corruption and one that is merely symbolic or "for show"?
8. At the time you were deciding whether to report, how sincere did you believe the government's commitment to fighting corruption was? What influenced that perception — for example, information in the media, news reports, online discussions, personal experiences, or stories from people you know?
9. In your decision-making process, did expectations about how authorities might treat individuals of different status influence you in any way?

Part D. Experience and Reassessment (Post-Reporting)

10. Could you describe the nature of the feedback or response you received after submitting your report? For example, was it a specific update on the investigation, or a more general acknowledgment?
11. Looking back, how has this specific experience influenced your overall trust in the authorities' commitment to justice?
12. If you were to encounter a similar situation in the future, would you choose to report it again? Why or why not?

B.4. Closing Questions (Used in Both Protocols)

Is there anything else you would like to share that we have not yet discussed?

Thank you again for your time and for your valuable contribution to this research.

B.4. Notes on Interview Procedure

Interviews were conducted in Russian or Kazakh according to the participant's preference. Direct quotations used in Chapter 5 were translated into English by the author. Transcripts were analyzed using inductive thematic coding (Braun and Clarke 2006).

B.6. Profile of Interview Participants

Table B 1. Profile of Interview Participants

Participant ID	Group	Age group	Gender	Place of residence	Employment sector	Channel of reporting
P1	Reporter	18-29	Male	Semey	Private sector	E-otinish, Department of Economic Investigations (DEI), Internal State Audit Committee (ISAC)
P2	Non-reporter	18-29	Male	Astana	Private sector	-
P3	Non-reporter	18-29	Female	Astana	Unemployed	-
P4	Non-reporter	18-29	Female	Astana	Unemployed	-
P5	Non-reporter	40-49	Female	Almaty	Government Sector	-
P6	Non-reporter	50+	Female	Pavlodar	Unemployed	-
P7	Non-reporter	18-29	Male	Zhezkazgan	Private sector	-
P8	Non-reporter	30-39	Female	Astana	Private sector	-
P9	Non-reporter	18-29	Female	Aktau		-

Note. Only gender, place of residence, age group, and employment sector were collected to protect participants' anonymity and reduce social desirability bias.

Appendix C. Supplementary Statistical Tables

This appendix presents additional statistical analyses supporting the main findings reported in Chapter 5.

Table C 1. Sample Characteristics by Treatment Condition

Characteristic	Total (N = 229)	Inconsistent & Partial (n = 56)	Inconsistent & Impartial (n = 61)	Consistent & Partial (n = 53)	Consistent & Impartial (n = 59)
Age, mean (SD)	22.1 (4.5)	22.0 (5.3)	21.5 (3.1)	22.0 (4.0)	22.8 (5.1)
Female (%)	59.0	58.9	60.7	64.2	52.5
Urban (%)	92.5	83.9	96.7	96.2	93.2
Large cities (%)	40.6	46.4	42.6	28.3	44.1
Bachelor's level (%)	72.1	75.0	73.8	75.5	64.4
Year 1 (%)	34.1	33.9	29.5	22.6	49.2
Corruption tolerance, mean (SD)	2.37 (0.84)	2.42 (0.94)	2.47 (0.82)	2.39 (0.88)	2.21 (0.72)
Corruption prevalence, mean (SD)	4.44 (0.71)	—	—	—	—
Personal corruption experience (%)	13.7	14.8	13.3	20.8	7.0
Unfamiliar with reporting (%)	38.0	—	—	—	—

Notes. Treatment conditions ordered from worst (inconsistent and partial) to best (consistent and impartial). Age and corruption tolerance reported as mean (SD). All other variables reported as percentages. Corruption prevalence and reporting familiarity are reported for the total sample only.

Table C 2. Randomization Balance Check

Variable	Inconsistent & Partial (Baseline)	Inconsistent & Impartial	Consistent & Partial	Consistent & Impartial	p-value
Age (mean)	22.00	21.51	22.00	22.76	0.489
Corruption tolerance (mean)	2.42	2.47	2.39	2.21	0.370
Female (%)	58.9	60.7	64.2	52.5	0.644
Urban (%)	84.0	96.7	96.2	93.2	0.035
Large cities (%)	46.4	42.6	28.3	44.1	0.211
Corruption experience (%)	14.8	13.3	20.8	7.0	0.233
Reporting familiarity (mean)	2.18	2.43	2.32	2.26	0.731

Notes. p-values from one-way ANOVA (continuous variables) and chi-square tests (categorical variables). Urban variable shows marginal imbalance ($p = .035$), likely due to small cell sizes. All pairwise Bonferroni-corrected t-tests confirm no significant differences between any pair of groups on age or corruption tolerance (all $p > .05$).

Table C 3. Mean Reporting Intention and Perceived Sincerity by Treatment Condition

Treatment condition	N	Reporting intention M (SD)	Perceived sincerity M (SD)
Inconsistent & Partial (Control)	56	3.32 (1.16)	2.23 (0.81)
Inconsistent & Impartial	61	2.95 (1.20)	1.98 (0.89)
Consistent & Partial	53	3.08 (1.05)	2.15 (0.79)
Consistent & Impartial	59	3.05 (1.15)	2.32 (0.96)
Total	229	3.09 (1.14)	2.17 (0.87)

Note. Reporting intention and perceived sincerity were measured on a 1–5 scale.

Table C 4. Robustness Checks: Treatment Effects among Attention Check Passers

Variable	(1) Reporting intention	(2) Reporting intention + controls	(3) Perceived sincerity	(4) Perceived sincerity + controls
CONSISTENT	−0.055 (0.153)	−0.089 (0.153)	+0.109 (0.118)	+0.137 (0.108)
IMPARTIAL	−0.167 (0.153)	−0.187 (0.152)	−0.027 (0.118)	−0.017 (0.108)
Corruption tolerance	—	−0.258** (0.093)	—	0.143* (0.066)
Corruption prevalence	—	0.081 (0.106)	—	−0.496*** (0.075)
Gender	—	0.044 (0.155)	—	0.018 (0.110)
Age	—	−0.001 (0.017)	—	0.004 (0.012)
Constant	3.228*** (0.134)	3.512*** (0.682)	2.111*** (0.103)	3.864*** (0.485)
R ²	0.006	0.043	0.004	0.184
F-statistic	0.663 (p = .517)	1.612 (p = .145)	0.455 (p = .635)	8.029*** (p < .001)
N	221	221	221	221

Notes. Standard errors are reported in parentheses. Sample restricted to respondents who passed the attention check. † p < .10; * p < .05; ** p < .01; *** p < .001.

Table C 5. Per-Protocol Analysis

Variable	(1) Reporting intention	(2) Reporting intention + controls	(3) Perceived sincerity	(4) Perceived sincerity + controls
CONSISTENT	0.118 (0.266)	0.446** (0.154)	0.084 (0.258)	0.463** (0.154)
IMPARTIAL	−0.400 (0.265)	0.448** (0.153)	−0.195 (0.265)	0.379* (0.158)
Corruption tolerance	—	—	−0.487** (0.168)	0.092 (0.100)
Corruption prevalence	—	—	0.341† (0.174)	−0.250* (0.104)
Gender	—	—	0.246 (0.248)	−0.028 (0.147)
Age	—	—	−0.036 (0.033)	0.000 (0.019)
Constant	3.282*** (0.179)	2.043*** (0.104)	3.521** (1.141)	2.957*** (0.678)
R ²	0.025	0.208	0.136	0.263
F-statistic	1.151 (p = .321)	11.72*** (p < .001)	2.227† (p = .048)	5.05*** (p < .001)
N	92	92	92	92

Notes. Standard errors in parentheses. Baseline = inconsistent and partial enforcement condition (CONSISTENT = 0, IMPARTIAL = 0). Per-protocol sample (n = 92, 40.2% of N = 229) includes respondents whose manipulation check scores fell in the expected direction relative to the median of the opposing condition group on both consistency and impartiality check items. Models (1) and (3) include treatment variables only. Models (2) and (4) add corruption tolerance, corruption prevalence, gender, and age as controls. † p < .10; * p < .05; ** p < .01; *** p < .001.

Table C 6. OLS Regression — Subgroup Analysis by Corruption Tolerance

Variable	(1) Reporting Intention	(2) Reporting Intention	(3) Perceived Sincerity	(4) Perceived Sincerity
	Low tolerance	High tolerance	Low tolerance	High tolerance
CONSISTENT	−0.087 (0.223)	−0.144 (0.210)	+0.186 (0.143)	+0.136 (0.160)
IMPARTIAL	−0.231 (0.223)	−0.137 (0.205)	−0.091 (0.143)	+0.008 (0.157)
Corruption prevalence	+0.142 (0.150)	−0.049 (0.149)	−0.665*** (0.096)	−0.271* (0.114)
Gender	−0.096 (0.226)	+0.189 (0.216)	−0.060 (0.145)	+0.079 (0.165)
Age	−0.000 (0.023)	+0.002 (0.027)	+0.002 (0.015)	+0.005 (0.021)
Constant	2.877** (0.875)	3.109*** (0.917)	4.972*** (0.561)	3.279*** (0.700)
R ²	0.021	0.016	0.295	0.069
N	123	106	123	106

Notes. Unstandardized OLS coefficients with standard errors in parentheses. Baseline = Inconsistent & Unequal (CONSISTENT = 0, IMPARTIAL = 0). Subgroups defined by median split of corruption tolerance (Low $\leq$ median, n = 123; High $>$ median, n = 106). Corruption tolerance was used as grouping variable and excluded as covariate. †p < .10. *p < .05. **p < .01. ***p < .001.