BABEŞ-BOLYAI UNIVERSITY

FACULTY OF PSYCHOLOGY AND EDUCATIONAL SCIENCES

APPLIED COGNITIVE PSYCHOLOGY DOCTORAL SCHOOL

Dark Personality and Interpersonal Deception in Prison:

A Qualitative and Quantitative Perspective

SUMMARY of the PH.D. THESIS

AUTHOR: Ph.D. candidate Andreea Turi

SCIENTIFIC ADVISOR: Prof. dr. habil. Laura Visu-Petra

CLUJ NAPOCA

2025

Table of Contents

Keywords: deception, attitudes towards deception, lying, dark personality, cues to deception, deception detection accuracy, DeceIT paradigm, prison, incarcerated individuals, qualitative research protocol, mixed methods approach.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my scientific supervisor, Prof. dr. Laura Visu-Petra for her significant contribution to this thesis. I express my deepest gratitude for her essential contributions to the theoretical and methodological framework, her extensive revisions, and her support in writing, analyzing, interpreting, disseminating, and publishing our findings. Furthermore, I am thankful for her patience, professionalism, and the guidance she offered throughout my PhD studies.

Next, I would like to express my gratitude and appreciation to my doctoral guiding committee members, lect. dr. Lavinia Cheie, lect. dr. Eva Kallay, and Professor dr. Adrian Opre, for their crucial input and contributions that allowed the completion of my thesis. I also extend my gratitude to senior lecturer PhD. Mircea Zloteanu for his significant contributions to the methodological framework and data analysis, and dr. George Visu for his role in the development of the experimental method.

In addition, I am thankful to the Gherla Penitentiary's management for approving and supporting me with my research, as well as to my colleagues and superiors who have shown their availability and encouraged the scientific process.

I am grateful to my colleagues PhD-candidates Liliana Hurezan, Raluca Mădălina Rebeleș, and Daria Solescu, for their valuable contribution to data collection, writing, analysis, and manuscript elaboration. Also, I express my kind appreciation to all my colleagues at the RIDDLE Lab for their assistance and to the students who contributed to the data collection.

Lastly, I want to thank my parents for their encouragements, acceptance, and support. And most importantly, I am deeply grateful to my friends (and mentors) Paul Vlad Muscaliuc and Alexandra Văidăhăzan for their help, unwavering trust, and unconditional support in my development throughout all these years.

THEORETICAL FRAMEWORK

1.1. Research Relevance

Defined as a communication strategy, *deception* is an intentional act of cultivating in another person a belief or understanding that the deceiver believes to be false, with the goal of manipulating the perceptions of others or gaining personal benefit (e.g., Masip et al., 2004). Acts of deception can take various forms, including verbal statements, omissions, or actions designed to mislead or create a false impression. Since interpersonal communication and empathic understanding do not always involve acknowledging the individual's true intentions, *deceptive behavior* could be interpreted positively and considered an adaptive strategy for optimal functioning in the social context (Janović et al., 2003). This is because the use of false information is not in itself an anti-social action, but *the way people define deception* and the motive behind differentiates between the pro- or antisocial nature of the untruths omitted or fabricated for different reasons. For instance, if lies are told for personal gain, people tend to classify them as anti-social. If they are used for the purpose of benefiting the receiver, they are more inclined to consider them more socially and morally acceptable or ethical (e.g., Hayashi et al., 2014; Turi et al., 2020).

Within this line of research, increasing attention has been paid to the role of *dark personality traits*, such as Narcissism, Machiavellianism, Psychopathy, and Sadism, grouped together under the umbrella terms **Dark Triad** or **Dark Tetrad**, due to common associations with a wide range of socially exploitative, self-serving, and manipulative behaviors (Paulhus & Williams, 2002; Buckels et al., 2013). Individuals with high levels of these traits were more likely to endorse *favorable attitudes towards deception*, lie more frequently, and perceive it as acceptable to use it instrumentally (Forsyth et al. 2021). Extending the conceptualization of dark personality, recent research has proposed a unifying framework: **the Dark Factor of Personality** (D), which conceptualizes all dark traits as manifestations of a general underlying tendency to maximize one's personal goals at the expense of others, even when this involves socially or morally questionable behavior (Moshagen et al., 2018). Within this model, a specific theme called ***Deceitfulness*** was identified as a central component, encompassing the intentional use of dishonest, manipulative, or illicit means for personal gain. Crucially, this "*theme*" has been shown to differentiate individuals with a criminal record from those without (Hurezan et al., 2024), underlining its relevance in forensic and correctional settings. Despite the potential relevance of the Dark Factor model for explaining the individual predisposition toward rule-breaking and unethical conduct, it has primarily been studied within general population samples (Moshagen et al., 2020) and few studies (García-Fernández et al., 2024; Hurezan et al, 2024) have examined this framework in prison populations.

Nevertheless, the **methodologies used by previous studies** to examine deception production and detection performances (e.g., DeceIT paradigm, Wright et al., 2012) were designed for community members. To date, as far as we know, there is no available paradigm to study interpersonal deception in prison. Thus, examining the interplay between dark personality traits and deception (beliefs, attitudes, and performances) among prisoners, can provide valuable insights into the **mechanisms** behind the deceptive behavior, explaining their willingness to use deception as a cognitively and emotionally driven response.

1.2. The Prison Environment

The prison environment, through its stimuli, social structures, and internal rules, fosters a distinct operational framework that reinforces behaviors congruent with institutional life (Crewe, 2012). Frequent exposure to unethical behavior, negative emotions, and morally disengaged justifications increases the likelihood that individuals will develop antisocial attitudes, internalize, and normalize antisocial conduct and might can gradually erode moral standards (Shulman et al., 2011).

Within this context, righteousness, ethical behavior, honesty, or law obedience can be perceived as signs of weakness and may not be tolerated by aggressive individuals. Without clear and immediate incentives for honest behavior, or in the presence of more immediate benefits from violating norms, many choose not to comply with them and rather obey the unwritten "*inmate code*" (Trammel, 2012) or adopt a *criminal identity* as a source of self-worth. In this mental frame, unethical strategies such as aggression or deception may function as strategies to preserve internal coherence, defend personal beliefs and fragile self-esteem, as well as protect against anticipated emotional pain (e.g., Laws, 2016; Maruna & Butler, 2009, 2013).

Thus, the prison environment may not only incubate antisocial behavior, but also cultivate a culture of deception, reinforced by the inmate code, and accentuate the display of dark personality traits, which have been associated with a higher propensity to deceive (e.g., Forsyth et al., 2021).

1.3. Personality[1]

If we examine human personality through the lens of the *Life History Theory (LHT,* Figueredo et al., 2021*),* we understand "*traits*" as adaptive strategies shaped by environmental challenges (Wiebe, 2004). According to this theory, personality development is affected by life adversities and influenced by stress (Birkás et al., 2018) and individuals develop specific strategies to ensure their survival via acquiring resources. If they are required to adapt to a secure and predictive environment, they are likely to develop a preference for *slow life strategies*, associated with the ability to delay gratification, long-term orientations, and positive personality traits, such as cooperation, long-term orientations, empathy, cooperation, and genuine concern for the wellbeing of others, encapsulated by the Five-Factor Model (Costa & McCrae 1992). In contrast, the unpredictable and aversive environments, tend to foster a preference for *fast life history strategies* that favor risk-taking, immediate gratification, and short-term oriented behaviors (Kaplan & Gangestad, 2005). These strategies are often preferred by individuals with dark personality traits, such as those composed of Dark Triad or Dark Factor Models of personality (Paulhus & Williams, 2002; Moshagen et al., 2018). Nevertheless, human personality is not entirely bright or dark, nor good, or bad, but a blend of positive (bright) and negative (dark) tendencies that coexist and can be activated in specific environmental contexts to benefit individual's adaptation (Thielman et al., 2020).

[1] This sub-chapter (pages 3-6) contains parts translated in English language that were originally published in the book chapter: Turi, A., G., Hurezan, L., & Visu-Petra, L. (2025). Delicvența la adult. Personalitatea și comportamentul infracțional. În Toma R. A & Sava F. A, (Eds) *Fundamentele psihologiei judiciare*. (p. 167-188), Polirom, Iași.

1.3.1. The Five-Factor Model (Big5)

The **Five-Factor Model** (FFM), also known as the **Big Five** (Costa & McCrae, 1992), groups the most important characteristics humans need to possess to successfully adapt to the social landscape, into five factors: *Openness to Experiences* (O), *Conscientiousness* (C), *Extraversion* (E), *Agreeableness* (A) and *Neuroticism* (N), relatively stable across time and culture. Three of these dimensions, *Agreeableness* and *Conscientiousness,* and the opposite pole of *Neuroticism* (also known as Emotional Stability) are considered the core of the desirable or bright personality or **Bright Triad** (Museck & Grum, 2021). These factors can act as protective factors against involvement in criminal activities, given that they equip the individual with prerequisites that allow individuals to adapt to the norms of society. Compared to the individuals situated at the opposite pole, who present emotional instability, suspiciousness, lack of empathy, and have low interest in following social rules and norms, low motivation by maintaining desirable relationships with others and are more concerned with the pursuit of self-interest, people characterized by high Bright Triad traits are more cooperative, confident, obedient (higher level of *Agreeableness*) organized, disciplined, future-oriented (higher level of *Conscientiousness*, Jackson et al, 2017) and emotionally stable (low *Neuroticism*, Wilt et al., 2017). Previous research supported the relationship between these personality factors, offending behavior (Liu, 2009), and recidivism (Weibe, 2004), but not all studies identified significant differences between offenders and non-offenders (Ciurbea & Dina, 2022).

1.3.2. The Dark Triad and Dark Tetrad Models

Over two decades ago, Paulhus and Williams (2002) introduced the concept of the **Dark Triad** (DT), which includes *Machiavellianism, Narcissism,* and *Psychopathy*. These traits have been grouped together because they share *a common tendency* toward egocentricity, insensitivity, and manipulation (Jakobwitz & Egan, 2006) in addition to the specific characteristics of each. *Narcissism* represents the pursuit of egoistic admiration for one's own characteristics, grandiosity, and sense of entitlement (Raskin & Terry, 1988; Raskin & Hall, 1979). *Machiavellianism* refers to a duplicitous interpersonal style, disregard for morality, and excessive focus on self-interest and self-gain (Christie & Geis, 1970). *Psychopathy* reflects diminished empathy and remorse, impulsivity, insensitive affect, chaotic lifestyle, and criminal tendencies (Paulhus et al., 2018).

This constellation of aversive personality traits situates on the crossline between typical personality traits and pathological personality traits due to their manifestation on a spectrum ranging from adaptive to dysfunctional. At moderate levels, dark traits can confer advantages in competitive contexts, such as business or politics, where they can facilitate leadership, effective strategy, and emotion-free decisions (Volmer et al., 2016). However, in their extreme forms, they may overlap with clinical personality disorders, such as Narcissistic Personality Disorder or Antisocial Personality Disorder, and display associations with dysfunctional and antisocial behaviors (Zeigler-Hill et al., 2017).

Regarding criminal behaviour, empirical data attests to the association between *Narcissism* and delinquent behaviors (Krusemark et al., 2018), and emphasizes that *Psychopathy* can predict lethal violence (Fox & DeLisi, 2018), sadistic and aggressive

tendencies (Shafqat et al., 2019), recidivism, crime intensity, cruelty (Međedović et al., 2012) and sense of entitlement to hurt others (Dargis et al., 2017).

The conceptualization of the dark personality through the lens of the Dark Triad has been extended in the **Dark Tetrad** (Buckels et al., 2013) to include the personality trait *Sadism*. Individuals high on Sadism were characterized as persons "*who humiliate others, show a longstanding pattern of cruel or demeaning behavior to others, or intentionally inflict physical, sexual, or psychological pain or suffering on others in order to assert power and dominance or for pleasure and enjoyment*" (O'Mera et al., 2011, p. 523). Triggered by their own suffering, they are likely to inflict disutility on others to compensate for their unhappiness (Mokros et al., 2011) and some even consider aggression as a shield against the maleficent world where are forced to live in (Coyne & Ostroy, 2018). Previous research indicated significant correlations Sadism and all the Dark Triad traits (Chabrol et al., 2009), due to their shared tendencies, including low empathic concern, low emotional involvement, and manipulative tendencies (Međedović & Petrović, 2015).

However, the inclusion of Sadism does not capture the full spectrum of dark traits and there are other traits associated with the tendency to satisfy personal self-interest at others' expense.

1.3.3. The Dark Factor Model

A more comprehensive approach to dark personality is the **Dark Factor of Personality Model** proposed by Moshagen, Zettler, and Hilbig (2018). This framework extended the Dark Tetrad to integrate additional aversive personality traits, such as *Egoism, Psychological Entitlement, Spitefulness, Greed, Cruelty, Frustration, Self-Centeredness, and Moral Disengagement*, all related to a wide range of unethical and morally unethical behaviors (Moshagen et al., 2018). According to this model, the common characteristic of all these dark personality traits is a "*general tendency to maximize one's individual utility - disregarding, accepting, or malevolently provoking disutility for others - accompanied by beliefs that serve as justifications"* (p. 657) known as the Dark Factor or D (Moshagen et al, 2018). This tendency, relatively stable across time and culture (Bonfá-Araujo et al., 2023), explains why some people are more willing to use antagonistic, malevolent, and socially questionable behaviors in the pursuit of their own interests, and have justifications that allow them to do so, without feeling remorse and guilt.

Bader et al. (2018) further extended the conceptual framework of the Dark Factor Model, suggesting that the tendency towards self-utility maximization at others' expense, common to the traits of D, is expressed through five interrelated themes: *Callousness, Deceitfulness, Narcissistic Entitlement, Sadism,* and *Vindictiveness*. In addition, they argued that framing the dark personality via these themes is a superior method compared to the trait approach, as they represent the commonalities observed across all the dark traits included in the Dark Factor Model. *Callousness* has been described as indifference towards the distress of others, low empathy, and a deficit in manifesting compassion. *Deceitfulness* has been defined as the willingness to use unlawful, illicit, and deceptive means for self-serving purposes. *Narcissistic Entitlement* was described through greedy tendencies and a desire for power and social status. *Sadism* reflected a tendency towards inflicting disutility on others for personal amusement and enjoyment, while *Vindictiveness* reflected a desire for vengeance, motivated

by resentfulness (Bader et al., 2021). These dark themes demonstrated their effectiveness in distinguishing individuals with criminal behavior from those with no criminal tendencies and negatively correlated with the Bright Triad traits, as presented by Hurezan et al. (2024). Their findings indicated that criminal personality may be described as a combination of low *Conscientiousness,* low *Agreeableness,* high *Neuroticism,* and elevated dark themes like *Callousness* and *Deceitfulness* (Hurezan et al. 2024).

1.3.4. The Pathological Personality Traits (PID5)

Maladaptive or *pathological* personality traits, as delineated in Section III of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, American Psychiatric Association, 2013), are strongly associated with criminality (Hopwood & Sellbom, 2013), the Dark Triad (Grigoras & Wille, 2017) and the Big Five (Strickland et al., 2013). Evaluated via the **Personality Inventory for DSM-5** (PID-5, APA, 2013, Krueger et al., 2012) they represent the maladaptive dimensions of Big Five personality: *Negative affectivity* (Neuroticism), *Detachment* (Extraversion), *Antagonism* (Agreeableness), *Disinhibition* (Conscientiousness) and lastly *Psychoticism* (Openness to Experience),

From an evolutionary standpoint, pathological traits emerged as adaptations to survival challenges (Harkness et al., 2014), facilitating individuals in the detection of immediate threats (Negative affectivity), allowing resource acquisition with little emotional engagement (Detachment), maximizing self-interest over other's wellbeing (Antagonism), considering the immediate costs of actions against the long-term benefits for others (Disinhibition), and altering external reality to align with one's own goals (Psychoticism). Supporting this theory, Zeigler-Hill et al. (2017) documented positive correlations between Negative Affectivity and vigilance towards potential threats, Antagonism, and an egoistic desire for self-fulfillment, Disinhibition, and potentially harmful behaviors towards others, as well as Psychoticism and difficulties in accurately depicting reality.

In addition, research on these pathological traits revealed their associations with criminal thinking styles (Zeigler-Hill et al., 2017), deficits in emotion regulation, interpersonal functioning (Wright et al. 2012), and aggression (Hopwood et al., 2013), highlighting their role in forensic settings (Niemeyer et al., 2022).

1.4. Deception[2]

1.4.1. Conceptual Clarifications

Deception is commonly defined as "*a deliberate attempt, whether successful or not, to mislead others by concealing, fabricating, or manipulating factual and/or emotional*

[2] This sub-chapter (pages 6-7) overlaps with parts of the published manuscript: Turi, A., Rebeleș, M. R., & Visu-Petra, L. (2022). The tangled webs they weave: A scoping review of deception detection and production in relation to Dark Triad traits. *Acta Psychologica*, 226, 103574. https://doi.org/10.1016/j.actpsy.2022.103574

information through verbal or nonverbal means, in order to induce or maintain a belief the communicator considers false" (Masip et al., 2004, p.148).

Deceptive *communication* shares many features with truthful communication, as individuals often combine accurate and inaccurate information to enhance persuasiveness (Burgoon & Buller, 1996). In contrast to truth-telling, where the sender communicates with no hidden agenda, deception requires the sender to mask their intent, suppress the truth, and manage the cues that could reveal dishonesty. This entails both the cognitive ability to distinguish truth from falsehood and the motivational willingness to mislead (Masip et al., 2004). Still, not all deceptive attempts are effective and the *success of deception* hinges on both the parties involved: (1) the communicator's ability to appear honest while concealing deception cues and (2) the receiver's ability to detect the falsity of the message, as well as on the context in which deception takes place (Burgoon & Buller, 2008).

1.4.2. Deception Motivations

People lie for different reasons, from avoiding interpersonal conflicts to preserving, protecting, or nurturing interpersonal relationships, to gaining social status and power. Evidence shows that the intention behind the lie and the motivation for which the liar engages in deception production determines the pro-social or anti-social nature of the lie (Visu-Petra et al., 2022). This is because people usually judge the character of a lie based on liar's degree of awareness, their motives, and effects on the parties involved (Knapp & Comadena, 1979).

De Paulo (1996) proposed that lies can be either (1) *others-oriented*, told to protect or enhance someone else's interests, usually referred to as white lies and (2) *self-oriented,* told to protect the liar's interests. Further, Bryant, (2008) provided a more nuanced understanding of lying revealing three categories of deception: *white lies, gray lies,* and *real lies* based on the qualitative examination of individuals' beliefs. More recently, Visu-Petra et al. (2022) proposed a taxonomy that categorizes lies based on two intersecting axes: *self-interest* and *other-interest*. This framework identifies four types of lies, each defined by the deceiver's intent and anticipated impact on others. Lies in the first category, marked by low self-interest and high other interest, are altruistic, empathetic and may even involve personal sacrifice (Nagar et al., 2020; Eisenberg et al., 2002). Lies in the second category, characterized by high self-interest and high other interest, are motivated by social desirability, and aim to preserve relationships (most white lies). In the third category, lies driven by high self-interest and low other interest serve primarily to avoid personal consequences, often at the expense of others (antisocial lies). Finally, the fourth category, low self-interest, and low other interest, comprises maladaptive lies that benefit neither the self nor others and are often used by what are called "pathological liars". This taxonomy provides valuable insights for understanding deception in prisons, where deception is both a survival strategy and a social tool (Dhami et al., 2007).

1.4.3. Deception General Ability

An important, yet relatively underexplored, avenue in deception research is the potential existence of *a deception general ability* which might facilitate both the production of

effective lies and the accurate detection of deception in others. This idea was proposed by Wright et al. (2012, 2015), who introduced a novel *ecologically valid* paradigm called the Deceptive Interaction Task (DeceIT) and used Signal Detection Theory, to provide empirical support for the hypothesis that deception skills may stem from a "**general deception ability**" rather than two isolated functions (Wright et al., 2012). The authors expanded the investigation in a follow-up study to examine how various individual differences, including personality traits like those from the **Dark Triad** (Pulhus & Williams, 2002) or maladaptive personality traits from the **PID-5 model** (Krueger et al., 2012), as well as emotional capabilities (e.g., TOM and alexithymia) may influence deception production and detection abilities (Wright et al., 2015). Surprisingly, the researchers found no significant correlation between Dark Triad traits and deception detection, despite individuals with high scores on these traits demonstrating overconfidence in their lie detection abilities. Conversely, a positive correlation was found between Detachment from **PID-5** (the maladaptive trait corresponding to low Extraversion) and lie detection accuracy, suggesting that a detached approach to social interaction may confer a lower response bias and a better detection accuracy. Building upon this foundation, Semrad et al. (2020) conducted a study involving federal police recruits, utilizing a modified version of the DeceIT paradigm. Contrary to Wright et al.'s (2012) conclusions, Semrad et al (2020) did not observe a significant correlation between deception detection and production abilities within their sample. This divergence raises questions about the universality of the "*deception general ability*" and suggests that the specific characteristics of the sample population may influence the interplay between lying and lie detection skills.

1.5. Theoretical Accounts for Interpersonal Deception

1.5.1. The Interpersonal Deception Theory (IDT, Burgoon & Buller, 2008)

The Interpersonal deception Theory (IDT, Buller & Burgoon, 1996) establishes an extensive theoretical framework for understanding deception in everyday interactions. Within this paradigm, deception has been defined as a message delivered by a sender with the goal of creating a false belief in one or more receivers (Buller & Burgoon, 1996), and portrayed as a dynamic, interdependent process that evolves through three main phases: pre-interaction, interaction, and post-interaction. Through this process, senders create and deliver deception while managing their strategic and non-strategic behaviors in such a way as to avoid detection and maintain credibility.To this message, receivers may respond with suspicion, or place their trust in the sender, unaware of any deceptive intentions, thereby affecting the sender's behavior. Because these cognitive processes take place simultaneously, there is a great likelihood of non-strategic behaviors or unintended indicators of deception (behavioral "leakage"), such as changes in speech patterns, signs of nervousness, or inconsistencies (Burgoon et al., 1995; Vrij, 2008). These indicators typically emerge from the cognitive and emotional processes associated with the act of deception (e.g., fear of being discovered, apprehension) and from contextual circumstances (e.g., high-stakes situations).

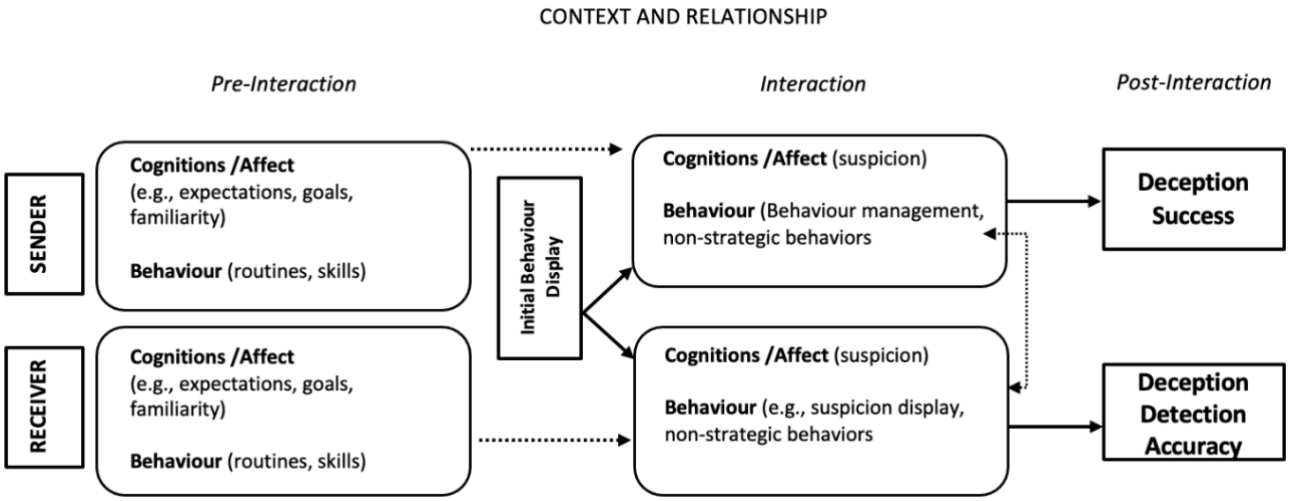

Figure 1.
Interpersonal Deception Theory model (simplified from Burgoon & Buller, 2008)

1.5.2. An Integrative Model of Interpersonal Deception

Although the IDT (IDT, Buller & Burgoon, 2008) provides a comprehensive understanding of deception in interactive settings, its application has predominantly been confined to the general population (e.g., Wright et al., 2012), leaving a significant gap in the literature concerning inmate population.

In the prison environment, individuals might encounter more opportunities to deceive, as well as receive deceptive messages and their abilities. Due to their personality predispositions, they may have different goals, expectations, and behaviors as compared to the general community. For instance, recent findings indicated that prisoners may present higher levels of dark personality traits (Dark Factor traits), especially *Deceitfulness* (i.e., a dark personality trait that might provide them with justifications for deception as required, acceptable, or fair for self-preservation or self-utility maximization), and low levels of Extraversion, indicating possible deficits in assertive communication and sociability (Hurezan et al, 2024).

While the IDT model accounts for the importance of cognitive/affective and behavioral factors across the pre-interaction and interaction phases, the beliefs, attitudes, prior experience with deception, and personality traits are not specifically addressed. To address this limitation, we aim to develop an **integrative model for understanding interpersonal deception** that builds upon the original framework established by Buller & Burgoon (2008), incorporating cognitive, dispositional, and affective elements across the pre-interaction, interaction, and post-interaction phases. These potential extensions are written with colors in **Figure 2.**

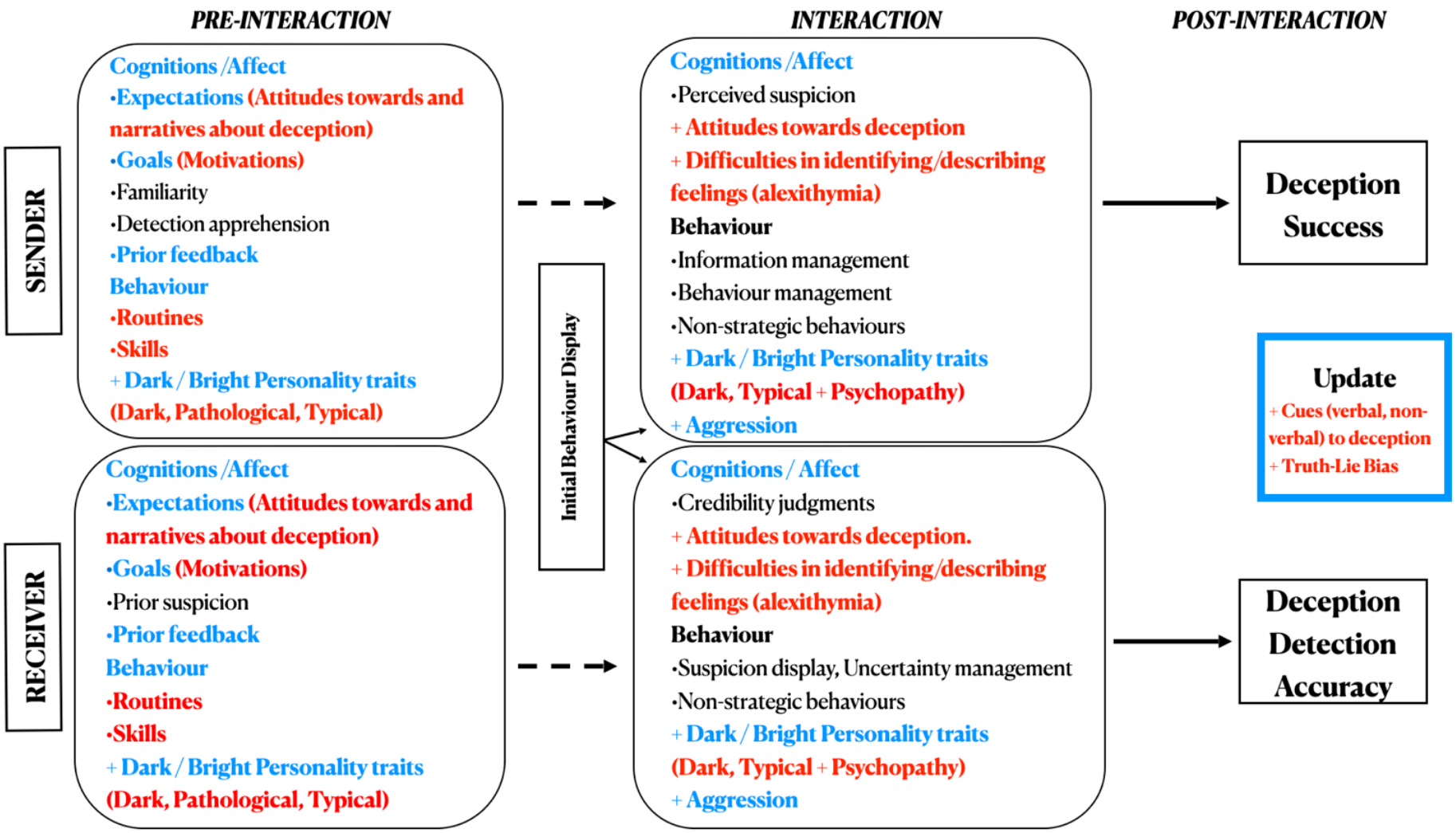

Figure 2.

An Integrative Model of Interpersonal Deception

Note. Possible extensions to the Interpersonal Deception Theory (IDT) Model are written with colors.

RESEARCH OBJECTIVES AND GENERAL METHODOLOGY

2.1. Theoretical Objectives

Aim 1:

The first aim of this thesis was to bring together the insights of the research relating **dark personality traits (Dark Triad)** to **deception production and detection** (*Study 1*) and examine the degree of alignment between individuals' perceptions (self-reports) and performances (experimental tasks) in deception production and detection.

Aim 2:

The second aim was to investigate the **predictive role of typical, pathological, and dark personality traits on lie acceptability** in inmates, assuming *attitudes towards deception* can be an explanatory mechanism for deceptive behaviors (*Study 2*).

Aim 3:

The third aim was to explore **inmate's beliefs about lying and deception detection** (*Study 3A and B*). In Study 3A we aimed to examine how inmates define deception, what narratives they attach to different types of lies, what is the degree of endorsing favorable or unfavorable attitudes towards lying, and self-reported lie frequency, for understanding how dispositional *Deceitfulness* can manifest behaviorally and how likely are individuals with dark traits to reframe antisocial deception as prosocial, morally justified, or necessary. In study 3B, we focused on their **beliefs about deception detection**, including their perceptions of successful deception, and reliance on specific cues, in an effort to identify the extent to which they endorse stereotypical opinions regarding the real indicators of deception.

Aim 4:

The final and most important aim of this thesis was to empirically verify the existence of a **general deception ability** hypothesis in both incarcerated and non-incarcerated individuals (*Study 4).* In addition to examining whether individuals who are proficient in deception production also demonstrate superior skills in deception detection, we examined the influence of both bright (e.g., *Agreeablenes, Conscientiousness*) and dark personality traits (e.g., *Callousness*, *Deceitfulnes*s, *Sadism*) on deception-related outcomes.

2.2. Methodological Objectives

Aim 1:

In Study 2, we aimed to examine the predictive value of the **5 Dark themes** (D70) for **lie acceptability** (a proxy for deception), over and above the **Big5 personality traits (**NEO-FFI) and the **PID-5 maladaptive personality domains**.

Aim 2:

Our second methodological aim was to develop and implement a **Qualitative and Quantitative Research Protocol** to examine incarcerated individuals' beliefs about deception production (Study 3A) and detection (Study3B) and employ the **Thematic Analysis** approach developed by Braun & Clarke (2006) to analyze inmates' responses.

Aim 3:

Our third methodological aim was to modify and adapt the **Game of DeceIT paradigm** (Wright et al., 2012) and establish its validity as a measure of deception production and detection in both the prison and the **general population** (*Study 4*). Additionally, we used **Signal Detection Theory** (SDT; Zloteanu, 2024)**,** to analyze the participants' performances.

Aim 4:

Our final methodological aim was to examine whether the **Dark Themes measured with D70** are associated with **two SDT outcomes**, as well as to validate the self-report measures for additional bright (Big5) and dark (Alexithymia and Aggression) and examine their association with deception success and deception detection accuracy.

The presentation of these theoretical and methodological objectives provides the foundation for the detailed studies discussed in the subsequent chapter. Prior to their description, it is essential to emphasize that all methods and procedures employed are in full compliance with both national and international regulations, including the guidelines of the American Psychological Association on responsible research conduct, as well as the recommendations of the European Committee for the Prevention of Torture and Inhuman or Degrading Treatment. These aspects have been elaborated in detail within the research project, which was approved by the Research Ethics Committee of Babeș-Bolyai University (no. 23553/17.12.2018, updated by no. 1823/09.02.2024).

ORIGINAL RESEARCH CONTRIBUTIONS

3.1. Study 1: A Scoping Review of Deception Detection and Production in Relation to Dark Triad Traits[3]

Deceptive behavior is a pervasive social strategy (Bryant, 2008) with major costs for both interpersonal relationships and for society (Markowitz, 2020). Some people perceive themselves as better liars (Wissing & Reinhard, 2019), and it has been suggested that some personality traits might be beneficial for lie production (Wright et al., 2015). To test the assumption that personality can explain individual differences in deceptive performance, several studies focused on the Dark Triad (DT) traits; because of their shared tendency toward the manipulation and exploitation of others for personal gain, including deploying a wide variety of deceptive tactics (Paulhus & Williams, 2002).

3.1.2. Method

Study aims: (1) To bring together the studies investigating self-reported and performance-based deception production and detection in the DT personality traits, and (2) To document the contribution of DT features in explaining individual differences in deception production and detection, beyond the Five-Factor model of personality.

We used Arksey and O'Malley's (2005) methodological framework for scoping reviews and followed five stages. To identify relevant studies, we conducted multiple computer-based searches on several electronic databases, including, but not limited to, ERIH PLUS, Psych Info, Web of Science and Science Direct, and Google Scholar, using the keywords such as deception, deception detection, deception production, lie production, lie detection, Dark traits, Dark Triad, etc. After identifying and screening the publications according to our inclusion-exclusion criteria and removing the duplicates, 18 publications remained eligible. From these, 7 used self-reports to assess perceived deceptive skills and 11 using experimental tasks to measure deceptive performances, with 1 using both assessment tools. See Table 2 and Table 3 for a visual summary of the significant findings reported by all studies included).

[3] The content of this sub-chapter (pages 13-23) overlaps with parts of the manuscript: "The tangled webs they weave: A scoping review of deception detection and production in relation to Dark Triad traits", published by Turi, A., Rebeleș, M. R., & Visu-Petra, L. (2022) in *Acta Psychologica*, 226, 103574. https://doi.org/10.1016/j.actpsy.2022.103574.

3.1.2. Results and Conclusions

Table 2.

The Dark Triad and Self-reported Deception

Authors (Year)	Dark traits (Measures)	Deception (Measures & Indexes)	Participants, Gender, Mean age	Main Findings
Deception production				
Baughman et al. (2014)	Dark Triad (SDT-3)	- Probability of lying in mating and academic contexts - Emotional state when lying (positive-negative) - Cognitive effort to lie successfully - Others' reactions (the degree to which they believed that the person being lied to: partner or lecturer, would believe their lie)	N = 462 130 men (19.4 years)	*Mating context:* - N was unrelated to the probability of lying (r = .07, n. s.) and correlated with positive emotions when lying (r = .25**), increased cognitive effort (r = .18**), and belief that the partner will believe their lie (r = .11*) - M correlated with the probability of lying (r =.10*), positive emotions when lying (r = .34**), increased cognitive effort (r = .15**), and belief that the partner will believe their lie (r = .13**) - P correlated to the probability of lying (r =.13**), positive emotions when lying (r = .46**), increased cognitive effort (r = .14**), and unrelated to the belief that the partner will believe their lie (r = .06, n. s.) *Academic context:* - N correlated with the probability of lying (r = .14**), positive emotions when lying (r = .28**), a belief that the lecturer will believe their lie (r = .19**), and unrelated with increased cognitive effort (r = .08, n. s.) - M correlated to the probability of lying (r =.25**), positive emotions when lying (r = .33**), increased cognitive effort (r = .28**) and belief that the lecturer will believe their lie (r = .16**) - P correlated to probability of lying (r =.19**), positive emotions when lying (r = .42**), increased cognitive effort (r = .10*) and the belief that the lecturer will believe their lie (r = .17**)
Jonason et al., (2014)	Dark Triad (SRP-III) (MACH-IV) (NPI-40)	Total number of lies in the last 7 days,	N = 447 from which 161 men	- N correlated with the total number of lies (r = .10*), self-gain lies (r = .20**), no reason for lies (r = .18*), and self-rated ability (r = .29**) and unrelated with the use of white lies (r = .06, n. s.) or the number of people lied to (r = .07, n. s.)

		number of people, and number of - self-gain lies - white lies - no reason for lies (Self-rated lying ability) (DMTS) for intersexual and intrasexual deception	(23.4 years)	- M correlated with total number of lies (r = .21**), the number of people lied to (r = .20**), self-gain lies (r = .12*), white lies (r = .13*), no reason for lies (r = .16*), and self-rated ability (r = .27**) - P correlated with total number of lies (r = .21**), number of people lied to (r = .25**), self-gain lies (r = .14*), no reason for lies (r = .26**), self-rated ability (r = .40**), unrelated with use of white lies (r = .07, n. s.) - Individuals scoring high on P and M were more similar in the "cheat strategy" (numerous correlations with intra/intersexual deceptive tactics, such as dominance, sincerity, superiority, and indifference). In contrast, N was associated only with intersexual deception for dominance and appearance.
Azizli et al., (2016)	Dark Triad (SDT-3)	(PTLQ): Lying behaviors and propensity to lie in 2 scenarios: (1) mating (2) academic (CMI)	N = 464, 131 males 333 females (19.5 years)	- N unrelated to a general propensity to lie (r = .03, n. s.) but correlated with lying in both mating (r = .15**) and academic (r = .17**) contexts. - M correlated with general propensity to lie (r = .12**) and lying in both mating (r = .21**) and academic (r = .30**) contexts. - P correlated with general propensity to lie (r = .15**) and lying in both mating (r = .19**) and academic (r = .21**) contexts. - All three DT traits correlated with the total score for misconduct, as follows: N (r = .16**), M (r = .22**), and P (r = .45**), and associated differently with CMI subscales, such as bullying, drug abuse, delinquency, and criminality.

Daiku et al., (2021)	Dark Triad (DTDD)	(Total number of lies in the last 24 hours)	N = 340 (19.6 years)	The mean for lie-telling in the last 24 hours was 2.14 lies (SD = 4.64) Distribution of results: - 45.4% of participants reported no lies - 47.4% reported one to five lies - 7.2% reported six or more lies, which accounted for 47.2% of the total reported lies (154 out of 326 lies) = the "*a few prolific liars*" Lying frequency was correlated with P (r = .14**), M (r = .10*) and unrelated to N (r = -.08, n. s.)
Forsyth et al. (2021)	Dark Triad (SDT-3) (SSIS)	Vignettes assessing the propensity to lie in three separate contexts: - professional - academic - relationship For each context: - Lying efficacy, - Cognitive load - Emotional response to lying (positive and negative)	N = 615 (26.8 years)	*Professional context:* - N correlated with propensity to lie (r = .18***), Lying Efficacy (r = .28***), cognitive load (r =- .24*), positive affect (r = .33***) and negative affect (r = -.09*) - M correlated with propensity to lie (r = .38***), Lying Efficacy (r = .36***), cognitive load (r =- .17***), positive affect (r = .37***) and negative affect (r = -.14***) - P correlated with propensity to lie (r = .23***), Lying Efficacy (r = .24***), cognitive load (r =- .25***), positive affect (r = .29***) and negative affect (r = -.14***) *Academic context:* - N correlated with propensity to lie (r = .19***), Lying Efficacy (r = .31***), cognitive load (r = -.18*), positive affect (r = .34***) and negative affect (r = -.09***) - M correlated with propensity to lie (r = .39***), Lying Efficacy (r = .25***), cognitive load (r = -.19***), positive affect (r = .39***) and negative affect (r = -.20***) - P correlated with propensity to lie (r = .23***), Lying Efficacy (r = .24***), cognitive load (r = -.27***), positive affect (r = .29***) and negative affect (r = -.17***) *Relationship context:*

				- N correlated with propensity to lie (r = .21***), Lying Efficacy (r = .18***), cognitive load (r = -.08*), positive affect (r = .29***) and unrelated to negative affect (r = -.03, n. s.) - M correlated with propensity to lie (r = .24***), Lying Efficacy (r = .24***), cognitive load (r = -.11***), positive affect (r = .24***) and negative affect (r = -.06***) - P correlated with propensity to lie (r = .15***), Lying Efficacy (r = .12***), cognitive load (r = -.19***), positive affect (r = .19***) and negative affect (r = -.11***)
Deception production and detection				
Zvi & Elaad (2018)	Narcissism NPI with 3 subscales: - LA - GE - EE	(LTAAS) Successfully lie-telling Convincingly truth-telling Successfully lie-detecting Believing other people Total number of lies in the last 7 days, number of people, and number of - self-gain, - altruistic, - no reason lies	N = 125 (25.7 years)	Total N correlated with successfully lie-telling (r = .57**), convincingly truth-telling (r = .52**), successfully lie-detecting (r = .52**) and believing other people (r = .38**). Total N correlated with the tendency to lie (r = .29*) the number of people lied to (r = .24**), and differences were observed for N's subscales, as follows: - LA correlated with tendency to lie (r = .20*) and telling lies for no reason (r = .18*) - GE correlated with tendency to lie (r = .31*) and the number of people lied to (r = .31*) - EE correlated with tendency to lie (r = .29*), the number of people lied to (r = .21*) and self-gain lies (r = .19*) N dimensions were unrelated to telling altruistic lies.

Wissing & Reinhard (2019)	Dark Triad (SDT-3)	Three 7-point Likert-type scales: - Perceived deception detection ability - Perceived deception production ability - Behavioral cues of deception (Hartwig and Bond, 2011)	N=205 58.5% male 41.4% female (22-70 years)	- N correlated with perceived deception detection ability (r = .16*) and deception production ability (r = .33***) - M was unrelated to perceived deception detection ability (r = .12, n. s.) and correlated with perceived deception production ability (r = .45***) - P correlated with perceived deception detection ability (r = .14*) and deception production ability (r = .44***) No sig correlations between any of DT and cue-based detectability.

Note: * p<.05. ** p < .01, *** p < .001, n. s. = not significant, n.m. = not mentioned
Personality: N = *Narcissism*, M = *Machiavellianism*, P = *Psychopathy*, S = *Sadism*, SDT-3 = *Short Dark Triad* (Jones & Paulhus, 2014), NPI = *Narcissistic Personality Inventory, with three subscales: LA – Leadership/Authority, GE - Grandiose Exhibitionism, EE – Entitlement/Explosiveness (*Raskin & Hall, 1988), MACH-IV = *Machiavellianism Scale Version IV* (Christie & Geis, 1970), DTDD = *Dark Triad Dirty Dozen* (Jonason & Webster, 2010), SSIS = *Short Sadistic Impulse Scale* (O'Meara, Davies, & Hammond, 2011), Deception: DMTS = *Deceptive Mating Tactics Scale* (Took & Camire, 1991), PTLQ = *Propensity to Lie Questionnaire* (Azizli et. al., 2016), CMI = *Comprehensive Misconduct Inventory* (Paulhus & Williams, 2002), LTAAS = *Lie-truth Ability Assessment Scale* (Elaad, 2009/2015).

Table 3.
The Dark Triad and performance-based Deception

Authors (Year)	Dark trait(s) (Measures)	Deception tasks	Participants (Mean age)	Findings (correlation coefficients, and statistical significance)
Deception production				
Geis & Moon, (1981)	Machiavellianism (MACH-IV)	*Lie production* (Group 1) One half denied the knowledge of a theft, in which they had just been directly implicated; the other half made the same denial truthfully. *Lie detection* (Group 2)	N = 360 64 Group 1 (G1) 64 Group 2 (G2) (n.m.)	Comparisons between individuals scoring high and low on M, revealed: - In the Lie condition, judges believed more the lies told by individuals scoring high on M, than they believed the lies told by those scoring low on M (p < .01). No significant difference was observed in the truth condition considered alone (n. s.)

		Watched the 1.25 min videotape clips in random sequence and judged the veracity of denials		- Judges discriminated less accurately between individuals scoring high on M lies and truths, than in the case of those scoring low on M $p < .001$)
Martin & Leach (2013)	Psychopathy (PPI-R)	Selected videos of 15 lie-tellers and 30 truth-tellers. Lie-tellers had been prompted to cheat by the confederate, cheated on the test, and denied cheating during the interrogation. Truth-tellers had not been induced to cheat, did not cheat, and denied cheating during the interrogation.	N = 117 from which 53 males, 64 females (19.8 years)	No correlations between the global P factor, individual content scale scores, and deception detection (all p >.05) Note: The sample included many individuals scoring high on P even when compared with criminal populations.
O'Reilly & Doerr (2020)	Narcissism (Resick) (NPI-16) (SINS)	(1) Lying = either failing to tell a car buyer about a faulty water pump (scenario S1) or telling a co-worker that a recommendation option 2 gives the co-worker more credit (scenario S2) (2) Cheating in an online game of rolling the dice (3) Self-reported willingness to steal a charger	N = 401 (34 years)	- The three N measures were inter-correlated: - SINS correlated with Lying in S1(r = .26**), but not in S2, also with Cheating (r = .15*) and Willingness to steal (r = .19**) - NPI-16 correlated with Lying in S1 (r = .23**) and S2 (r = .24**), but unrelated to Cheating and Willingness to steal. - Resick was unrelated to Lying in both scenarios, but correlated with Cheating (r = .19*) and Willingness to steal (r = .27**)
Michels et al. (2020)	Dark Triad N (NARQ) M (MACH-VI) P (SRP-4) Intelligence (WAIS-IV)	*Lying ability* (LA) = indirectly measured by G2 (no of raters successfully misled / no of raters that judged the subjects' stories) *Lie production:* 50 students narrated three short stories that took place in the last 24 hours, 2 true and 1 fictional, while videotaped in laboratory settings *Lie detection:* 13 raters were informed that one of the three stories was fabricated	N = 50 students (Group 1) N = 13 academic assistants/interns (Group 2) (22.6 years)	- P (Antisocial Behavior Subscale) correlated with LA (r = .30*), whereas other subscales did not. - M was unrelated to LA (r = −.02, n. s.) - N was unrelated to LA (r = −.09, n. s.) - Intelligence was unrelated to LA (r = .15, n. s.)

		and requested to watch and judge which story was deceptive		
Deception detection				
DePaulo & Rosenthal (1979)	Machiavellianism (MACH)	*Deception production:* Participants were videotaped while describing someone they liked, disliked, and felt ambivalent and indifferent about. *Deception detection:* Participants returned to judge one of these videotapes. They always judged a videotape on which they did not appear.	N = 40 Students (n.m.)	- The ability to recognize deception when the speaker is hiding positive affect is not significantly related to the ability to recognize deception when the speaker is hiding negative affect (r = -.18, n. s.) - Speakers who get caught lying by women also tend to get caught lying by men (r=.62** for positive affect and r-.54** for negative affect) - High M were more successful at getting away with their lies than low M (F = 3.04, p =.09; d =.58) - High M were especially successful at deceiving when pretending to dislike someone they genuinely liked, (F = 5.96, p < .05; d = .79)
Lyons et al. (2013)	Psychopathy (SRP-III) Primary P Secondary P	*Deception detection* Online experiment presenting 26 clips (real-life high stakes appeals, 13 truthful and 13 lies, from international missing person websites. Truth/False judgments required	N =150 (21.1 years)	Sex moderated the relationship between P and lie detection: - In men, primary P was positively correlated with lie detection (r = .26**) - In women, primary P was negatively correlated with lie detection (r = -.24*) - Secondary P was unrelated to lie detection in both men (r = -.22, n. s.) and women (r = .18, n. s.)
Lyons et al., (2017)	Dark Triad (SDT-3)	*Deception detection* Online experiment presenting 20 clips with real-life high-stakes appeals, 10 truthful and 10 lies, from international missing persons websites Truth/False judgments were required.	N = 347, from which 98 men (25.7 years)	In high-stakes deception, results revealed sex differences when judging the veracity of emotional lies: - In men, only N correlated negatively with deception detection accuracy (β = −0.24, t = −2.30, p<0.02). - In women, only M correlated positively with deception detection accuracy (β = 0.28, t = 3.46, p<0.001)

Schindler et. al. (2019)	Dark Triad (NARQ) (MACH-IV) (SRP-III)	*Lie production:* (prior to the study) Truth condition: 10 participants witnessed a confederate stealing 10 euros. Lie condition: The other 10 participants were asked to steal. Next, liars and truth-tellers were asked about the missing money, *Lie detection:* Students and offenders watched video recordings and judged the veracity of the presented statements.	N = 20 males (n.m. /students) (liars/truth-tellers) N = 76 males Offenders (29.7 years) N= 43 males (25.2 years)	No significant correlations between DT and classification accuracy (all p > .35), judgment confidence (all p > .06), and correct beliefs about deception (all p > .35) Offenders had higher levels of N (F = 7.21, p <.008), M (F = 6.96, p <.009) and P (F = 53.05, p <.001) and both groups showed an equally strong truth bias (n.m.) Confidence in own judgment was negatively correlated with classification accuracy (r = -.23*) and messages judged as true (r = .24*)
Deception production and deception detection				
Wright et al. (2015)	Dark Triad (NPI-16) (MACH-IV) (SRP-SF)	*Deception production and detection* Computer administered deceptive interactive task DeceIT, participants (Senders) took turns making true or false 20-30 second verbal statements, and they judged the statements produced by others (Detectors) While each participant completed 80 trials in the role of Sender, while the rest of the participants were lie Detectors. Self-deception subscale (BIDR) Lie acceptability (RLAS)	N = 75 28 males 47 females (27.2 years)	DT traits were unrelated to the ability to produce lies which others found difficult to discriminate from the truth (deception production) or to discriminate truth from lies when judging others (deception detection). - M was correlated with Lie acceptability (r = .38**) - N was correlated with Self-deception (r = .25*) - Lie Acceptability was correlated with deception production (r = -.24*).
Semrad & Scott-Parker (2020)	Dark Triad (SD3)	*Deception production and detection* Face-to-face DeceIT task, in which the participant (Sender) speaks (either the truth or a lie, depending on the card instruction) for approximately 20 seconds, attempting to convince other participants that it reflects their true	N = 50 Australian Federal Police recruits (30.2 years)	No significant relationships were found between any DT measure and neither truth nor lie production.

		opinion. The rest of the participants: lie Detectors		
Elaad et al., (2020)	Narcissism (NPI With 3 subscales: - LA - GE - EE)	(LTAAS) Successful lie-telling Convincing truth-telling Successful lie-detecting Believing other people *Deception production* A modified version of the original ultimatum game which investigates deception, where participants were requested to deceive another to gain more points	N = 70 From which 42 males 28 females (24.6 years)	- The lie-telling ability assessment was the only significant predictor of deception production (β = .45, t = 3.38, p = .001) - All three N dimensions predicted deception. - LA contributed to self-assessed lie-telling (β = .29, t = 2.54, p = .012), lie-detecting (β = .29, t = 4.26, p < .001) and truth-telling (β = .40, t = 3.64, p < .001) but not to truth-detecting ability - GE contributed to self-assessed lie-telling (β = .40, t = 3.60, p = .001), lie-detecting (β = .47, t = 4.47, p < .001) and truth-telling (β = .38, t = 3.44, p < .001) but not to truth-detecting ability - EE contributed to self-assessed lie-telling (β = .31, t = 2.76, p = .007), lie-detecting (β = .40, t = 3.67, p <.001) and truth-telling (β = .45, t = 4.21, p<.001) but not to truth-detecting ability

Note: * p<.05. ** p < .01, *** p < .001, n. s. = not significant, n.m. = not mentioned

Personality: N = *Narcissism*, M = *Machiavellianism*, P = *Psychopathy*, S = *Sadism*, SD3 = *The D3 Short* (Paulhus & Williams, 2002) SDT-3 = *Short Dark Triad* (Jones & Paulhus, 2014), NPI = *Narcissistic Personality Inventory, with three subscales: LA – Leadership/Authority, GE - Grandiose Exhibitionism, EE – Entitlement/Explosiveness (*Raskin & Hall, 1988), NARQ = *Narcissistic Admiration and Rivalry Questionnaire* (Back et al., 2013), WAIS-IV = Wechsler Adult Intelligence Scale – Fourth Edition (Petermann, 2012), Resick = *eight adjectives descriptive of Narcissism (arrogant, assertive, boastful, conceited, egotistical, self-centered, show-off, and temperamental)* (Resick et. al., 2009), NPI-16= *Narcissistic Personality Inventory* (Ames et al., 2006), SINS = *Single-Item Narcissism Scale* (Van der Linden & Rosenthal, 2016), MACH-IV = *Machiavellianism Scale Version IV* (Christie & Geis, 1970), MACH-IV= *Machiavellianism Scale VI* (Jones & Paulhus, 2009), SPR-4 = *The Self-report Psychopathy Scale – Forth Edition* (Paulhus et. al, 2016), SRP-III = *Self-Report Psychopathy Scale-III* (Paulhus et. al., 2009), SRP-SF = *Sub-Clinical Self-Report Psychopathy Questionnaire Short-Form* (Paulhus et. al., 2015); BIDR (Self-Deception Scale of the Balanced Inventory of Desirable Responding, Paulhus & Reid, 1991). Deception: DeceIT = *The Deceptive Interactive Task* (Wright et al., 2012), LTAAS = *Lie-truth ability assessment scale* (Elaad, 2009/2015)

Deception Production

The majority of studies using **self-reports** showed Machiavellianism and Psychopathy were frequently associated with different aspects of deception, such as increased lie frequency and propensity to lie across various contexts and reported a similar "cheat strategy" (e.g., Baughman et al., 2014; Daiku et al., 2021). Narcissism on the other hand, was unrelated to the general propensity to lie (Azizli et al., 2016) and lie frequency (Daiku et al., 2021) in two studies, while in another, it was associated with a willingness to lie in professional and academic contexts (Forsyth et al. 2021). In terms of **objective measurements**, DT traits were unrelated to deception production performances, measured separately (Wright et al. 2015) and as a unitary construct (using SDT-3, Semrad & Scott-Parker, 2020), in both general and prison populations (Schindler et al. 2019). Investigated separately, all dimensions of Narcissism predicted lie telling (Elaad et al., 2020) and three inter-correlated measurement tools for Narcissism related differently to deception production, cheating, and willingness to steal (O'Reilly & Doerr, 2020). There was only one study assessing Machiavellianism as a separate dimension, showing that individuals with higher levels of Machiavellianism were more successful in producing convincing lies than those with low levels (DePaulo & Rosenthal, 1979), but did not differ in producing truthful statements. When Psychopathy was assessed with PPI-R, no correlations were observed between Psychopathy and lie production (Martin & Leach 2013), compared to SRP-4, where a positive association was found between the antisocial aspects of the construct and lie telling (Michels et al. 2020).

Deception Detection

Two studies investigated DT's **perceptions of deception detection skills**, and most of the literature focused on investigating deception production. As observed, individuals scoring high on Narcissism and Psychopathy perceived themselves better at successfully detecting lies, but not those scoring high on Machiavellianism (Wissing & Reinhard 2019). In the case of Narcissism, results were replicated when assessed as a separate construct, documenting a positive association between Narcissism, self-rated lie-detecting ability, and increased confidence in their lie detection skills (Zvi & Elaad, 2018). When deception detection performances were assessed with **experimental tasks,** in face-to-face interaction, Wright et al., (2015) failed to find a correlation between DT and increased deception detection performance (assessing personality with separate instruments: NPI, MACH-IV, SRP-III), similar to Semrad and Scott-Parker (2019) who assessed DT traits as a unitary construct (SD-3). Contrary to these results, when judging high-stakes deceptive statements, males scoring low on Narcissism and women scoring high on Machiavellianism, proved to be better at deception detection (Lyons et al., 2017). The only two studies investigating the link between Psychopathy and deception detection obtained divergent results (e.g., Martin & Leach, 2013).

To conclude, an ideal way to study deceptive behaviors in relation to personality would require using both self-reports and peer evaluations for the dark personality features and both self-reports and experimental tasks for deception production and detection. In addition, to enhance the predictive power of DT assessment over typical personality traits, it might be helpful to study both typical (e.g., FFM) and aversive (i.e., DT, Dark factor) personality traits, in relation to subjective and objective measurements of deception in high-stake situations (e.g., prison environment).

3.2. Study 2: Dark Personality and Lie Acceptability as a Proxy for Interpersonal Deception in Prisoners

Prior research investigating the relationship between personality and deception indicated that people lie for various reasons, and personality traits may explain differences in lying behavior (Turi et al., 2020). For instance, the adaptive traits of Openness, Conscientiousness, Extraversion, and Agreeableness demonstrated a negative relationship with the frequency of lying (Hart et al., 2020), whereas the Dark triad traits and Sadism demonstrated a positive association with the propensity to lie (Forsyth et al., 2021). Furthermore, individuals with elevated levels of Narcissism, Machiavellianism, Psychopathy, and Antagonism have been more likely to think of themselves as effective deceivers and skilled lie detectors; though, neither the dark nor maladaptive traits improved actual deception detection accuracy (Wissing & Marc-André Reinhard, 2017, 2019).

These suggest an inaccurate perception of effective deception abilities and a higher lying propensity among those possessing dark personality characteristics. In addition, the dark traits have been repeatedly associated with more lenient attitudes towards deception (Wright et al., 2012), also known as lie acceptability (Oliveira & Levine, 2008). The acceptance of deception has been shown to predict both the occurrence and efficacy of deceptive behavior (Levine et al., 2010; Halevy et al., 2014; Hart et al., 2023). Individuals who consider lying appropriate are more inclined to engage in deception and do so successfully (McLeod & Genereux, 2008). When deception is consistently rewarded, which often happens in prison, these behaviors may become permanent potentially encouraging pathological tendencies such as habitual lying or manipulativeness, demonstrated by prolific liars (Serota et al., 2010). Thus, the acceptability of lying may not only indicate an individual's attitude towards deceit but also act as a proxy of the deceptive behavior itself.

While past research suggested that personality traits might significantly impact the perception and justification of deception as acceptable or necessary, thus shaping deceptive behaviors, these studies mostly examined personality frameworks in separate ways. To date, no research has investigated the collective contribution of typical, dark, and maladaptive traits on predicting individuals' attitudes towards deception, which constituted a significant gap in the literature. The present study aims to address this limitation by examining the relationship between normal (adaptive), pathological (maladaptive), and dark personality characteristics and the acceptability of lying among incarcerated individuals, who are presumed to have a higher propensity towards lying.

Our first hypothesis (**H1**) concerned the typical traits of personality represented by the **Big 5 Factors:** *Openness to Experiences* (O), *Conscientiousness* (C), *Extraversion* (E), *Agreeableness* (A) and *Neuroticism* (N). *Openness to Experience*. We expected that people with elevated levels of Agreeableness and Conscientiousness will report lower lie acceptability. These features demonstrated positive associations with empathy, moral integrity, and prosocial behavior (Musek & Grum, 2021) while showing negative relationships with manipulative tendencies and dishonesty (De Vries et al., 2010). (*Exploratory Hypothesis 1*).

Our second hypothesis **(H2)** referred to the **pathological personality traits** assessed through the **PID-5:** *Negative affectivity*, *Detachment*, *Antagonism*, *Disinhibition* and lastly *Psychoticism*, evaluated via the **Personality Inventory for DSM-5** (PID-5, APA, 2013,

Krueger et al., 2012)**.** We anticipated that higher *Antagonism* and *Disinhibition*, characterized by manipulativeness, impulsivity, and hostility, would be associated with greater acceptance of lying. (*Confirmatory Hypothesis 2*).

Our third hypothesis **(H3)** referred to the **dark personality traits** as defined within the **Dark Factor** framework (Hurezan et al., 2024). Building upon prior research on the Dark Triad (Yarbrough & Hart, 2025), we anticipated that elevated levels of the **Dark Factor themes** (*Callousness, Deceitfulness, Narcissistic entitlement, Sadism,* and *Vindictiveness*) will be positively associated with lie acceptability (*Exploratory Hypothesis H3*).

Out last hypothesis **(H4)** concerned the **combined contribution of typical, pathological, and dark personality traits on the prediction of lie acceptability**, aiming to establish an integrative model that considers multiple personality traits in predicting deceptive tendencies, which is especially relevant in high-stakes environments like prisons where deception may serve adaptative purposes and thus, reinforced. (*Exploratory Hypothesis 4*).

3.2.1. Method

Participants

In this study there were included **94 incarcerated offenders** (71 males and 23 females) from the Maximum-Security Gherla Penitentiary. Participants ranged in age from 21 to 67 years (M = 38.98, SD = 10.48), with a mean IQ of 93. The majority (63%) had been convicted of violent offenses.

Measures

First, participants completed an individual interview covering personal information such as age, sex, family, and educational background. Supplementary data, including offense type, sentence length, prison conduct, and recidivism, were extracted from institutional records.

Next, participants completed several self-report measures. Dark personality traits have been assessed through the **Dark Factor Inventory** (D70; Moshagen et al., 2018), the typical personality has been measured via the **NEO Five-Factor Inventory** *(*NEO-FFI; Costa & McCrae, 1992) and the pathological personality has been assessed through the **Personality Inventory for DSM-5** (PID-5; Krueger et al., 2013). Lie acceptability was measured through the **Revised Lie Acceptability Scale** (RLAS; Oliveira & Levine, 2008) and we also asked participants to complete the **Paulhus Deception Scales – Impression Management subscale** (PDS - IM; Paulhus, 1998) to assess their socially desirable tendencies.

3.2.2. Results and Conclusions

Supporting our hypothesis, the correlational analyses (see **Table 1**) revealed significant associations between Lie Acceptability and all five **Dark Factor themes**: Callousness ($r = .48$, $p < .01$), Deceitfulness ($r = .44$, $p < .01$), Narcissistic Entitlement ($r = .44$, $p < .01$), Sadism ($r = .40$, $p < .01$), and Vindictiveness ($r = .36$, $p < .01$).

In addition, Lie acceptability correlated with all maladaptive PID-5 domains: Negative Affectivity ($r = .40$, $p < .01$), Detachment ($r = .30$, $p < .01$), Antagonism ($r = .29$, $p < .01$), Disinhibition ($r = .37$, $p < .01$), and Psychoticism ($r = .32$, $p < .01$). These associations remained significant even after controlling for socially desirable responding.

Table 1

Correlations (and partial correlations) between Lie acceptability, the 5 Themes of D, the PID-5 domains and the Big 5 Factors.

		1.	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**	**11.**	**12.**	**13.**	**14.**	**15.**	**16.**
1.	Lie acceptability	-	.48**	.44**	.44**	.40**	.36**	.40**	.30**	.29**	.37**	.32**	.23*	-.05	-.15	-.29*	-.31
2.	D5: Callousness	.49**	-	.63**	.60**	.60**	.59**	.36**	.33**	.49**	.46**	.29*	.27**	-.20	-.23*	-.57**	-.40**
3.	D5: **Deceitfulness**	.45**	.65**	-	.54**	.69**	.34**	.35**	.42**	.52**	.36**	.29*	.22*	-.17	-.20	-.40**	-.33**
4.	D5: Narcissistic Entitlement	.45**	.62**	.57**	-	.55**	.37**	.30**	.31**	.41**	.34**	.26*	.18	-.06	-.09	-.43**	-.31**
5.	D5: Sadism	.41**	.61**	.71**	.58**	-	.09	.20	.30**	.29*	.23*	.09	.22*	-.18	-.29*	-.23*	-.33
6.	D5: Vindictiveness	.36**	.59**	.34**	.37**	.10	-	.42**	.31**	.39**	.50**	.36**	.25*	.01	.02	-.49**	-.15
7.	PID-5: Negative Affectivity	.39**	.34**	.31**	.27**	.18	.41**	-	.69**	.60**	.81**	.84**	.47**	.15	-.09	-.18	-.15
8.	PID-5: Detachment	.27**	.28**	.32**	.24*	.24*	.29**	.69**	-	.59**	.57**	.63**	.24*	-.28*	-.18	-.21	-.24*
9.	PID-5: Antagonism	.28**	.47**	.48**	.38**	.27**	.39**	.60**	.58**	-	.77**	.69**	.20	-.10	-.08	-.44**	-.13
10.	PID-5: Disinhibition	.36**	.44**	.33**	.31**	.22*	.49**	.81**	.57**	.77**	-	.81**	.30**	-.03	-.07	-.34**	-.13
11.	PID5: Psychoticism	.31**	.26*	.25*	.23*	.07	.35**	.84**	.64**	.69**	.81**	-	.36**	-.08	.06	-.17	-.12
12.	Big 5: Neuroticism	.23*	.26*	.20*	0.17	.21*	.25*	.47**	.24*	.20	.30**	.36**	-	-.23*	-.17	-.19	-.34**
13.	Big 5: Extraversion	-.05	-.19	-.16	-.06	-.18	.01	-.15	-.28**	-.10	-.03	-.08	-.23*	-	.28*	.12	.56**
14.	Big 5: Openness	-.16	-.26*	-.24*	-0.13	-.31**	.01	-.07	-.13	-.07	-.06	.08	-.17	.28**	-	.17	-.36**
15.	Big 5: Agreeableness	-.30**	-.58**	-.42**	-.44**	-.25*	-.49**	-.16	-.17	-.43**	-.33**	-.16	-.18	.12	.19	-	. 31**
16.	Big 5: Conscientiousness	-.33**	-.43**	-.38**	-.35**	-.37**	-.16	-.13	-.17	-.11	-.11	-.10	-.32**	.54**	.39**	.33**	-

Note. N = 94; p < .05 = *; p < .01 = ** Correlations above the diagonal are conducted while controlling for socially desirable tendencies

In relation to the typical personality traits, lie acceptability demonstrated significant negative correlations with Agreeableness ($r = –.29$, $p < .05$) and Conscientiousness ($r = –.31$, $p < .01$), but weak and non-significant negative correlation with Openness ($r = –.15$). However, no significant relationship was observed between Lie Acceptability and Extraversion ($r = –.05$) and a small but statistically significant positive association with Neuroticism ($r = .23$, $p < .05$), though, this relationship diminished after controlling for social desirability.

To test hypothesis H4, we examined the combined contribution of dark, pathological, and typical personality traits in predicting attitudes toward deception, employing a hierarchical multiple regression analysis. As it can be observed in **Table 2**, our findings revealed that he five Dark Factor themes: Callousness, Deceitfulness, Narcissistic Entitlement, Sadism and Vindictiveness, accounted for a significant proportion of variance in Lie Acceptability, $R^2 = .306$, Adjusted $R^2 = .266$, $F (5, 88) = 7.76$, $p < .001$ and the additional contribution of pathological and normative personality traits was non-significant.

Table 2

Regression Model Summary Predicting Lie Acceptability from Dark, Pathological, and Typical Personality Traits

Model	R	R^2	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1. Dark themes	0.553	0.306	0.266	6.129	0.306	7.755	5	88	<.001
2. Dark themes + PID5	0.598	0.357	0.28	6.072	0.052	1.333	5	83	.258
3. Dark themes + PID5 +BIG5	0.617	0.38	0.261	6.15	0.023	0.577	5	78	.717

Note. N = 94, *Model 1* included Dark Factor themes (Vindictiveness, Sadism, Narcissistic Entitlement, Deceitfulness, Callousness); *Model 2* added PID-5 maladaptive domains (Detachment, Disinhibition, Antagonism, Psychoticism, Negative Affectivity), *Model 3* added Big Five traits (Neuroticism, Agreeableness, Conscientiousness, Extraversion, Openness). Dependent variable: Lie Acceptability. The Durbin-Watson coefficient for the model was 2.285.

The associations between lie acceptability and all the five Dark Themes (Callousness, Deceitfulness, Narcissistic Entitlement, Sadism, and Vindictiveness), indicate that emotional, cognitive, and motivational components captured by these themes might favor interpersonal deception in high-risk populations. For instance, individuals who are emotionally indifferent to the suffering of others (Callousness) or believe they are entitled to special treatment (Narcissistic Entitlement), may be more inclined to endorse (and use) deception as an interpersonal strategy. Of all dark traits, Callousness followed by Deceitfulness exhibited the strongest associations with lie acceptability, reinforcing their centrality to the prediction of manipulation and interpersonal exploitation. Further, the pathological personality traits, as measured through the PID-5, also correlated significantly with lie acceptability, particularly Antagonism and Disinhibition. These domains encompass traits such as manipulativeness, impulsivity, irresponsibility, and rule-breaking, all of which have been linked to criminal behavior and poor impulse control in forensic samples (Anderson et al., 2021). Similarly, Detachment and Psychoticism, though to a lesser extent, were positively associated with lenient attitudes towards lying, suggesting that emotional distancing and reality bending, may also

contribute to reduced concern for honesty in social interactions. In contrast to the positive associations between both dark and maladaptive traits, we found negative correlations between typical traits, particularly Agreeableness and Conscientiousness, and lie acceptability, supporting the previous studies highlighting their roles in promoting empathy, rule-following, and moral behavior (Musek & Grum, 2021). However, their influence was significantly weaker and less consistent.

When we investigated the combined effect of this broad spectrum of characteristics and traits, the Dark Factor themes alone accounted for a substantial proportion of variance in lie acceptability ($R^2 = .31$), indicating that aversive interpersonal dispositions are foundational to attitudes endorsing deception. The inclusion of the five PID-5 domains of pathological personality failed to significantly improve model fit, despite their previous associations with perceived deception ability (Wissing & Reinhard, 2019). This finding may reflect conceptual overlap between the dark and pathological traits. For instance, Callousness and Deceitfulness are found in both the Dark Factor and the Antagonism domain of the PID-5, and individuals with higher levels of these traits may share a willingness to maximize their own gain at the expense of others. An interesting, yet expected finding was that the addition of the Big Five traits did not significantly enhance the prediction of lie acceptability. A possible explanation for these results may be that the typical traits, also referred to as bright traits may be more relevant for the prediction of pro-social behavior rather than for explaining individual differences in manipulative, deceptive, and socially aversive tendencies.

Importantly, the present findings align with previous research by Wissing and Marc-André Reinhard (2017, 2019), who demonstrated that individuals high in dark and maladaptive traits perceive themselves as more effective deceivers. In our sample, the strong association between Deceitfulness and lie acceptability suggests that those with aversive dispositions may not only be more inclined to deceive but also morally justify or rationalize such behavior.

3.3. Study 3A: A Mixed Methods Analysis of Inmates' Narratives Shaping their Attitudes towards Interpersonal Deception[4]

The few existing studies on deception in the prison population predominantly relied on quantitative methods, such as surveys and questionnaires, to explore inmates' beliefs about lie detection and their self-perceptions of their ability to produce and detect lies. The findings from these studies have been mixed. For example, teenage offenders demonstrated significantly greater accuracy in evaluating statement credibility than their non-offending peers (Jupe et al., 2016). In addition, inmates exhibited less stereotypical beliefs about deception indicators compared to prison personnel and students (Granhag et al., 2004). Similarly, Vrij and Semin (2006) found that prisoners' beliefs about nonverbal lying behavior were more accurate than those of police officers, prison guards, and students, which supports Ulatowska's (2005, 2009) findings that inmates might have a better understanding of deception than non-criminals. Other findings revealed that inmates were less accurate at detecting lies than students, showing similar truth biases as the general public (Schindler et al., 2021). While inmates may be proficient at detecting lies, they struggle with identifying truths due to a pronounced lie-bias (Bond et al., 2005). Preliminary results from a recent behavioral study comparing prisoners and community members found that prisoners were less capable than non-inmates at detecting deceit in others, and this ability was negatively related to their higher levels of self-reported deceitfulness (Turi et al., 2025). They were as truth biased as non-prisoners and surprisingly, reported less lenient attitudes toward deception (Turi et al., 2025).

This is particularly relevant because studies have shown that inmates' attitudes toward dishonesty are closely linked to behavior outside of prison. For instance, a study examining the attitudes of prison inmates toward dishonest behaviors found significant correlations between these attitudes and parole violations after release (Lilienfeld et al., 1994). This suggests that inmates who internalize the convict code, values, and norms that often conflict with societal standards, are more likely to repeat antisocial behaviors upon reintegration into society (Gendreau et al., 1996). In another study, Cohn et al. (2013) used a coin-tossing task to examine dishonesty among inmates. They found that dishonesty was correlated with violations of in-prison regulations, such as aggression and weapon possession. Furthermore, inmates who believed cheating was detectable were less likely to engage in dishonest behavior, and their attitudes toward dishonesty predicted their risk aversion.

These findings suggest that while inmates may be willing to use self-serving lies, their decision to do so is influenced by the perceived advantages and risks of lying in a given context. In this context, the present study aims to explore, for the first time in the literature to our knowledge, the personality traits, beliefs, narratives, attitudes, and self-reported behaviors related to deception and various types of lies among incarcerated individuals. Using a mixed-methods approach, we seek to leverage the strengths of both quantitative and qualitative methodologies to illuminate the potential of narratives in shaping favorable or unfavorable

[4] The content of this sub-chapter (pages 29-40) overlaps with the manuscript: "Through the Looking Glass: A Mixed Methods Analysis of Inmates' Narratives Shaping their Attitudes Towards Interpersonal Deception" written by Turi, A. & Visu-Petra, L. (2025) and submitted for publication [Manuscript under review].

attitudes toward deception and the likelihood of engaging in deceptive behaviors. To achieve this goal, we formulate four main research questions:

Q1: H*ow do incarcerated individuals define deception and how do they justify, and categorize different types of lies?*

Q2: *What is the degree of endorsing favorable or unfavorable attitudes towards interpersonal deception and what are the justifications behind these attitudes in incarcerated individuals?*

Q3: *What is the self-reported frequency of overall lying and lying in different situations among incarcerated individuals?*

Q4: *What are the interrelationships between attitudes towards various types of lies and measures of aversive traits (dark personality, aggression, alexithymia)?*

3.3.1. Method

To achieve our study aim, we developed a QQ protocol based on questions extracted from an interview on the typology of deception (Bryant, 2008) and quantitative methods investigating attitudes towards types of lies (Lundquist et al. 2009) to address our research questions. This individual written approach was preferred given the sensitivity to the topic and to minimize the effects of social desirability tendency that may have been more pronounced within individual interviews or focus groups, providing participants' time to think and express their beliefs privately. The protocol contained both generous spaces for free text and multiple-choice responses (See Appendix 1 for the QQ Protocol administered to inmates).

Participants

The research protocol was completed by **42 incarcerated individuals** who agreed to participate in the study (21 females) with ages ranging from 23 to 67 years; $M = 40.55$, $SD = 10.51$) and returned to the investigators after 3 days.

Procedure

Participants' responses were analyzed using a ***Thematic Analysis*** (TA, Braun & Clarke, 2006) by the two investigators working together. The process of analyses began with one investigator, reading the data multiple times to identify relevant patterns and code the central features of the data. Next, both investigators revised the codes and identified the main themes based on the proposed theoretical framework, while also remaining open to new meanings and interpretations of the unique aspects of inmates' perspectives on lying.

3.3.2. Results and Conclusions

The results of the analysis to Q1 and Q2, are presented in the tables below, along with examples of participants' answers (translated into English).

Q1: How do incarcerated individuals define deception and how do they justify, and categorize different types of lies?

Table 1

Topic 1: What is deception*?*

Q1. What is deception? (Themes)	**Subjective definitions of deception (participant's answers)**

Falsehood	• *"A lie means an untruth, an invention, a fabrication." (A2)* • *"Untruth. The deliberate concealment of the truth." (E4)* • *"A modified truth." (F4)* • *"A lie = a word, a phrase that is not true, presented and shown to the public as a big truth." (MC4)* • *"I believe that a lie is the opposite of truth." (MF1)*
Individual Trait	• *"To me, a lie means a lack of personal responsibility." (A1)* • *"Dishonesty, a lack of accountability, a lack of trust." (MA1)* • *"Deception makes you a nobody, a man of no one." (MD1)* • *A trait of selfish, hypocritical people, but sometimes also of protective people." (B1)*
Strategy	• *"Intentional avoidance of the truth, aiming to hide an act or decision that would be unfavorable to me." (MC1)* • *A solution (which is best used only when the situation demands it, to ensure a beneficial outcome for you, your family, close ones, etc.). MD3)* • *"An 'efficient' way to get out of certain extreme situations." (MD2)*

Note. Key themes and examples of participants' answers (individual codes in parentheses) to the question: "*Please define the term 'deception' in your own words. Write down whatever comes to mind. What does lying mean to you?*"

Table 2

Topic 2: Why deception occurs?

Q1. Why deception occurs? (Themes)	**The most important reasons behind people's deceptive behavior (participants' answers)**	**Type of lie**
Protect others or avoid negative consequences.	• "*To avoid hurting loved ones."* (E3) • "*They lie so they don't offend others*." (MC4) • "*To avoid a conflict or an unpleasant situation."* (F3)	OI: high SI: high
Instant Gratification Manipulation for Self-Gain Social Desirability	• "*To create an easier advantage, to avoid and shorten the correct path."* (B1) • "*To obtain something, to get on someone's good side by lying."* (C3) • "*They lie to achieve their goals."* (A4) • "*To satisfy different needs, whether material, informational, or emotional."* (D4) • "*People may lie to attract a loved one (for example, sex in return for favors), to become rich, they pretend to be who they're not, attempting to gain more and more without working."* (MD1) • "*To appear honest, to stand out."* (MA2) • "*Many people lie out of fear, insecurity, or lack of involvement."* (A2) • "*People lie because they're afraid of what others might think of them*." (MF3) • "*People might lie from self-doubt."* (A1) • "*Out of fear*." (C1)	OI: low SI: high

Personal Enjoyment Habit of Lying	• "*Out of habit*." (E4) • "*Some lie because they're used to it or simply like to embellish things*." (F4) • "*Others lie without any reason … they're sick, etc*." (ME3) • "*For some people, lying is a chronic disease*." (MA4) • "*There are people who lie just for personal amusement*." (MF4) • "*Some people lie for pleasure (clear purpose to harm others); for them, lying is a way of life*." (A2)	OI: low SI: low

Note. Key themes and participants' answers (individual codes in parentheses) to the question: "*Why do people lie? Which cause do you think is the most important?*". OI = other interest, SI = self-interest.

Table 3

Topic 3: When is lying justified?

Q1. When is lying justified? (Themes)	Examples of situations in which lies are justified (participants' answers)	Type of lie
Help, save, or protect others. Create positive emotions	• "*A justified lie is one that helps someone, that saves someone from a conflict*." (MD2) • "*When you solve a problem or situation for someone else and don't boast about it*." (E1) • "*To save someone's life, to avoid hurting, to avoid harming someone*." (D1) • "*I think that sometimes even delaying the delivery of bad news equates to a lie. It's justified by that feeling of wanting to protect, to spare the person receiving the news from some of the suffering."* (B1) • "*I suppose that lies are justifiable when you want to do good. But deep down you know it's wrong*." (C4) • "*A justified lie is when you're preparing a surprise for loved ones and need to find out details or invent something to make the surprise happen*." (MA2)	OI: high SI: low
Prevent conflicts and negative events. Relationship maintenance	• "*To prevent two people from fighting or something worse*." (A2) • "*Yes, when you can avoid a big scandal or something very serious*." (F2) • "*When you want to cover for someone, to help, to keep other people out of trouble*." (C3) • "*When telling the truth would trigger an event or a conflict with serious consequences, potentially even murder or a grave catastrophe*." (MC1) • "*I lie to my family every day about the conditions in the prison*." (ME2) • "*Sometimes you're forced to lie to avoid hurting loved ones*." (E3)	OI: high SI: high
Avoid taking responsibility.	• "*Yes…! When you or your loved ones stand to gain money without hurting anyone … when you save yourself or another person … when you avoid prison*!" (ME3)	OI: low SI: high
-	-	OI: low SI: low

Note. Key themes and participants' answers (individual codes in parentheses) to the question: "*Are lies sometimes justified? Can you give examples of such situations?";* OI = other interest, SI = self-interest

Table 4

Topic 4: What are white lies?

Q1. What are white lies? (theme)	Responses that illustrate the themes (participants' answers)	Type of lie
Lies told to protect others	• "*You go to visit someone terminally ill in the hospital. You have to smile, restore their confidence in life, maybe even let them dream of recovery, brighten up perhaps their final moments*." (A2) • "*When my brother died, my mother didn't tell me to protect me*." (A4)	OI: high SI: low
Lies told to navigate complex social landscapes.	• "*Yes, for example, when you don't want to respond but have to say something so that no one is affected."* (C4) • "*Returning a compliment."* (D4) • "*When you don't want to get involved in others' situations and are asked if one person spoke badly about another, you just say you don't know to avoid getting involved*." (A3) • "*A behavior directed at someone like: 'You look good today,'" "you've lost weight,' 'What a nice outfit you bought*.'" (MD2) • "*I lied when I said: 'you look great today' because I felt they needed the appreciation to lift their spirits, seeing they were having a bad day*." (B1) • "*The most common white lie in my opinion is 'I'm fine' when asked, 'How are you?*'" (MC3) • "*To create a good atmosphere in a group or family setting*." (G78) • "*Yes, a white lie is a kind of lying without lying. For example, Santa Claus brings gifts, but in fact, there's no Santa Claus*." (MF1) • "*When you're planning a surprise for a loved one*." (MA2)	OI: high SI: high
Lies told to make excuses, conceal private things, or boost social image.	• "*A white lie is when someone asks me for something, and I say I don't have it. Actually, I do have it but just don't want to give it to them*." (A1) • "*Possibly when you're not in the mood for 'colleagues' and say you have a letter to write or a book to read*." (E1) • "*People who lie harmlessly are sometimes just trying to stand out, to be the center of attention. When someone wants to be the center of attention, they attribute qualities to themselves that they don't have, trying to impress; often it's by claiming certain qualities, events, or deeds*." (MA3) • "*I'm supposed to meet someone at a set time, but something more important comes up*." (MA4) • "*I'm coming from work' (when actually coming from a drink with friends)."* (MC4) • "*You were supposed to go somewhere but didn't want to, so you call and say you're sick."* (ME4)	OI: low SI: high

Note. Key themes and participants' answers (individual codes in parentheses) to the question "*Are you familiar with the term "white lie" (also known as an innocent lie)? Can you give an example of a white lie?";* OI = other interest, SI = self-interest.

Table 5

Topic 5: What distinguishes white lies from real lies?

Q1. What distinguishes white lies from real lies? (theme)	Illustrative quotes (participant's answers)
Lie acceptability	• "*White lies are more easily acceptable, while normal lies are harder to accept.*" (A1) • "*There are differences between the two types of lies, starting with their purpose and the quantitative and qualitative analysis of the harm or good produced by that white lie. I do not accept normal lies, those that come from malice; they seem dishonorable to me. I accept lies told out of kindness, those with a favorable result, harmless lies that bring a bit of value to someone's spirit.*" (B1) • "*A white lie is acceptable when it doesn't produce negative effects.*" (D1)
Consequences of the lies	• "*In the case where we realize it is a lie, its acceptability depends on how it affects us and the consequences, regardless of whether it's a normal or a white lie.*" (F1) • "*Yes, the difference is that a normal lie can have serious consequences compared to a white lie.*" (D4) • "*The differences are very large, as [a normal lie] can even lead to death. It seems acceptable only up to a point, but if it gets out of hand or repeats, that's it. White lies are only acceptable when they are jokes, humor, anecdotes.*" (MD1) • "*Wars have been won with the help of white lies.*" (ME2)
No difference	• "*Even if we lie for different reasons, I don't think it's acceptable. But if they aren't told to hurt someone, white lies are more acceptable. Still, I don't believe there are differences between the two types of lies.*" (D1) • "*In the end, neither type is acceptable, but sometimes we do it out of convenience.*" (A3) • "*There is no acceptable lie.*" (E4) • "*No one accepts lying, but we all do it, whether we like it or not.*" (MD2)

Note. Key themes and participants' answers (individual codes in parentheses) to the question, "*How acceptable do you think lying is? From your perspective, do you believe there is a difference between how acceptable a normal lie is compared to a white lie? In what situations are white lies acceptable*

Q2: What is the degree of endorsing favorable or unfavorable attitudes towards interpersonal deception and what are the justifications behind these attitudes in incarcerated individuals?

Table 6

Attitudes Towards Lies and Acceptability of Lying by Lundquist et al. (2009).

No. item	Attitude	N	Endorsement (%)	Key Beliefs (theme and description)	Illustrative quotes of positive (+) and negative (-) attitudes (participant's answers)
1.	*"A small lie that makes someone feel good is*	40	**+ 37.5%** **- 62.5%**	**Key Beliefs:** Small lies can provide short-term emotional boosts but may have long-term costs.	+ • "*In general, an insignificant lie that adds something positive to someone's life improves their day and can put*

	always a good thing."			**Description:** Participants are divided on whether "small lies" that make others feel better are acceptable. Some view them as harmless and even helpful in encouraging others, while others argue that small lies might damage relationships if uncovered, creating distrust and a habit of dishonesty.	*them in a better mood, so it's not a bad thing.*" (B1) • "*Because life is ugly anyway. At least you can tell an overweight person that they've lost weight, encouraging them to put in more effort.*" (D3) - • "*It's not a good thing because, at that moment, they might feel good about being lied to, but when they find out I lied, they'll be upset. So, it's better to tell the truth*". (A1) • "*Although that lie makes the person feel good, it burdens your conscience and gradually affects your character, getting you used to lying more and more often*" (MC3)
2.	*"I would be willing to lie to help someone, even if it would negatively affect me."*	40	**+ 55%** **- 44.5%**	**Key Beliefs:** Willingness to lie for others varies, with concerns about personal costs. **Description:** Opinions are nearly split on whether lying to help someone else is acceptable. Those who approve of such lies cite a desire to help loved ones, sometimes at their own expense. However, others emphasize the potential personal disadvantages, such as harming one's reputation or facing social rejection.	+ • "*Because I like to help the people around me, even if it's to my own disadvantage*". (E3) • "*I would agree to lie for the health or life of my children, my partner… no matter what happens to me afterward…*" (ME3) - • "*Because in the future, this will put me at a disadvantage with others.*" (MC1) • "*Because even though I'm helping someone, I end up losing in two ways: first through the lie itself, and then when the truth eventually comes out*". (MF4)
3.	*"The more I stand to gain from a lie, the more tempted I am to lie."*	40	**+ 20%** **- 80%**	**Key Beliefs:** Lying for personal benefit is largely condemned, as it can damage trust. **Description:** Most participants disapprove of lying for personal gain, viewing it as morally questionable and harmful to interpersonal trust. While some acknowledge that lies can offer easy, fast rewards, they believe these benefits are short-lived and often lead to greater losses in the long run.	+ • "*It's human nature to want to obtain more and to do so more easily*". (MB3) - • "*You can't gain anything from lying. Instead, you risk losing people's trust*". (F3) • "*Even if I gain something from a lie, it would only be temporary. The more I stand to gain from a lie, the more I ultimately lose*". (D1)

4.	*"If there's a chance I might get caught, I'd rather not lie."*	40	**+ 85%** **- 15%**	**Key Beliefs:** Fear of consequences discourages lying. **Description:** A strong majority agree that they would avoid lying if there's a risk of being caught, citing potential shame and the mental strain of maintaining a lie. A few participants, however, might lie if high-stake benefits outweigh the risks, such as escaping severe consequences.	+ • "*The shame you feel when you're found out*". (MD2) • "*For my own peace of mind and conscience, and so that it won't be worse in the end*". (MF2) - • "*If I had lied, I wouldn't have ended up in prison, but I didn't do it*". (MB4)
5.	*"The greater the risk of being found out, the less tempted I am to lie."*	39	**+ 89.7%** **- 10.3%**	**Key Beliefs:** High risk of exposure discourages lying due to possible social exclusion and loss of credibility. **Description:** Many of the participants fear social exclusion and damage to their credibility, while a minority would still consider lying if the rewards are substantial enough to justify the risk.	+ • "*If you don't have the talent and skill to lie and you're caught, the person or group will marginalize you*" (MA4) • "*Surely, we lose credibility with others, and there will be situations where we tell the truth, but we won't be believed*". (MD4) - • "*If the consequences aren't worse than the benefits, I keep going until the end, even if I have only a slim chance (1 in 100) of making it*". (ME3)
6.	*"A small lie is still a lie."*	39	**+ 100%** **- 0%**	**Key Beliefs:** Participants largely believe all lies, regardless of size, are morally equal. **Description:** This view reflects a strict moral stance that small lies may lead to a habit of dishonesty and are equally wrong as larger lies.	+ • "*There's no big lie or small lie. It's all just a lie*". (A3) • "*I completely agree with this statement, first because "a small lie" leads to more, and secondly, because I don't agree with lying in any form*". (MD3)
7.	*"If I promised someone, I'd tell the truth, it's very hard for me to lie to that person."*	39	**+ 100%** **- 0%**	**Key Beliefs:** Keeping one's word is tied to personal honor and trustworthiness. **Description:** All participants value honor and trustworthiness and breaking a promise of honesty feels morally wrong.	+ • "*I promise to tell the truth, I do it because to me, a given word and a kept promise mean honor*". (D1) • "*Because I don't want to lose the trust that person has given me*". (E3)
8.	*"I lied when I did something I didn't want*	38	**+ 73.7%** **- 26.4%**	**Key Beliefs:** Lying is acceptable when concealing private matters, though some prefer accountability.	+ • "*Not everyone needs to know what I'm doing*". (D2)

	others to find out about."			**Description:** The majority feel justified in lying to protect personal privacy, especially if the lie doesn't harm others. A notable minority believes in taking responsibility for their actions and prefers honesty over concealment.	• "*If my actions don't affect others, I don't consider not telling them to be lying*". (MB3) - • "*I don't hold back; if I want to do something, I do it, and if others don't like it, I prefer they keep their opinions to themselves. I take responsibility for what I say and do*". (F3)
9.	*"I lied when I was questioned by the authorities. "*	38	**+ 44.7%** **- 55.2%**	**Key Beliefs:** Participants are wary of lying to authorities but may do so out of fear. **Description:** Slightly more participants reject lying to authorities, citing honesty as beneficial in the long term. However, nearly half admit to lying in this context, usually to avoid negative consequences.	+ *of fear*". (MF2) - *n't lie, so I received a smaller sentence, and ocess went very quickly*". (MB4)
10.	*"I lied when I turned down a friend who asked me for help."*	39	**+ 35.9%** **- 64.1%**	**Key Beliefs:** Honesty is preferred over lying to friends, as it strengthens respect. **Description:** Most participants avoid lying to friends, believing honesty is better appreciated, even if it results in refusal. A smaller group finds it acceptable to lie, often due to personal grievances or conflicting priorities.	+ • "*There are times in life when other priorities take precedence*". (MC1) - • "*I find it hard to refuse a friend who asks for help, but if that help would require me to do something wrong or beyond my abilities, I prefer to tell them the truth behind my refusal*". (B1) • "*I believe honesty is much more appreciated*". (MF3)

Note. Key themes, description of the theme, and participants' answers (individual codes in parentheses) to the items of the self-reported questionnaire *Attitudes Towards Lies and Acceptability of Lying* by Lundquist et al. (2009).

Q3: What is the self-reported frequency of overall lying and lying in different situations among incarcerated individuals?

The self-reported frequency of lying behaviors (ranging from "Never" to "Frequently" and referring to the present occurrence) across various relational contexts was assessed to determine if inmates' actual behaviors aligned with their beliefs and attitudes toward lying. Our findings revealed that inmates reported a notably *low frequency of lying* across various contexts, indicating that honesty was generally a valued norm within their social interactions. Inmates were least likely to lie to *authority figures*, with the majority reporting that they never or rarely lie in these situations, possibly reflecting a pragmatic approach to maintaining favorable outcomes and avoiding potential repercussions. A similar pattern of infrequent lying was observed in *intimate relationships*, suggesting respect for honesty within close personal connections, though half of the participants acknowledged occasional or rare lying. When it came to interactions with *friends, family*, *and fellow inmates*, lying was also largely reported as a rare occurrence. These relationships appeared to allow for slightly more flexibility in truthfulness, perhaps as a strategy to preserve harmony or navigate social dynamics. However, lying for *personal enjoyment* was nearly absent, with most inmates asserting they never lie purely for amusement, pointing to general disapproval of lying without purpose or necessity. Specific percentage rates are presented in Table 9.

Table 7

Self-reported lying frequency across everyday life situations

Lying frequency	N	Never	Rarely	Sometimes	Frequently
Overall lying frequency	38	28.9%	52.6%	18.4%	0
Lying to friends or family	40	35%	47.5%	17.5%	0
Lying to authority	37	56.8%	32.4%	10.8%	0
Lying for personal enjoyment	39	87.2%	10.3%	2.6%	0
Lying to an intimate partner	40	50%	42.5%	7.5%	0
Lying to other inmates	39	33.3%	59%	7.7%	0

Note. Self-reported overall *lying frequency* was measured through item 11 from *Attitudes towards Lies and Acceptability of Lying* (Lundquist et al., 2009)

Q4: Interrelations between attitudes towards lies and aversive traits.

As presented in Table 8, Lie acceptability, the primary variable in our study was positively correlated with the Dark Factor of Personality, $r(42) = .43$, $p = .004$, even after controlling for Impression Management, $r_p(42) = .35$, $p = .026$. This relationship was particularly strong for Deceitfulness, $r(42) = .49$, $p = .001$, Narcissistic Entitlement $r(42) = .33$, $p = .033$, and Vindictiveness, $r = .49$, $p < .001$. However, the association with Narcissistic Entitlement was no longer significant when controlling for Impression Management, $r_p(42) = .27$, $p = .091$.

Table 8

Correlations (and partial correlations) between personality variables and lie acceptability.

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1.	Lie Acceptability	1	**.35***	.09	**.43****	.27	.19	**.46****	.20	.10	.21	-.05	.26	-.06	-.13	-.07
2.	Dark Factor (D)	**.43****	1	.80**	.82**	.75**	.83**	.79**	.38*	.27	.60**	-.28	-.07	-.54**	-.31*	-.33*
3.	*Callousness*	.20	.85**	1	.50**	.59**	.60**	.48**	.37*	.24	.55**	-.33*	-.22	-.55**	-.25	-.26
4.	*Deceitfulness*	**.49****	.82**	.54**	1	.56**	.64**	.68**	.25	.25	.48**	-.16	-.09	-.41**	-.31	-.38*
5.	*Narcissistic Entitlement*	**.33***	.77**	.65**	.57**	1	.55**	.44**	.07	.20	.30	-.31	.19	-.53**	-.10	-.23
6.	*Sadism*	.27	.81**	.61**	.71**	.55**	1	.47**	.30	.12	.48**	-.24	-.02	-.30	-.26	-.19
7.	*Vindictiveness*	.**49****	.83**	.64**	.64**	.53**	.47**	1	.41**	.28	.52**	-.11	-.10	-.43**	-.30	-.29
8.	Alexithymia	.27	.48**	.47**	.34*	.18	.37*	.50**	1	.36*	.39*	.01	-.37*	-.26	-.37*	.01
9.	Aggression	.16	.35*	.31	.34*	.25	.21	.36*	.41**	1	.46**	.03	-.14	-.42**	-.08	-.14
10.	Neuroticism	.23	.59**	.57**	.42**	.36*	.43**	.55**	.42**	.45**	1	-.18	-.26	-.36*	-.41**	-.29
11.	Extraversion	-.11	-.36*	-.40**	-.23	-.36*	-.28	-.22	-.07	-.02	-.22	1	.29	.18	.34*	.09
12.	Openness	.17	-.18	-.28	-.20	.11	-.11	-.18	-.41**	-.19	-.26	.32*	1	.20	.49**	.13
13.	Agreeableness	-.14	-.60**	-.60**	-.47**	-.57**	-.37*	-.49**	-.34*	-.46**	-.38*	.24	.25	1	.25	.21
14.	Conscientiousness	-.19	-.37*	-.30	-.37*	-.15	-.32*	-.33*	-.41**	-.14	-.40**	.37*	.52**	.30	1	.30
15.	Self-deceptive Enhancement	-.10	-.31*	-.23	-.42**	-.21	-.25	-.21	-.01	-.19	-.23	.10	.16	.23	.33*	1
16.	Impression Management	-.24	-.41**	-.28	-.51**	-.18	-.41**	-.29	-.24	-.27	-.01	.16	.24	.26	.22	.22

Note. N = 42; p < .05 = *; p < .01 = ** Correlations above the diagonal are conducted while controlling for Impression Management.

Taken together, these results (see Figure 1 for a summary) illustrate that inmates' beliefs and attitudes toward deception are shaped not only by moral considerations but also by the strategic imperatives of their environmental challenges. Deception for them is not merely a matter of personal morality but a complex interplay of social adaptation, personal survival, and reputational concerns. This complexity challenges simplistic views of dishonesty as inherently good or bad, suggesting instead that the acceptability of lying is deeply embedded in the social and institutional contexts in which it occurs. These narratives and attitudes towards deception are reflected in inmates' self-reported frequency of lying across various contexts (Q3), suggesting that they limit deception to specific situations where it is perceived as necessary. Previous research suggests that the more seriously individuals perceive deception, the less likely they are to engage in it (Ennis et al., 2008). This was evident in our study, where inmates' strong stance against deception was reflected in their relatively low self-reported frequency of lying.

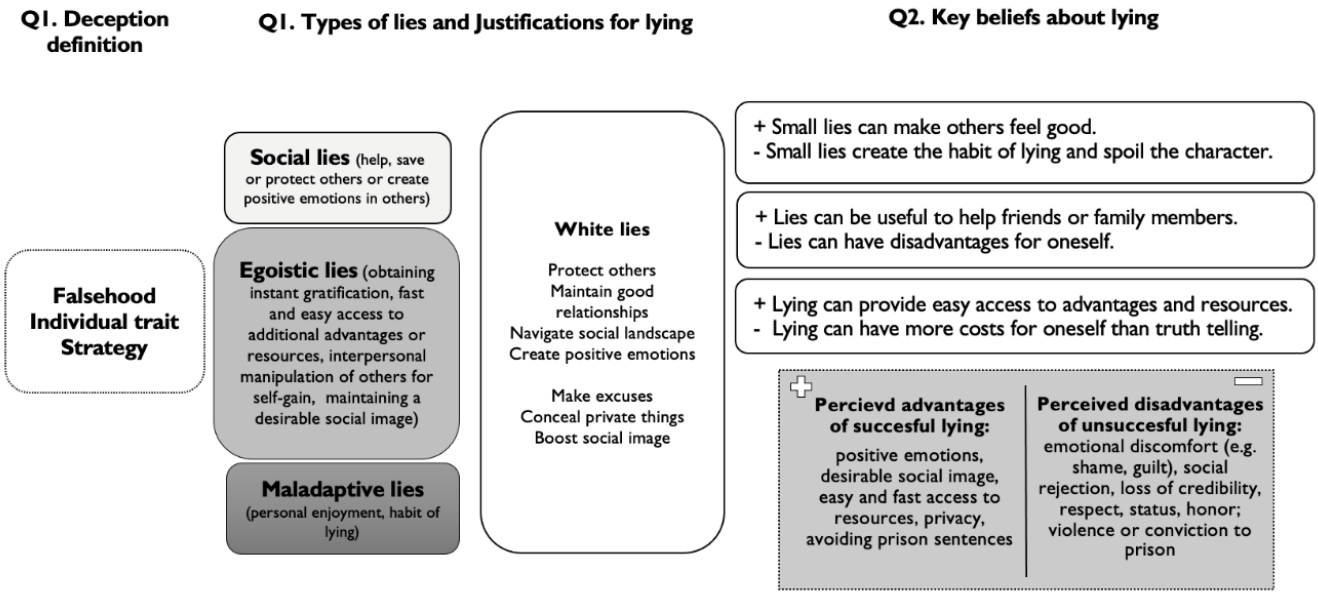

Figure 1. The thematic map presenting the themes identified in Q1 and Q2 (deception's definition, types of lies, use of white lies, and summarization of the positive (+) and negative (-) attitudes towards small lies, real lies, and motives for using or not using deception supported by the possible advantages and disadvantages of lying).

Finally, the attitudes towards deception of the prisoners were related to a variety of aversive traits (Q4), ranging from dark personality to alexithymia or aggression. Only the Dark Factor of personality had a significant association with attitudes towards deception, especially the Deceitfulness and Vindictiveness themes, which remained significant even when controlling for their social desirability (unlike Narcissistic Entitlement which was no longer significant after adding this control). This suggests that inmates higher in their propensity to deceive and be vengeful are more lenient towards using lies as a tactic for reaching their goals.

3.4. Study 3B: A Mixed-Methods Investigation of Inmates' Beliefs About Lie Detection[5]

Several studies suggested that prisoners possessed less stereotypical and more accurate beliefs about the cues than indicate deception (Vrij & Semin, 1996; Granhag et al., 2004; Hartwig et al., 2004; Strömwall et al., 2004), as they were more accurate lie detectors compared to other groups of people, including correctional staff, students, and prison personnel. Moreover, they were more likely to use verbal cues to identify liars, which are believed to be more reliable (Hartwig et al., 2004; Bogaard & Meijer, 2018). Similarly, Jupe et al., (2016) showed that juvenile offenders exhibited greater accuracy in credibility judgments compared to their non-offending peers. However, conflicting evidence challenges the notion that prisoners are inherently better at detecting deception. While some studies suggest that inmates are skilled at identifying liars, they may also be prone to a lie-bias (i.e., a tendency to assume deceit even when none is present, Bond et al., 2005). Some studies have shown that prisoners perform no better than students in accurately classifying true and false statements, suggesting that their abilities may not be superior to those of the general population (Schindler et al., 2021).

Given these mixed findings, there remains a lack of consensus regarding whether inmates are indeed better at detecting lies, or predominantly rely on stereotypical beliefs about deception. Although a few studies have investigated prisoners' beliefs using quantitative methods, such as self-reports or surveys (e.g., Hartwig et al., 2004; Vrij & Semin, 2006), these studies did not allow them to freely express their beliefs and comment on their deception detection strategies. The pioneering qualitative research focused primarily on inmates' beliefs about deception or the strategies they use, revealing that inmates consider the ability to deceive and manipulate as useful assets and may use them to mask hurtful emotions and foster self-preservation (Laws, 2016). These findings align with Maruna and Butler's research (2009; 2013), which emphasizes the role of personal rationalizations and narratives in shaping inmates' propensity for engaging in aversive behaviors, such as manipulation or aggression. Their work suggests that prisoners often construct self-narratives that justify manipulative or aggressive behaviors as necessary responses to perceived disrespect or social threat.

The current study aims to extend the limited literature on this topic, by exploring prisoners' beliefs about deception detection through a mixed-methods approach, addressing their opinions about four main questions:

- **Topic 1:** How good are people at detecting deception and do some lies pass undetected?
- **Topic 2:** What are the ingredients of successful lying and the characteristics of good liars?
- **Topic 3:** What are the cues that help deception detection? (freely reported by inmates).
- **Topic 4:** What is the degree of endorsing stereotypical beliefs about the real indicators of deception? (questionnaire report).

[5] The content of this sub-chapter (pages 48-55) overlaps with the manuscript: "Unveiling Deception Behind Bars: A Mixed-Methods Investigation of Inmates' Beliefs About Lie Detection" written by Turi, A. & Visu-Petra, L. (2025) and submitted for publication [Manuscript under review].

3.4.1. Method

In this study we used the same methodology from Study 3A, on the same sample of **42 inmates** (21 females), aged between 23 and 67 years (M = 40.55, SD = 10.51).

However, we focused exclusively on the multiple questions related to deception detection extracted from the QQ Protocol Part 1. In addition, we included in our analysis participants' responses to a self-report questionnaire designed to measure stereotypical beliefs regarding seven categories of cues to deception detection (Strömwall & Granhag, 2003; Hartwig et al. 2004 - generating QQ Protocol Part 2). This questionnaire has been completed right after participants experienced deception detection and being deceived in the experimental task DeceIT (Wright et al., 2012; see Turi et al., 2025 for the entire data set).

3.4.2. Results and Conclusions

The analysis of incarcerated individuals' perceptions regarding our research topics together with answers to the questions are presented in the tables below.

Topic 1: How good are people at detecting deception and do some lies pass undetected?

Table 1

Frequencies and percentages for the inmate's beliefs regarding people's ability to detect deception and undetected lies.

Theme	N	%	Codes	Example quotes (participant's answers)
Most lies are detected	19	45.23%	▪ Nonverbal cues and gut feelings might improve detection. ▪ People detect lies but prefer to oversee them to avoid conflicts. ▪ Lies eventually get discovered	• "*I believe people notice when they are being lied to, but they choose to overlook it to avoid a fight.*" (F2) • "*In principle, people realize when they are being lied to, although sometimes it goes unnoticed, but eventually, or after some time, it is discovered."* (MC1) • "*The higher the level of education, the easier it is to spot a lie."* (B1)
Sometimes lies get detected, other times they pass unnoticed	12	28.57%	▪ People tend to default to the truth. ▪ The social status of the liar matters in credibility. ▪ Depends on the individual characteristics of both parties and the type of lie.	• "*Sometimes they realize it, sometimes they don't. When you trust someone and believe them to be honest, you tend to believe what they say until proven otherwise."* (A3) • "*It depends on who is saying it and who is listening... If the lie is told by magistrates, leaders, or people in authority, it is often believed."* (ME3) • "*In most cases, yes, but some lies may go unnoticed if they are of minimal importance."* (MA1) • "*It depends on how much attention you pay or the type of lie."* (MB3)
Lies often pass unnoticed	10	23.80%	▪ People believe lies more often than truths. ▪ Depending on the liars' skills.	• "*Unfortunately, lies go unnoticed, and people are more willing to accept a lie than the truth."* (D1) • "*A lot of the time, people believe lies because the truth can be harsh*." (D3)

				• "*In most cases, people do not realize when they are being lied to. A concrete example is the large number of fraud cases that end up in court.*" (MD3)

Note. From the total sample (N = 42), 2.5% of responses were missing or inconclusive. Codes present participant's beliefs regarding the factors that influence deception detection.

Topic 2: What are the ingredients of successful lying and the characteristics of good liars?

Table 2

Frequencies and percentages for the inmate's beliefs about the ingredients of successful lying and the characteristics of good liars

Theme	N	%	Codes	Example quotes (participant's answers)
Liar's behaviors	15	35.71%	▪ Use verbal tactics such as repeating, and detailing. ▪ Using justifications, arguments, and truths ▪ Use gestures, maintain eye contact, and control voice tone. ▪ Avoiding direct answers, changing the subject, and fabricating evidence.	• *"They try to give as many details as possible and arguments, even if they repeat them sometimes." (MA1)* • *"They should not insist, not repeat, look the person in the eyes, and have the same reactions as if they were telling the truth." (ME3)* • *"A liar will go into detail to give credibility to their words or will sprinkle some truth into their lies, which will make them credible." (MF4)*
Liar's traits	16	38.09%	▪ Believe their own lies. ▪ Are persuasive. ▪ Are willing to do anything. ▪ Have extensive experience in deception	• *"They are very persuasive; they begin to live what they say." (A4)* • *"They are experts in arguing falsehoods." (A2)* • *"They have great persuasive power." (F4)* • *"They are willing to do anything." (E3)* • *"They must have vast experience or be a born liar." (MA4)*
Various Traits & Behaviors	4	9.52%	▪ Have a good memory. ▪ Carefully manage their emotions and non-verbal behavior ▪ Use emotional appeals and psychological tactics to mislead others	• *"Chronic liars generally have a good memory and know very well how to control their emotions, non-verbal language, hand gestures, and eye contact." (MD2)* • *"They make you enter their mind, into their story, and act on that sensitive string we all have: 'pity.'" (MF2)* • *"They try to manipulate others with the arguments they make, and even if it's not true, their nature is not made to regret." (MD4)*
Other Aspects	5	11.90%	▪ People are too trusting or do not verify information. ▪ High occurrence of lying ▪ The receiver is uninterested or lies cannot be checked.	• *"The lie will go unnoticed if it is not exaggerated and has no unnecessary details." (MF4)* • *"Either they give the impression that they really know what they are talking about, or others trust them." (ME4)* • *"Probably, lying is a daily occurrence for some." (A3)*

Note. From the total sample (N = 42) 4,76% of responses were missing or inconclusive.

Topic 3: What are the cues that help deception detection? (freely reported by inmates).

Table 3

Frequencies and percentages for all categories of cues used by inmates to identify deception in others.

Theme (total % of endorsement)	Specific cues (codes)	N	%	Example quotes (participant's answers)
A mix of verbal and nonverbal cues	changes in facial expressions, including gaze aversion, changes in speech patterns, and more body movements	17	40.47%	• "*The way of expressing oneself, facial expressions, the desire to argue as much as possible*." (MA1) • "*Avoiding eye contact, body language, lack of obvious details, avoiding certain aspects."* (MA3) • "*If he looks into my eyes but at the same time is agitated. Nervous, restless, uneasy. He presents reality as a beautiful story."* (D3)
Nonverbal cues	avoiding eye contact	24	57.14%	• "*If they avoid my gaze, I know they are lying to me*." (D1) • "*First of all, when someone is lying to you, they don't look you in the eye, they avoid eye contact."* (A3) • "*Avoiding or not maintaining eye contact, with eyelids widening and a more fixed gaze than usual."* (MD2)
	touching the face, nose, fidgeting, posture shifts, various hand gestures	15	35.71%	• "*Usually, a person who is lying will touch their nose or mouth (an unconscious gesture covering the mouth that tells lies)."* (B1) • "*Widening of the eyelids and a more fixed gaze than usual, bringing the hand to the nose or rubbing the eyes, lowering the shoulders and leaning the hands forward with faster gestures, in children, a slight tilting of the head."* (MD2)
	emotional reactions, such as nervous laughter, panic, fear, or physiological signs of nervousness, such as babbling, trembling	13	30.95%	• "*They are agitated. Nervous, restless, uneasy*." (D3) • "*Some laugh, some panic, some avoid eye contact*." (D2) • "*They blush, their chin tenses, and they tremble*." (MF3)
	changes in face mimic (usually associated with other movements or gestures)	10	23.80%	• "*Facial expressions*." (A3) • "*I look at his gaze, observe the gestures he makes, and I can tell from his facial expressions*." (A4) • "*Through gestures, facial expressions, and voice*." (E3)
Verbal cues	telling the story in different ways, altering details through avoiding or exaggerating certain aspects, not providing	14	33.33%	• "*I would realize that someone is lying if a version is told in different ways*." (F1) • "*When the false statement is not supported by any demonstration of truthfulness*." (B1)

	arguments or evidence to support the lie			• "*They get flustered when they have to say the same thing a second time, slightly changing what they originally said*." (D1)
	liar's way of talking, changing the subject, repeating, over-explaining, justifying, arguing, jumping between contexts, praising, promising, even cursing or using flattery or trying to convince	13	30.95%	• "*They flatter, repeat the same thing many times."* (MA2) • "*They talk a lot*." (MB4) • "*Through their way of expressing themselves, compliments, promises*." (MA4) • "*They jump from one context to another, gesturing a lot, using a lot of insults*." (ME1)
	changes in voice tone, voice inflections, way of talking	5	11.90%	• "*A person who lies changes their tone of voice, sentence structure, and intonation*." (MC3) • "*Their tone of voice changes*" (ME4)
Other	receiver's gut feeling sender's interests, self-confidence, and previous experience	5	11.90%	• "*Gestures, expressions, attitude, their actions, and presence*."(ME2) • "*Most of all, I listen to the response of my own body to what I've heard. I listen to my own feelings*." (B4)

Note. For each category, the percentage is computed from 100% = 42 respondents, 2 responses (4.76%) were inconclusive.

Topic 4: What is the degree of endorsing stereotypical beliefs about the real indicators of deception (questionnaire report)?

Table 5

Percentage distribution of chosen alternatives to the survey (Hartwig et al., 2004)

Type of cue (item)	A (<) less	B (-) neutral	C (>) more	D (don't know)	Missing	Correct response (in the literature)
Details	**38.1%**	11.9%	40.5%	7.1%	2.4%	< less during deception
Eye contact	73.8%	**2.4%**	11.9%	11.9%	-	no differences
Consistency	**71.4%**	**4.8%**	16.7%	4.8%	2.3%	< less during deception / no differences
Body movements	**9.5%**	9.5%	59.5%	19.0%	2.5%	< less during deception
Pitch of voice higher	9.5%	23.8%	**42.9%**	23.8%	-	> more during deception
Video recordings vs. Face-to-face detection accuracy	**69.05%**	14.29 %	14.29%	-	-	Detecting deceit in face-to-face interactions is harder.
Verbal cues vs. nonverbal detection accuracy	50.0%	31.0%	**4.8%**	9.5%	4.7%	Verbal cues are more reliable.

Note. Italics signal stereotypical beliefs held by prisoners and the general population. Bolded values in the table represent the option documented in the scientific literature (Strömwall et al., 2004).

The findings from this study provide novel and contextually grounded insights into incarcerated individuals' beliefs about deception detection, a group that remains notably underrepresented in deception research. Understanding how inmates conceptualize lying and deception not only expands the theoretical understanding of deception detection across diverse social environments, but also carries important practical implications for institutional management, rehabilitation strategies and social reinsertion of ex-offenders. By shedding light on the beliefs behind cognitive and social mechanisms through which inmates interpret and respond to deceptive behavior, these insights can inform the development of targeted interventions aimed at improving detection accuracy, which in turn, might aid in interpersonal communication, reducing conflict, and fostering prosocial behavior within correctional settings. Moreover, enhancing inmates' understanding of reliable versus unreliable deception cues may contribute to smoother reentry into society by equipping them with more accurate social perception skills and perhaps a better theory of mind, assets in both personal relationships and employment contexts.

Our results support and extend frameworks such as Truth-Default Theory (TDT) and Interpersonal Deception Theory (IDT) by illustrating that detection accuracy is not merely a function of perceptual abilities but might be influenced by contextual and social-cognitive factors, such as institutional norms, perceived threat levels, and interpersonal dynamics. For example, the inmates' reliance on nonverbal cues, despite demonstrating accurate insight into verbal indicators and the limited diagnostic validity of nonverbal ones (Bond Jr & DePaulo, 2006), underscores how environmental pressures - such as the need for rapid threat assessment in overcrowded and tense settings - can bias individuals toward heuristic processing. Such biases, in turn, may entrench stereotypical beliefs about deception indicators, thereby complicating both self-monitoring and the evaluation of others' communicative behaviors.

Furthermore, the interplay between individual differences (e.g., social status, prior experience, and personality traits) and situational variables (e.g., contextual norms, interpersonal trust, and the inmate code) suggests that deception detection is a multifaceted construct. This reinforces the idea that the close to chance level accuracy rates (e.g., 54%) in deception detection are highly context-dependent and should be interpreted within a broader spectrum of cognitive and social influences (Levine et al., 2025; Blair et al., 2010). The convergence of these factors calls for an integrated theoretical model that accounts for both the micro-level processes (such as cue utilization) and the macro-level influences (such as cultural and situational norms) on deception detection. Practically, these findings have significant ramifications for managing social dynamics within correctional settings. Excessive vigilance and reliance on potentially misleading nonverbal cues can lead to a "fight-or-flight" response, in which inmates may misinterpret deceptive behaviors as signs of disrespect or treason (e.g., Butler & Maruna, 2009; Maruna & Butler, 2016). Such misinterpretations can escalate conflicts, weaken social bonds, and further destabilize the already fragile social fabric of prison communities. Correcting entrenched stereotypical beliefs about nonverbal behaviors may not only enhance detection accuracy but also facilitate a more sophisticated understanding of social dynamics. This, in turn, may contribute to reducing conflict and fostering an environment where trust and vigilance are balanced in a manner conducive to both security and rehabilitation.

3.5. Study 4: Dark and Bright Personality Traits and Deception Production or Detection in Prisoners and Community Members[6]

Study 4 explored the complex interplay between personality traits (dark and bright dimensions), deception production and detection abilities among incarcerated individuals and the general population. By employing the modified Deceptive Interactive Task (DeceIT, Wright et al., 2012) for the first time in a prison environment, we examined the relation between ***lie production and detection abilities***, comparing prisoners with civilians in terms of their detectability, detection accuracy, and bias. Based on the study of Wright et al., (2012) we hypothesized that civilians would present a "deception-general" ability, reflected by the moderate correlation between deception production and detection abilities, as well as a pronounced truth bias (*Confirmatory Hypothesis 1a*). Similar to the study of Semrad et al. (2020) on federal police recruits, we had no grounds to expect a similar association between lie production and lie detection abilities in prisoners, yet we aimed to explore their potentially higher lie bias due to increased prison-specific suspiciousness (*Exploratory Hypothesis 1b*). On the contrary, we hypothesized that civilians would exhibit a truth bias (*Confirmatory Hypothesis 1c*), and given the mixed findings in previous studies, we sought to explore the presence of a potential truth bias in prisoners as well (*Exploratory Hypothesis 1d).*

Next, we investigated the relationship between ***dark personality tendencies*** and the ability to produce and detect deception in both prisoners and civilians, testing two contrasting possibilities. Based on the findings of both Wright et al. (2015) and Semrad & Scott-Parker (2020) concerning the Dark Triad traits, we expected no consistent association between the Dark Factor and the ability to lie and detect lies, neither in prisoners nor in civilians (*Exploratory Hypothesis 2a*). Conversely, based on Hurezan et al., (2024) showing higher levels of the Dark Factor and especially of self-reported Deceitfulness in prisoners, we expected a significant positive association between the Dark factor, its themes, and the ability to lie convincingly and to detect the lies produced by others (*Exploratory Hypothesis 2b*).

Regarding the additional individual differences in ***dark and bright tendencies*** that we measured in the prison sample, we expected that dark tendencies, such as Psychopathy, Aggression, and Alexithymia would negatively influence the ability to detect lies (*Exploratory Hypothesis 3a*). Finally, we explored whether protective, bright traits such as low Neuroticism, high Openness to experience, Extraversion, Agreeableness, and Conscientiousness, would facilitate or hinder lie production and detection (*Exploratory Hypothesis 3b*).

3.5.1. Method

Participants

A total number of **140 individuals** (60 civilians and 80 prisoners) with ages ranging from 18 to 67 years old (M = 39.26, SD = 11.20) participated to the study. The **community sample** included 60 individuals (36 males and 24 females) from the general population (regular employees, other than students) all wearing facial masks. The **prison sample** included 80

[6] The content of this sub-chapter (pages 47-52) overlaps with the manuscript: "Behind bars and lies: Dark and bright personality traits and deception production or detection in prisoners and community members" written by Turi, A., Zloteanu, M., Solescu, D., & Visu-Petra, L. (2025) and submitted for publication in *Applied Cognitive Psychology* [Manuscript under revision]. Open Science Framework. https://osf.io/my6cp/

individuals (49 males and 31 females) of which 43 wearing facial masks and 37 with no facial mask. Of these, 51% had no prior criminal record, 12% had a criminal record, and 17% were recidivists. See Table 1 for a detailed analysis of the types of offenses.

Table 1.
Convictions by crime category for prison sample

Category of crime (examples)	**N**	**%**
murder or serious injury (*murder, manslaughter, homicide, assault*)	35	43.8
sex-related crimes (*sexual assault, rape, child molestation*)	9	11.3
theft or robbery (*theft, burglary, armed robbery, and robbery*)	5	6.3
drug-related crimes (*drug traffic, drug consumption*)	7	8.8
forgery crimes (*fraud, money laundry, traffic influence, corruption*)	9	11.3
other crimes (*firearm possession, blackmail, kidnapping, cigarette/human trafficking, organized crime, pimping*).	15	18.8

Note. N = 80

Measures and procedure

Data collection was conducted in three sessions. First, each individual completed an individual interview assessing a variety of personal information such as age, sex, family and educational background, employment, marital status, alcohol and drug use, criminal history, and mental health issues. For the inmate sample, additional data were gathered from prison files, such as offense type, sentence length, prison conduct, and recidivism rates.

Next, the participants completed the **self-report measures** and rated each item on a 5-point Likert scale. **The Dark factor inventory** (D70) (Moshagen et al., 2018) was used to measure the dark tendencies. **The NEO Five-Factor Inventory** (NEO-FFI, Costa & McCrae, 1992) was used to measure bright or normative personality traits. **Self-Report Psychopathy Scale-III** (SRP-III, Paulhus et al., 2009, Mahmut et al., 2011) was used to measure sub-clinical psychopathy. **The Aggression Questionnaire** (AQ, Buss & Perry, 1992) was used to measure aggression as a personality trait. **Toronto Alexithymia Scale–20** (TAS-20, Bagby et al., 1994a; 1994b) was used to measure alexithymia. **Revised Lie Acceptability Scale** (RLAS, Oliveira & Levine, 2008) was used to measure participants' favorable or unfavorable attitudes towards deception. **Paulhus Deception Scales** (PDS, Paulhus, 1998) – the 20-item Impression Management subscale was used to measure social desirability.

In the third stage of data collection, participants were divided into groups of six and completed the **Deceptive Interactive Task** (DeceIT; Wright et al., 2012; 2013), an experimental task designed to measure the ability to produce convincing truthful and deceptive statements and to accurately detect such statements. At the outset, participants were informed that they would be participating in a 'communication skills experiment.' Those who took part during the COVID-19 pandemic were required to wear facial masks. Participants were instructed that their goal was to compete with others to appear credible while accurately identifying lies and truths, with two prizes awarded: 'Best Storyteller' and 'Best Lie Detector'. The experiment commenced with all participants completing a 30-question survey (e.g., "Smoking should be banned in public places"), responding with either "agree" (A) or "disagree" (D). This survey established a baseline for their true opinions. Subsequently, each participant received a set of 10 cards, each containing a statement from the opinion survey

paired with a game task (e.g., “tell a lie”). While one participant assumed the role of the Sender and made truthful or deceptive statements based on the cards they drew, the remaining participants acted as Receivers, evaluating the veracity of the Sender’s statements. Each Sender had 20 seconds to deliver their statement and then rated their perceived credibility. Meanwhile, the Receivers assessed the Sender’s statement as either true or false on a 4-point Likert scale (1 = “not at all likely” to 4 = “very likely”). This process continued until each Sender completed 10 trials, ensuring a balanced 50:50 ratio of lies to truths. Finally, all participants rated the perceived difficulty, guilt, anxiety, and cognitive load associated with lying and truth-telling, also using a 4-point Likert scale.

Data analysis

Performance in the Sender and Receiver roles was analyzed in the Signal Detection Theory framework (Wright et al., 2009) using a generalized mixed effects probit model that accounts for the current design (SDT, Zloteanu & Vuorre, 2024). Separate SDT measures were calculated for the Receiver/Sender roles: the Receiver’s capacity to discriminate lies from truths is indexed by *d' Receiver*; the corresponding measure of bias, *C Receiver*, corresponds to Truth Bias (with negative values indicating the tendency to judge statements as deceptive regardless of veracity, while a positive value indicates a bias to classify messages as truthful). The discriminability of the Sender’s truths and lies is indexed by *d' Sender*. The corresponding measure of bias, *C Sender*, indicates the perceived overall credibility of a Sender, regardless of their veracity. It is often termed Demeanor Bias within the deception literature and positive scores indicate higher credibility regardless of the statement's veracity. With these measures, successful deception is indicated by more negative values of d′ Sender, and better lie detection is reflected by higher positive d′ Receiver values.

3.5.2. Results and Conclusions

The two groups were similar in terms of basic socio-demographic variables, as t-tests revealed no significant differences between prisoners and civilians in terms of age ($t(138) = 1.352$, $p = 0.089$), gender ($t(138) = -.149$, $p = 0.441$), and education ($t(138) = 1.329$, $p = 0.093$). However, prisoners scored higher on the Dark Factor compared to civilians ($t(138) = -5.182$, $p < 0.001$), and presented less lenient attitudes toward lying, ($t(138) = -5.927$, $p < .001$).

H1: Deception-general ability in prisoners and civilians

We found weak evidence for *a deception-general ability in civilians*, indicated by a small negative correlation between their detection accuracy and detectability (r = -0.28 [-0.52, -0.02], pd = 98.25% ROPE = 6.08%). However, we found moderate evidence for *the absence of a deception general ability in prisoners* (r = 0.04 [-0.18, 0.27], pd = 64.92%, ROPE = 61.16%).

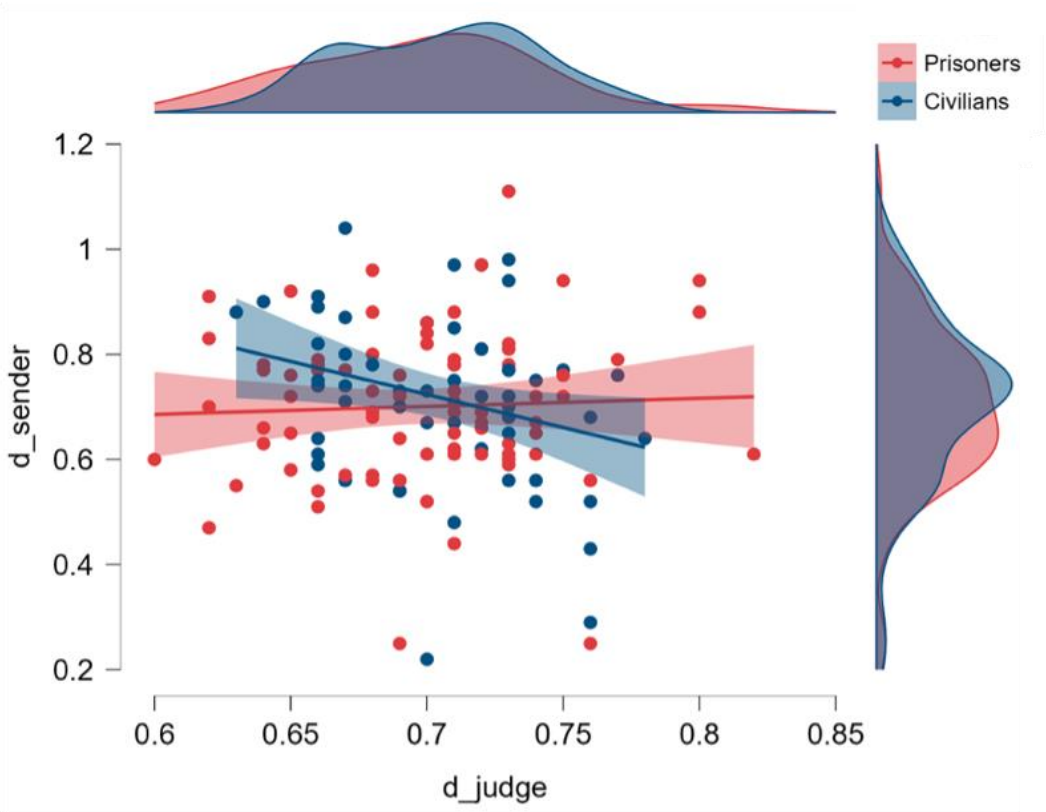

Figure 1.
Association between the ability to discriminate in the Receiver role (d' Receiver) and detectability in the Sender role (d' Sender) for prisoners and civilians.

H2: Relationships between Dark factor or personality, the five dark themes, the ability to produce (d' Sender) and detect lies (d' Receiver) in both prisoners and civilians.

We found no evidence for the influence of either the Dark Factor or its themes on the civilians' ability to correctly identify lies in the Receiver's role (d' Receiver), but strong evidence for a negative relationship between the ability to correctly identify lies in the Receiver's role (d' Receiver) and the Dark Factor (r(80) = -.20 [-.34, -.05], p = .010), Callousness (r(80) = -.16 [-.28, -.01], p = .044), Deceitfulness (r(80) = -.22 [-.35, -.08], p = .005), Narcissistic Entitlement (r(80) = -.20 [-.32, -.05], p = .010) and Sadism (r(80) = -.17 [-.31, -.01], p = .028) in prisoners. (See **Table 2).**

Table 2
Table of partial correlations (controlled for age and social desirability) between the Dark Factor and its themes, and the ability to produce (d' Sender) and detect lies (d' Receiver) in prisoners and civilians.

Dark Personality	PRISONERS		CIVILIANS	
	Lie production: d' *Sender*	Lie detection: d' *Receiver*	Lie production: d' *Sender*	Lie detection: d' *Receiver*
Dark Factor	.04	-.20**	.05	-.20

Callousness	.07	-.16*	-.08	-.15
Deceitfulness	.01	-.22**	.02	-.22
Narcissistic Entitlement	-.03	-.20**	-.02	-.19
Sadism	-.01	-.17*	-.00	-.17
Vindictiveness	.01	-.10	.03	-.08

Note. *p >.05, **p>.01

H3: Relationships between additional dark and bright tendencies and prisoners' ability to produce (d' Sender) and detect lies (d' Receiver)

We found anecdotal evidence for a negative correlation between prisoners' lie detection accuracy (d' Receiver) and Interpersonal Manipulation facet of Psychopathy (r(80 = -.17 [-.32, -.02], p = .027, overall Aggression (r(80) = -.20 [-.34, -.04], p = .011), Anger (r(80) = -.15 [-.31, -.00], p = .045), strong evidence for Physical Aggression (r (80) = -.26 [-.40, -.12], p < .001), Alexithymia (r(80) = -.26 [-.40, -.12], p < .001), Difficulties Describing Feelings (r(80) = -.26 [-.39, -.12], p < .001) and moderate evidence for Difficulties Identifying Feelings (r(80) = -.21 [-.35, -.07], p = .006), supporting *Exploratory Hypothesis 3a*.

Regarding the bright tendencies, results show no evidence for a correlation between neither the Big Five Factors nor the ability to detect lies, failing to support *Exploratory Hypothesis 3b* (See **Table 3**).

Table 3

Table of partial correlations (controlled for age and social desirability) between additional dark tendencies: Psychopathy and its subscales: Callous Affect, Erratic Lifestyle, Interpersonal Manipulation, Criminal Tendencies, Aggression and its subscales: Anger, Verbal Aggression, Physical Aggression, Hostility, Alexithymia and its subscales: Difficulties Describing Feelings, Difficulties Identifying Feelings, Externally Oriented Thinking, bright tendencies: Openness to experience, Conscientiousness, Extraversion, Agreeableness, Neuroticism, Lie acceptability and the ability to produce (d' Sender) and detect lies (d' Receiver) in prisoners.

	Lie production: d' *Sender prisoners*	Lie detection: d' *Receiver* prisoners
Psychopathy	-.58	-.12
Callous Affect	.15*	-.61
Erratic Lifestyle	.79	-.92
Interpersonal Manipulation	.11	-.17*
Criminal Tendencies	-.17	-.61
Aggression	-.15	-.20*
Anger	.74	-.15*
Verbal Aggression	-.82	-.42
Physical Aggression	.38	-.26***
Hostility	-.62	-.67
Alexithymia	.59	-.26***
Difficulties Describing Feelings	-.12	-.26***
Difficulties Identifying Feelings	.90	-.21**
Externally Oriented Thinking	.13	-.11

Typical personality: Five-factor Model			
	Neuroticism	-.14	-.35
	Extraversion	-.18	.52
	Openness	.49	.07
	Agreeableness	.40	.80
	Conscientiousness	.09	-.14
Lie acceptability		.13	-.18*

Note. *p >.05, **p>.01, ***p>.00).

The findings of this study offer important theoretical and practical implications for understanding the relationship between personality traits and deception abilities. The theoretical insights highlight the complex interplay between personality traits and cognitive-emotional processes in shaping deception abilities, underscoring the importance of considering both personality factors and contextual influences when studying deception in realistic settings.

From a practical standpoint, these findings suggest that interventions aimed at enhancing emotional awareness, specifically targeting issues such as alexithymia (difficulty identifying and describing emotions), could significantly improve deception detection skills in incarcerated populations. By teaching individuals to gain insight into their emotional states and interpret emotional cues, these interventions may help individuals recognize deceptive behaviors more accurately. Additionally, interventions designed to address (physical) aggression and promote better emotional awareness, could mitigate impulsive judgments, and hostile attributions and enhance prisoners' ability to assess others' honesty more effectively.

Being able to accurately read the behavior of others can influence how one responds to it, especially for individuals with dark personality traits. For instance, those who wrongly assume deception in others may be more prone to react aggressively due to perceived betrayal or manipulation. By enhancing deception detection skills, prisoners might develop a more accurate understanding of others' intentions, potentially reducing unnecessary aggressive responses. This improved perception of honesty or deception could also foster greater trust and reduce instances of conflict, ultimately contributing to lower levels of aggression and more respectful interpersonal interactions.

THEORETICAL, EMPIRICAL AND METHODOLOGICAL CONTRIBUTIONS

The main objective of this PhD thesis was to explore the **interplay between dark personality traits and deception** among prisoners. To this end, we examined deception both as a *personality trait* (**Deceitfulness**), and a *behavioral act* of producing and detecting lies, connected to individuals' attitudes towards deception.

We also considered the role of additional **cognitive** (e.g., subjective definitions of deception, justifications, lie categorizations, stereotypical thinking about cues to detection), **affective** (e.g., difficulties in identifying and describing personal emotions), and **personality factors** (Dark Factor Themes, Big 5 Factors, PID-5 domains, aggression).

The results of our investigations enable us to advance ***an integrative model for interpersonal deception*** that extends the one developed by Burgoon & Buller (2008) to account for the role of personality traits in shaping the outcomes of deception and provide a better understanding of the per-interaction factors (e.g., expectations, motivations, goals, and knowledge) that influence the initial display of the deceptive behavior.

4.1. Theoretical and Empirical Contributions

Study	Aim	Main Theoretical and Empirical Contributions
Study 1	Summarize findings from existing studies on the relationship between Dark Triad personality traits and deception across diverse population.	▪ Provide support for a **link between Dark Triad traits, lie frequency, lie acceptability, and self-reported deceptive abilities**, but inconsistent findings in the relation to deceptive performances measured via experimental tasks.
Study 2	Examine the predictive role of typical, maladaptive, and dark personality traits in shaping attitudes toward deception (lie acceptability)	▪ Provide theoretical and empirical support for **associations between typical, pathological, and dark personality traits and lie acceptability.** ▪ Identify the five **Dark Themes: Callousness, Deceitfulness, Narcissistic Entitlement, Sadism, and Vindictiveness, as key predictors of lie acceptability**, over the typical (Big 5) and maladaptive personality traits (PID-5), addressing a significant gap in the previous literature.
Study 3	Explore inmates' conceptualizations of deception, self-reported lie frequency, attitudes toward different types of lies, and reliance on subjective versus objective cues	▪ Establish a **new** theoretical framework for **understanding underlying beliefs and perceptions of deception** in prison. ▪ Support the applicability of the **orthogonal taxonomy of lies by Visu-Petra et al. (2022)** within the prison population. ▪ Partially support the **feedback hypothesis** advanced by Granhag et al. (2004) by demonstrating prisoners possess accurate knowledge about verbal cues to deception but provide **novel evidence on the presence of stereotypical beliefs** about the non-verbal cues. ▪ Document **inmate's accurate beliefs about the ingredients of successful deception,** hence supporting Vrij et al., (2010) theories concerning the characteristics of good liars.

Study 4	Investigate the associations between lie production, lie detection, lie acceptability, and personality traits among both prisoners and civilians	▪ Support the theoretical conceptualization of **a general deception ability in civilians** and provide preliminary evidence for the **absence of this ability in prisoners**. ▪ Provide new evidence regarding the **negative influence of Dark Personality Themes, Alexithymia, and Aggression on deception detection accuracy** in prison population. ▪ Offer empirical validation for the **Integrative Interpersonal Deception Model** advanced in this thesis, highlighting the role of individual predispositions in shaping deception detection accuracy.

4.2. Methodological Contributions

Study	Aim	Main Methodological Contributions
Study 1	Conduct a **scoping review** on the research concerning Dark Triad traits in relation to deception production and deception detection	▪ Analyses the findings of **18 publications** and documents the congruence between self-reported and performance-based deception production and detection abilities in individuals with Dark Triad traits, following **Arksey & O'Malley's (2005) methodological framework.** ▪ It also highlights the inconsistencies that emerge across different types of measurement tools.
Study 2	Examine the combined **contribution of typical, pathological, and dark personality traits on lie acceptability** in prison population	▪ **Validates the D70** as a useful tool to measure the dark themes of personality in prison population. ▪ Employs **Regression analysis** to demonstrate the predictive role of dark personality themes for lie acceptability above the typical and pathological traits, a previously unaddressed gap in the literature.
Study 3	Develop and employ a **mix-methods approach** to study deception related beliefs in prison	▪ Develops and validates a **new** tool (**Qualitative and Quantitative Research Protocol**) for studying deception in prisoners. ▪ Employs **Thematic Analysis** (TA, Braun & Clarke, 2006) to examine inmate's beliefs about deception production and detection, for the first time in the literature.
Study 4	Modify and validate the **DeceIT paradigm** as an applicable method for examining interpersonal deception in both prison and community environments	▪ Introduces an **ecologically valid experimental task** (**DeceIT**, Wright et al., 2012) for measuring deception production and detection performances, applicable in both prison and community settings. ▪ Employs **Signal Detection Theory** (SDT, Zloteanu & Vuorre, 2024) to examine inmate's deception production and detection performances, an approach that has been not used in prison before.

LIMITATIONS, IMPLICATIONS, AND CONCLUSIONS

Despite the valuable contributions of this PhD thesis, several limitations must be acknowledged, so that future research can expand the theoretical and empirical findings of this thesis.

First, the **generalizability** of our findings. Although our scoping review has been conducted on various populations and the Dark Factor Inventory has been validated on a large pool of participants from the general population (Hurezan et al., 2024), in Study 2 and Study 3 we had no control group. Only in Study 4 did we have participants from both populations. This limits our capacity to expand our findings to the general population. Thus, future research should address this limitation by employing more civilians.

A second limitation related to the **type of offense the inmates** were convicted for. Although we tried to have a heterogeneous group of participants convicted for various types of offenses, due to the prison's specificity (maximum security), most of our participants were convicted for serious violent offenses such as murder. Future studies should employ this methodology on inmates convicted of non-violent crimes and investigate if they obtain the same results.

A third limitation is the reliance on **self-report measures for personality**. Even when combined with performance-based tasks, these measures may not fully overcome the influence of impression management or self-deceptive tendencies, particularly relevant in dark trait assessments. A solution to this may be the use of informants, such as prison guards or other inmates, but still, they should be selected with caution to avoid contaminating the possible findings with their own subjective beliefs.

A fourth limitation is related to the **contextual factors specific to prison settings** (e.g., degree of endorsing the unwritten rules of the inmate code, the amount of stress imposed by incarceration, the group status of the inmates) that may affect not only the way they think, but also the way they behave within social interactions.

As a general conclusion, this PhD thesis embarked on a series of studies exploring the intricate relation between dark personality and the complex phenomenon of interpersonal deception, through *qualitative* and *quantitative* methodologies. We started by reviewing the literature on dark personality traits and deception and identified a ***mismatch between self-reported and performance-based deception production and detection abilities*** (in **Study 1**). We also noted that individuals may possess a general deception ability which grants them success in both deception production and detection, and that their attitudes towards lying can impact their frequency.

To advance our understanding of both dark personalities, we adapted and validated a self-report instrument designed to measure the Dark factor of personality (defined as the tendency of maximize self-utility at the expense of others), which allowed us to identify "***Deceitfulness***" a personality trait that distinguishes individuals with criminal behavior from those without one (Hurezan et al., 2024). In **study 2**, we utilized this instrument along with other measures of personality (e.g., The Big Five and PID-5) to investigate a broad spectrum of dark, pathological, and typical traits in relation to lie acceptability. This analysis revealed that ***Dark Themes of personality alone explained 31% of the variance in individuals'***

attitudes towards deception, meaning that they had predictive value over and above the other personality models.

Hence, we considered it relevant to develop a qualitative and quantitative protocol to explore these attitudes in detail, as well as to identify inmates' beliefs about deception detection, aiming to provide a more comprehensive understanding of the insights they may have gathered during previous encounters with deception. In **Study 3**, we applied this protocol and identified similarities between their and researchers' conceptualizations and categorizations of deception. However, we noticed negative ***attitudes towards all types of deception***, especially in relation to their family members and friends. They admitted lying to authority despite knowing it is morally wrong. An interesting finding was the belief that constant deception can erode one's character, and that being labeled "liar" can distance oneself from others, a price most of them were not willing to pay for the possible advantages of successful deception. They acknowledged that good liars must possess a combination of traits and behaviors previously identified by the literature (see Vrij et al., 2010) and reported ***accurate beliefs about the verbal indicators of deception, but stereotypical beliefs about the non-verbal ones***.

Building upon these findings and the limitations we identified in our scoping review, in **Study 4**, we advanced our research and investigated the existence of *a general deception ability* in both prisoners and civilians using a modified version of the experimental task DeceIT, originally developed by Wright et al., (2012). The major advantage of this task is that it allows interaction between participants, creating an ecological setting where deception production and detection can be measured simultaneously within the same participant. We also examined the role of the Dark themes of personality in deception success and deception detection accuracy, as well as the possible influence of typical personality traits, alexithymia, aggression, and lie acceptability. The findings revealed significant differences between prisoners and civilians, indicating a ***general deception ability in the general population***, but not in prison. Also, they demonstrated the ***negative influence of dark personality traits on deception detection accuracy in the prison population***, suggesting that inmates with pronounced levels of dark personality traits, aggression, and alexithymia are more likely to misjudge others and fail to identify when they are being deceived.

An **overview** of the links explored through the studies in the current thesis and the Integrative Interpersonal Deception Model that we advance can be found in the figure below.

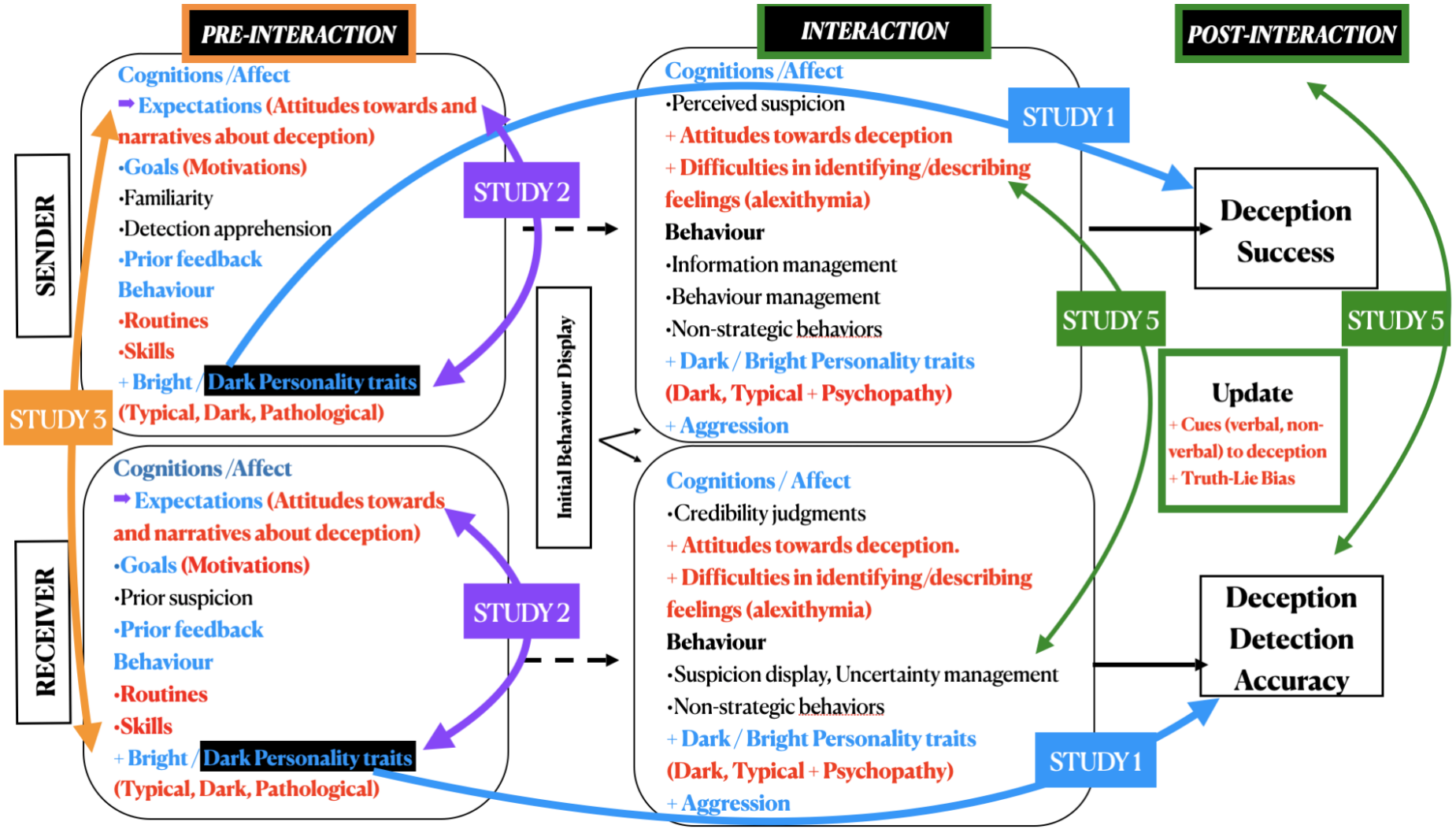

Figure 1.
Overview of the links explored in the studies supporting the Integrative Interpersonal Deception Model proposed in this PhD thesis.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi/book/10.1176/appi.books.9780890425596

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. International journal of social research methodology, 8(1), 19-32. https://doi.org/10.1080/1364557032000119616.

Azizli, N., Atkinson, B. E., Baughman, H. M., Chin, K., Vernon, P. A., Harris, E., &Veselka, L. (2016). Lies and crimes: Dark Triad, misconduct, and high-stakes deception. Personality and Individual Differences, 89, 34-39. https://doi.org/10.1016/j.paid.2015.09.034.

Bader, M., Hartung, J., Hilbig, B. E., Zettler, I., Moshagen, M., & Wilhelm, O. (2021). Themes of the dark core of personality. *Psychological Assessment, 33*(6), 511–525. https://doi.org/10.1037/pas0001006

Baughman, H. M., Jonason, P. K., Lyons, M., & Vernon, P. A. (2014). Liar liar pants on fire: Cheater strategies linked to the Dark Triad. Personality and Individual Differences, 71, 35-38. https://doi.org/10.1016/j.paid.2014.07.019.

Birkás, B., Pátkai, G., &Csathó, Á. (2020).The mediating role of the dark triad between life history strategy and perceived stress factors.Psychological Reports, 123(2), 252-265. https://doi.org/10.1177/0033294118818095.

Bonfá-Araujo, B., Ferreira, L. B., Jesuíno, A. D. S. A., Hauck-Filho, N., & Iglesias, F. (2023). Measuring the dark core: A Brazilian adaptation and comparison between the general population and incarcerated men. *Journal of Criminal Justice*, *89*, 10213 https://doi.org/10.1016/j.jcrimjus.2023.102133

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3(2)*, 77–101. https://doi.org/10.1191/1478088706qp063oa

Bryant, E. M. (2008). Real lies, white lies and gray lies: Towards a typology of deception. Kaleidoscope: A Graduate Journal of Qualitative Communication Research, 7, 23. Retrieved from https://digitalcommons.trinity.edu/cgi/viewcontent.cgi?article=1006&context=hct_faculty.

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral Confirmation of Everyday Sadism. *Psychological Science*, *24*(11), 2201-2209. https://doi.org/10.1177/0956797613490749.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x

Burgoon, J. K., & Buller, D. B. (2008). Interpersonal deception theory. In *Engaging theories in interpersonal communication: Multiple perspectives* (pp. 227-239). https://doi.org/10.1111/j.1468-2885.1996.tb00127.x

Burgoon, J. K., Buller, D. B., & Guerrero, L. K. (1995). Interpersonal deception: IX. Effects of social skill and nonverbal communication on deception success and detection accuracy. *Journal of Language and Social Psychology*, *14*(3), 289-311. https://doi.org/10.1177/0261927X9514300

Butler, M., & Maruna, S. (2009). The impact of disrespect on prisoners' aggression: Outcomes of experimentally inducing violence-supportive cognitions. *Psychology, Crime & Law, 15(2-3)*, 235-250. https://doi.org/10.1080/10683160802190970

Chabrol, H., Van Leeuwen, N., Rodgers, R., & Séjourné, N. (2009). Contributions of psychopathic, narcissistic, Machiavellian, and sadistic personality traits to juvenile

delinquency. *Personality and individual differences*, *47*(7), 734-739. https://doi.org/10.1016/j.paid.2009.06.020

Christie, R., & Geis, F. L. (1970). Studies in machiavellianism. New York: Academic Press.

Ciurbea, F. E., Dina, M. M., & Rada, C. (2022). Extraversion and Agreeableness–the role of personality dimensions in the criminal act. *Revista de Psihologie*, *68*(3), 243-256.

Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders, 6*(4), 343–359. https://doi.org/10.1521/pedi.1992.6.4.343

Coyne, S. M and Ostrov, J. M,. 2018. The development of relational aggression, oxford university press.

Crewe, B. (2012). *The prisoner society: Power, adaptation and social life in an English prison*. OUP Oxford.

Daiku, Y., Serota, K. B., & Levine, T. R. (2021). A few prolific liars in Japan: Replication and the effects of Dark Triad personality traits. PloS one, 16(4), e0249815. https://doi.org/10.1371/journal.pone.0249815.

Dargis, M. & Koenigs, M. (2017). Two Subtypes of Psychopathic Criminals Differ in Negative Affect and History of Childhood Abuse. Psychological Trauma: Theory, Research, Practice, and Policy. 10. 10.1037/tra0000328.

DePaulo, B. M., & Rosenthal, R. (1979). Telling lies. Journal of personality and social psychology, 37(10), 1713. DOI:10.1037/0022-3514.37.10.1713.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979–995. https://doi.org/10.1037/0022-3514.70.5.979

Dhami, M. K., Ayton, P., & Loewenstein, G. (2007). Adaptation to imprisonment: Indigenous or imported?. *Criminal Justice and Behavior, 34(8)*, 1085-1100. https://doi.org/10.1177/00938548073020

Eisenberg, N., Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., & Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal study. *Journal of Personality and Social Psychology, 82(6)*, 993–1006. https://doi.org/10.1037/0022-3514.82.6.993

Elaad, E., Hanania, S. B., Mazor, S., &Zvi, L. (2020). The relations between deception, narcissism and self-assessed lie-and truth-related abilities.Psychiatry, psychology and law, 27(5), 880-893. https://doi.org/10.1080/13218719.2020.1751328.

Figueredo, A. J., Hertler, S. C., & Peñaherrera-Aguirre, M. (2021). The biogeography of human diversity in life history strategy. *Evolutionary Behavioral Sciences, 15*(1), 10–26. https://doi.org/10.1037/ebs0000198

Forsyth, L., Anglim, J., March, E., & Bilobrk, B. (2021). Dark Tetrad personality traits and the propensity to lie across multiple contexts. *Personality and Individual Differences, 177*, 110792. https://doi.org/10.1016/j.paid.2021.110792

Fox, B. & DeLisi, M. (2019). Psychopathic killers: A meta-analytic review of the psychopathy-homicide nexus. *Aggression and Violent Behavior, 44*, 67-79, https://doi.org/10.1016/j.avb.2018.11.005.

García-Fernández, J., Postigo, Á., Cuesta, M., & Moshagen, M. (2024). Spanish adaptation of the D70—The dark core of personality in Spanish-speaking countries. *European Journal of Psychological Assessment.* Advance online publication. https://doi.org/10.1027/1015-5759/a000836

Geis, F. L., & Moon, T. H. (1981). Machiavellianism and deception. Journal of personality and social psychology, 41(4), 766. https://doi.org/10.1037/0022-3514.41.4.766.

Grigoras, M., & Wille, B. (2017). Shedding light on the dark side: Associations between the dark triad and the DSM-5 maladaptive trait model. *Personality and Individual Differences*, *104*, 516-521. https://doi.org/10.1016/j.paid.2016.09.016

Harkness, A. R., Reynolds, S. M., & Lilienfeld, S. O. (2014). A review of systems for psychology and psychiatry: Adaptive systems, personality psychopathology five (PSY–5), and the DSM–5. *Journal of personality assessment*, *96*(2), 121-139. https://doi.org/10.1080/00223891.2013.823438

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Andersson, L. O. (2004). SUSPICIOUS MINDS: CRIMINALS'ABILITY TO DETECT DECEPTION. *Psychology, Crime and Law, 10*(1), 83-95. https://doi.org/10.1080/1068316031000095485

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly, 7*(4), 429–456. https://doi.org/10.1177/1098611104264748

Hayashi, A., Abe, N., Fujii, T., Ito, A., Ueno, A., Koseki, Y., ... & Mori, E. (2014). Dissociable neural systems for moral judgment of anti-and pro-social lying. *Brain Research*, *1556*, 46-56. https://doi.org/10.1016/j.brainres.2014.02.011

Hopwood, C. J., & Sellbom, M. (2013). Implications of DSM-5 personality traits for forensic psychology. *Psychological Injury and Law*, *6*, 314-323. https://doi.org/10.1007/s12207-013-9176-5

Hurezan, L., Turi, A., Ion, A., & Visu-Petra, L. (2024). Dark and bright personality dimensions as predictors of criminal behavior and recidivism. *Scientific Reports, 14*(1), 18565. https://doi.org/10.1038/s41598-024-69288-5

Jackson, J. J., & Hill, P. L. (2019). Lifespan development of conscientiousness. In D. P. McAdams, R. L. Shiner, & J. L. Tackett (Eds.), *Handbook of personality development*, 153–170. The Guilford Press.

Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual differences*, *40*(2), 331-339. https://doi.org/10.1016/j.paid.2005.07.006

Janović, T., Ivković, V., Nazor, D., Grammer, K., & Jovanović, V. (2003). Empathy, communication, deception. *Collegium antropologicum*, *27*(2), 809-822.

Jonason, P. K., Lyons, M., Baughman, H. M., & Vernon, P. A. (2014). What a tangled web we weave: The Dark Triad traits and deception. Personality and Individual Differences, 70, 117-119. https://doi.org/10.1016/j.paid.2014.06.038.

Kaplan, H. S., & Gangestad, S. W. (2015). Life history theory and evolutionary psychology. *The handbook of evolutionary psychology*, 68-95. https://doi.org/10.1002/9780470939376.ch2

Knapp, M. L., &Comadena, M. E. (1979). Telling it like it isn't: A Review of theory and research on deceptive communication. Human Communication Research, 5, 270-285.

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychological Medicine, 42(9), 1879–1890. doi:10.1017/S0033291711002674

Krusemark, E. A., Campbell, W. K., Crowe, M. L., & Miller, J. D. (2018). Comparing self-report measures of grandiose narcissism, vulnerable narcissism, and narcissistic personality disorder in a male offender sample. *Psychological Assessment, 30*(7), 984–990. https://doi.org/10.1037/pas0000579

Laws, B. (2016). *Fronting, masking, and emotion release: An exploration of prisoners' emotional management strategies* [Unpublished doctoral dissertation]. University of Cambridge.

Liu, J. (2009). Asian criminology–challenges, opportunities, and directions. *Asian Journal of criminology*, *4*, 1-9. https://doi.org/10.1007/s11417-009-9066-7.

Lyons, M., Croft, A., Fairhurst, S., Varley, K., & Wilson, C., (2017). Seeing through crocodile tears? Sex-specific associations between the Dark Triad traits and lie detection accuracy.Personality and individual differences, 113, 1-4. https://doi.org/10.1016/j.paid.2017.03.008.

Lyons, M., Healy, N., & Bruno, D. (2013). It takes one to know one: Relationship between lie detection and psychopathy. *Personality and Individual Differences, 55*(6), 676–679. https://doi.org/10.1016/j.paid.2013.05.018

Martin, K., & Leach, A. M. (2013). Psychopathy and deception detection. *Personality and mental health, 7*(2), 154-159. https://doi.org/10.1002/pmh.1215

Maruna, S., & Butler, M. (2013). Violent Self-Narratives and the Hostile Attributional Bias. In D. Youngs (Ed.), Behavioural Analysis of Crime: Studies in David Canter's Investigative Psychology (pp. 27-48). Ashgate Publishing Ltd.

Masip, J., Garrido, E., & Herrero, C. (2004). Facial appearance and impressions of 'credibility': The effects of facial babyishness and age on person perception. *International Journal of Psychology, 39*(4), 276–289. https://doi.org/10.1080/00207590444000014

Međedović, J. (2012). Topography of dishonesty: Mapping the opposite pole of honesty-humility personality domain. *Primenjena psihologija*, *5*(2).

Michels, M., Molz, G., & Maas genannt Bermpohl, F. (2020). The ability to lie and its relations to the dark triad and general intelligence. *Personality and Individual Differences, 166*, 110195. https://doi.org/10.1016/j.paid.2020.110195

Mokros, A., Osterheider, M., Hucker, S. J., & Nitschke, J. (2011). Psychopathy and sexual sadism. *Law and human behavior*, *35*, 188-199. https://doi.org/10.1007/s10979-010-9221-9

Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The dark core of personality. *Psychological Review, 125*(5), 656–688. https://doi.org/10.1037/rev0000111

Moshagen, M., Zettler, I., & Hilbig, B. E. (2020). Measuring the dark core of personality. *Psychological Assessment*, *32*(2), 182. https://doi.org/10.1037/pas0000778

Musek, J., & Grum, D. K. (2021). The bright side of personality. *Heliyon*, *7*(3). DOI: 10.1016/j.heliyon.2021.e06370

Nagar, P. M., Caivano, O., & Talwar, V. (2020). The role of empathy in children's costly prosocial lie-telling behaviour. *Infant and Child Development*, *29*(4), e2179. https://doi.org/10.1002/icd.2179

Niemeyer, L. M., Grosz, M. P., Zimmermann, J., & Back, M. D. (2022). Assessing maladaptive personality in the forensic context: Development and validation of the Personality Inventory for DSM-5 Forensic Faceted Brief Form (PID-5-FFBF). *Journal of Personality Assessment*, *104*(1), 30-43.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological Assessment, 23*(2), 523–531. https://doi.org/10.1037/a0022400

O'Reilly III, C. A., &Doerr, B. (2020). Conceit and deceit: Lying, cheating, and stealing among grandiose narcissists. Personality and Individual Differences, 154, 109627. https://doi.org/10.1016/j.paid.2019.109627.

Oliveira, C. M., & Levine, T. R. (2008). Lie acceptability: A construct and measure. *Communication Research Reports, 25*(4), 282–288. https://doi.org/10.1080/08824090802440170

Paulhus, D. L. (1998). *Paulhus deception scales (PDS): The Balanced Inventory of Desirable Responding-7: User's manual*. Multi-Health Systems, Incorporated.

Paulhus, D. L., Curtis, S. R., & Jones, D. N. (2018). Aggression as a trait: The Dark Tetrad alternative. *Current opinion in psychology*, *19*, 88-92. https://doi.org/10.1016/j.copsyc.2017.04.007

Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45*(2), 590. https://doi.org/10.2466/pr0.1979.45.2.590

Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54*(5), 890–902. https://doi.org/10.1037/0022-3514.54.5.890.

Schindler, S., Reinhard, M. A., Dobiosch, S., Steffan-Fauseweh, I., Özdemir, G., & Greenberg, J. (2019). The attenuating effect of mortality salience on dishonest behavior.Motivation and Emotion, 43(1), 52-62.

Semrad, M., & Scott-Parker, B. (2020). Police, personality and the ability to deceive. *International Journal of Police Science & Management, 22*(1), 50–61. https://doi.org/10.1177/1461355719880568

Shafqat, A., Majeed, S. & Malik, F. (2019). Sadistic Impulsiveness and Violent Behaviour in Prisoners of Lahore. *Pakistan Journal of Criminology, 11*(01), 28-42. https://www.pjcriminology.com/wp-content/uploads/2019/06/3.PRISONERS-ARTICLE.pdf

Shulman, E. P., Cauffman, E., Piquero, A. R., & Fagan, J. (2011). Moral disengagement among serious juvenile offenders: a longitudinal study of the relations between morally disengaged attitudes and offending. *Developmental psychology*, *47*(6), 1619.

Strickland, C. M., Drislane, L. E., Lucy, M., Krueger, R. F., & Patrick, C. J. (2013). Characterizing psychopathy using DSM-5 personality traits. *Assessment*, *20*(3), 327-338. https://doi.org/10.1177/1073191113486691

Trammell, R. (2012). Enforcing the convict code: Violence and prison culture. Boulder, CO: Lynne Rienner Publishers.

Turi, A., Rebeleș, M. R., & Visu-Petra, L. (2022). The tangled webs they weave: A scoping review of deception detection and production in relation to Dark Triad traits. *Acta Psychologica*, *226*, 103574. https://doi.org/10.1016/j.actpsy.2022.103574

Turi, A., Zloteanu, M., Solescu, D., & Visu-Petra, L. (2025). Behind bars and lies: Dark and bright personality traits and deception production or detection in prisoners and community members [Manuscript submitted for publication]. *Open Science Framework*. https://osf.io/my6cp/

Visu-Petra, L., Prodan, N., & Talwar, V. (2022). Children's Lies: Intersecting Cognitive Development, Theory of Mind, and Socialization. *The Wiley-Blackwell Handbook of Childhood Social Development*, 668-685. https://doi.org/10.1002/9781119679028.ch36

Volmer, J., Koch, I. K., & Göritz, A. S. (2016). The bright and dark sides of leaders' dark triad traits: Effects on subordinates' career success and well-being. *Personality and Individual Differences*, *101*, 413-418. https://doi.org/10.1016/j.paid.2016.06.046

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley & Sons.

Wiebe, R. P. (2004). Delinquent behavior and the five-factor model: Hiding in the adaptive landscape?. *Individual differences research*, *2,* 38-62.

Wilt, J. A., Grubbs, J. B., Pargament, K. I., & Exline, J. J. (2017). Religious and spiritual struggles, past and present: Relations to the big five and well-being. *International Journal for the Psychology of Religion, 27*(1), 51–64. https://doi.org/10.1080/10508619.2016.1183251 .

Wissing, B. G., & Reinhard, M.-A. (2017). The Dark Triad and the PID-5 maladaptive personality traits: Accuracy, confidence, and response bias in judgments of veracity. *Frontiers in Psychology, 8.* https://doi.org/10.3389/fpsyg.2017.01549

Wissing, B. G., &Reinhard, M. A. (2019).The dark triad and deception perceptions.Frontiers in psychology, 10, 1811. https://doi.org/10.3389/fpsyg.2019.01811.

Wright, G. R. T., Berry, C. J., & Bird, G. (2012). "You can't kid a kidder": Association between production and detection of deception in an interactive deception task. *Frontiers in Human Neuroscience, 6.* https://doi.org/10.3389/fnhum.2012.00087

Wright, G. R. T., Berry, C. J., Catmur, C., & Bird, G. (2015). Good liars are neither 'dark' nor self-deceptive. *PLOS ONE, 10*(6), e0127315. https://doi.org/10.1371/journal.pone.0127315

Zeigler-Hill, V., Mandracchia, J. T., Dahlen, E. R., Shango, R., & Vrabel, J. K. (2017). Pathological personality traits and criminogenic thinking styles. *Personality and Individual Differences*, *110*, 41-48. https://doi.org/10.1016/j.paid.2017.01.021

Zloteanu, M., & Vuorre, M. (2024). A tutorial for deception detection analysis or: How I learned to stop aggregating veracity judgments and embraced signal detection theory mixed models. *Journal of Nonverbal Behavior, 48*(1), 161–185. https://doi.org/10.1007/s10919-024-00456-x

Zvi, L., & Elaad, E. (2018). Correlates of narcissism, self-reported lies, and self-assessed abilities to tell and detect lies, tell truths, and believe others. Journal of Investigative Psychology and Offender Profiling, 15(3), 271-286. https://doi.org/10.1002/jip.1511.