The Ethical Mentor or the Hidden Manipulator in Higher Education? Machiavellianism Behind a Caring Façade

Abdelaziz Alowais * and Abubakr Suliman

Posted Date: 11 February 2026

doi: 10.20944/preprints202602.0850.v1

Keywords: Machiavellianism; ethical mentorship; trust and morale; manipulative strategies; moral disengagement; higher education institutions (HEIs)

 doi:10.20944/preprints202602.0850.v1

Article

The Ethical Mentor or the Hidden Manipulator in Higher Education? Machiavellianism Behind a Caring Façade

The Ethical Mentor or the Hidden Manipulator? Machiavellianism Behind a Caring Facade

Abdelaziz Abdalla AlOwais * and Abubakr Suliman

The British University in Dubai, Faculty of Business and Law, Dubai, United Arab Emirates

* Correspondence: abdelaziz.alowais@icloud.com; 23000066@buid.ac.ae

Abstract

Introduction: Ethical mentorship is commonly portrayed as a mechanism for guidance, care, and professional development. However, within hierarchical settings, ethical language may also be strategically deployed by Machiavellian leaders to mask self-serving motives. This study examines the paradox of ethical mentorship when care-based rhetoric is used to legitimize control, obligation, and dependency. Building on Alowais and Suliman (2025), the study extends existing work on dark leadership traits by focusing specifically on mentorship dynamics within Higher Education Institutions (HEIs), contexts known for bureaucratic complexity and symbolic moral discourse. **Methods:** A qualitative research design was adopted using semi-structured interviews with employees working in HEIs. This approach enabled an in-depth exploration of lived experiences, emotional ambivalence, and relational power dynamics that are not readily captured through quantitative methods. Data were analysed thematically to identify recurring patterns related to ethical rhetoric, manipulation, and trust. **Results:** The findings revealed three interrelated themes. First, Machiavellian mentors strategically employed professional and ethical language to enhance legitimacy, making manipulative practices difficult to detect. Second, mentees experienced persistent emotional ambivalence, characterized by simultaneous feelings of gratitude and exploitation driven by norms of moral duty and reciprocity. Third, once manipulation became apparent, fractures in trust emerged and spread across teams, undermining morale and organizational culture. **Discussion:** The study contributes to leadership and ethics scholarship by demonstrating how mentorship can be weaponized under the guise of care, transforming ethical guidance into a subtle mechanism of control. The findings underscore the importance of distinguishing genuine support from performative ethics within mentoring relationships. Practically, the study highlights the need for institutional vigilance, targeted training, transparent feedback mechanisms, and safeguards that protect mentees from coercive dependency. Without such measures, ethical mentorship risks becoming an instrument of domination rather than development.

Keywords: Machiavellianism; ethical mentorship; trust and morale; manipulative strategies; moral disengagement; higher education institutions (HEIs)

1. Introduction

Dark leadership remains a pressing concern, particularly when mentorship is used as a mask for manipulation. Jung's Shadow Theory shows how hidden destructive motives can be projected as care (Jung, 1959/1990), while the Toxic Triangle frames how such behaviors thrive in permissive environments (Padilla, Hogan, & Kaiser, 2007). Trait Activation Theory explains how situational cues trigger these dark traits (Tett & Burnett, 2003). Building on recent work that highlights the spread of dark traits in higher education (*Alowais & Suliman, 2025*) and the importance of ethical philosophies

 doi:10.20944/preprints202602.0850.v1

for sustainable strategies (*Alowais, 2024*), this study examines the paradox of mentorship as both duty and manipulation.

This study delves deeper into the dynamics identified by *Alowais and Suliman (2025),* who demonstrated how Leader Dark Triad (LDT) traits can cascade downward and be mirrored in Employee Dark Triad (EDT) behaviors within higher education institutions. While their work highlighted the contagious nature of dark leadership, the present paper explores a specific paradox within this terrain: the Machiavellian mentor who cloaks manipulation in the guise of care. Ethical philosophies, as argued in *Alowais (2024),* remain central to any sustainable business or institutional strategy, yet this study reveals how such philosophies can be co-opted and twisted, turning mentorship from a pathway of empowerment into an instrument of control. By examining this paradox, the research underscores how ethics, when weaponized, not only lose their purpose but also destabilize trust, morale, and organizational culture.

Mentoring is a highly encouraged business ethic where organizations are meant to guide the inexperienced employees on their way to professional development. Managers will be expected to offer guidance, encouragement, and assistance, making sure that the knowledge will be transmitted and career growth will be supported (Allen et al., 2017). These practices also aid in the development of trust in the workplace, the development of skills, and the overall moral aspect of organizational life. However, regardless of this optimistic framing, mentorship is not necessarily altruistic. Some leaders, particularly those with Machiavellian tendencies, strategically adopt the appearance of supportive mentors while using the relationship for self-serving ends. This paradox is the centerpiece of the current research: the moral obligation of mentorship and the sophistry of masquerading it.

The leaders of Machiavellianism are described as manipulative, tactical, and individualistic (Collison et al., 2018). The latter is used in leadership to make a leader seem caring and principled without overt reliance, extracting loyalty, or advancing their own interests (Belschak et al., 2018). Ethical branding of mentorship is thus camouflage-producing trust superficially and draining it out in actual practice. This tension between supportive appearance and exploitative reality highlights an urgent need for deeper analysis of mentorship in organizational settings.

The paper follows three interrelated aims as a reply to this paradox: (1) To understand how ethical Machiavellian political leaders apply mentorship narratives to conceal manipulation. High Machiavellian leaders can use the language of ethics and professional duty as a tool to build credibility and to persuade mentees. (2) To learn about the reconciliation of the feelings of gratitude with the feelings of exploitation by mentees. Mentees tend to listen to and take advice as well as like the chances that they have, estimating that at the same time, they start feeling oppressed by being manipulated, which mingles gratitude and resentment. (3) To assess whether such "ethical" manipulation produces long-term trust erosion in the organization. While manipulative mentorship may yield short-term productivity, the eventual recognition of manipulation can fracture trust, damaging morale and team cohesion. These aims not only focus on the experiences of mentees but also on the wider cultural and ethical outcomes of the influence and cultural practice of mentorship in organizations.

Based on the above objectives, the following three major research questions will be explored: (RQ1) What are practitioners making of mentorship that seems ethical yet manipulative? (RQ2) Which organizational cues allow leaders to sustain this two-facade image? (RQ3) What is the impact of this paradox of care and control on team morale? Such questions will attempt to record the lived life of mentees as well as the organizational circumstances that foster mentorship with such contradiction as to prosper.

The research relies on dark leadership and organizational ethics theories to conduct its investigation. Prior research suggests that Machiavellian leaders often thrive in ambiguous contexts, where ethical narratives can be co-opted for strategic purposes (Cai et al., 2024). These shows a paradox of care and control to mentorship. Mentees make themselves present but might later find some hidden agendas. Three likely themes frame this tension: (1) "The Wolf in a Scholar's Robe." This metaphor explains that professional ethics and mentorship discourse are disguised as

 doi:10.20944/preprints202602.0850.v1

camouflage in the face of manipulation. Leaders look and act interested in their mentees but have an ulterior agenda. (2) "Debts That Never End." This sense of obligation, by eliminating the norm of reciprocity, can be established by mentorship. Mentees may feel indebted to their mentors for guidance, even when that guidance is exploitative or self-interested. This repetitive moral debt is an engine of control. (3) "Trust Fractures." Mentees and teams feel disappointed when so-called ethical branding turns out to be a lie. What was once a source of trust and support was becoming evidence of deception, creating long-term organizational cynicism and reduced morale. These themes indicate that mentorship is not only an interpersonal practice but a structural dynamic capable of reinforcing or weakening organizational trust.

This study complements Alowais and Suliman (2025), which demonstrated how Leader Dark Triad traits cascade into employee behavior, and Alowais (2024), which stressed the centrality of ethical philosophies in sustainable strategy. By focusing on Machiavellian mentorship, it shows how ethics can be weaponized to disguise manipulation. Together, these works highlight a coherent agenda: understanding how dark leadership spreads, why ethics matter, and how hidden manipulation erodes trust and performance. The paper makes three contributions to the areas of leadership and organizational ethics. It introjects into the field of research on dark leadership by contextualizing Machiavellianism in the conditions of mentorship, which is viewed as an ethical field. Second, it gives a deeper understanding of the way mentees overcome the psychological and emotional inconsistencies involved in mentorship by finding a balance between gratitude and suspicion. Third, it discusses the organizational consequences as the manifestations of the lack of trust in connection with the presence of manipulative mentorship, and the impact on morale, retention, and ethical culture. By so doing, the research notes the necessity of critically revisiting the practices of mentorship instead of presuming it to be ethically sound.

2. Literature Review

Leadership and mentorship are often framed as ethical duties, yet research shows they can also conceal manipulation and self-interest. Theories such as Jung's Shadow Theory (1959/1990), the Toxic Triangle (Padilla, Hogan, & Kaiser, 2007), and Trait Activation Theory (Tett & Burnett, 2003) provide lenses to uncover how dark traits emerge in mentorship settings. Recent studies highlight how these dynamics are particularly acute in higher education, where leader traits cascade into employee behaviors (*Alowais & Suliman, 2025*) and where ethical philosophies, though vital, are vulnerable to distortion (*Alowais, 2024*). This review explores mentorship as ethical duty, the dark side of leadership, ethics as camouflage, mentee experiences of gratitude versus exploitation, and the broader organizational consequences of such dynamics.

2.1. Mentorship as Ethical Duty

Mentorship as an ethical duty can mask hidden motives, as Jung's Shadow Theory shows leaders projecting unconscious drives under a caring façade (Jung, 1959/1990). The Toxic Triangle explains how destructive mentors exploit willing followers in permissive contexts (Padilla, Hogan, & Kaiser, 2007). Trait Activation Theory adds that situational cues, such as crises or career opportunities, trigger Machiavellian behaviours disguised as support (Tett & Burnett, 2003).

Mentorship can be flagged as a pillar of ethical leadership, and there might be organizational life. Professional codes of ethical conduct in medicine, law, and education reveal the responsibility of more senior professionals in mentoring their younger counterparts in adhering to necessary knowledge maintenance and professional ethics across generations (Allen et al., 2017). Mentorship is a popular part of organizations structured as part of their official code of ethics and training as an established moral obligation in correspondence with corporate social responsibility (CSR). Incidentally, CSR focuses on the concept of external accountability to the communities, and the idea of internal duty to the development and well-being of employees (Macassa et al., 2021). Mentorship is considered a leadership role that aids in developing employees while ensuring equality and equity in accessing the career ladder.

 doi:10.20944/preprints202602.0850.v1

The authority and care in professional responsibility demonstrates that leaders are expected to be moderate in mentorship. According to the ethical leadership theory, a leader's role is to be a good example of fairness, honesty, and caring about others (Brown & Treviño, 2006). Through mentoring, the leaders exhibit readiness towards an investment in the people, making the act not just a managerial decision but an ethical decision in line with the organizational integrity compounds.

The beneficial efforts of mentorship are well written. It has been found through empirical studies that mentoring promotes career growth through knowledge, skills, and social capital to those involved (Eby et al., 2013). Mentees face the advantages of faster promotions, higher job satisfaction, and better performance results (Allen et al., 2019). Another essential advantage is psychosocial support because mentors tend to offer encouragement, role modeling, and reassurance, leading to the establishment of professional identity and resilience at work (Mann & Gaufberg, 2016).

Mentorship also helps to associate with organizational culture. Successful mentoring relationships induce a sense of inclusiveness, strengthen mutual values, and instill the norms of an institution (Baker et al., 2020). In systems thinking terms, mentorship enhances the social glue of the organizations and inculcates moral principles and teamwork in organizational activities on a rank-order scale. These advantages explain why mentorship has been given much focus in leadership models in organizations and initiatives in corporate social responsibility. However, despite championing the moral imperative of mentorship, researchers have been concerned about making blanket predictions that mentorship practices are always bad (Pizzolato & Dierickx, 2023). The trust and reliance that are created by mentorship can turn into the venue of exploitation when leaders with darker characteristics use the relationship to their own benefit.

2.2. The Dark Side of Leadership & Mentorship

The dark side of mentorship reflects Jung's Shadow Theory, where leaders conceal destructive impulses behind guidance and care (Jung, 1959/1990). The Toxic Triangle explains how toxic leaders, vulnerable mentees, and permissive institutions sustain manipulative mentoring (Padilla, Hogan, & Kaiser, 2007). Through Trait Activation Theory, organizational pressures activate Machiavellian, narcissistic, or psychopathic traits that turn mentorship into exploitation (Tett & Burnett, 2003).

Christie and Geis (1970) stated that the term Machiavellianism was subsequently translated into the broader and encompassing philosophy of the Dark Triad of personality traits (namely related to narcissism and psychopathy) (Paulhus & Williams, 2002). Machiavellians are scheming, scheming, and strategic characters who tend to advance their interest instead of being ethically responsible. Machiavellian leaders in organizations play with ambiguity, leading people to become dependent upon them through charm and persuasion to conceal selfish interests (Belschak et al., 2018).

More recent research emphasizing Machiavellian leaders supports impression management by leaders as a means to look competent and benevolent and to manipulate employees to make them loyal and obedient (Jaiswal & Bhal, 2014). It is tempting to think that they act in a supportive way in mentorship, yet, in fact, they are using it either to acquire political allies, sensitive information, or promote their own personal status (Greenbaum et al., 2017). This two-sided leadership justifies the fact that not all mentorship is altruistic.

The term toxic mentoring summarizes the situations in which leaders can abuse the mentee role to dominate them or take advantage (Scandura, 1998). This kind of mentorship can be characterized by restrictions on autonomy, dependence, or loyalty to engage more than what is considered professional. In the case of mentees, this relationship may lead to stress, low confidence, and loss of faith in the company/ company values (Eby et al., 2013). These ideas have been further developed in more recent studies that have been used to connect manipulative mentorship and the literature on abusive supervision (Tepper, 2007). Such studies indicate that Machiavellian leaders can wrap up their egoism under the pretense of developmental instructions by taking advantage of their mentees' gratitude and loyalty (Walter, 2025). Such duality forms a paradox, as mentorship is an opportunity to grow and a trap on the way to it.

 doi:10.20944/preprints202602.0850.v1

The existence of both affirming actions and predatory intentions leads to theorizing the issue of a paradox. At least formally, mentorship is consistent with organizational ethics and CSR; at the same time, in practice, this relationship may become the vehicle of control and self-serving manipulation (F Swargiary , 2024). The drips in the form of signals have that overt praise and encouragement can be accompanied by implicit pressure or self-advantageous requests from mentees. This erosion of psychological safety is caused by such contradictions that propel mentees into the uncertainty of whether the intentions of the mentor are genuine.

2.3. Ethics as Camouflage

Ethics may serve as camouflage when leaders use moral language to mask manipulative intent, consistent with Jung's Shadow Theory, which highlights the projection of hidden drives under virtuous appearances (Jung, 1959/1990). The Toxic Triangle shows how such ethical fades thrive in contexts where followers are dependent and environments tolerate deception (Padilla, Hogan, & Kaiser, 2007). Trait Activation Theory explains how crises or mentoring situations trigger Machiavellian tactics, with "ethics" deployed as a cover for control (Tett & Burnett, 2003).

Impression management theory (Goffman, 2023) will be very effective in considering how leaders manage an impression of virtuous mentorship yet still have other hidden feelings. In the context of organizations, impression management implies a conscious self-presentation while aiming at the desired impact on other people's perceptions regarding the character of a specific individual and his intentions (Bolino et al., 2016). Most notably, Machiavellian leaders take advantage of impression management, whereby they position themselves as caring mentors who uphold organizational values, despite the reasons behind such arguments being self-serving. Such leaders use their role in perpetuating their image as well-wishers through their choice of carefully selected words, as opposed to symbolic gestures, and their affiliation with organizational rituals. This forms a safe cocoon in which it is hard to be accused by the mentees or workmates without jeopardizing their reputation.

Researchers observe the similarity of such an interaction with the phenomenon of surface acting in their emotional labor when their leaders can show empathy or caring at the surface but stay detached internally (Humphrey, et al., 2025). The ethical persona, therefore, is a variety of disguises-an organizationally approved performance that disguises manipulation. As we can see, a leader can endorse mentorship programs and place them publicly, but they can be used to increase personal power or guarantee allegiance.

Recent studies point out how ethical branding of leaders may add to this disguise. Brown and Treviño (2004) state that ethical leaders may strategically enact the act to increase their legitimacy instead of acting in compliance with it. Ethical rhetoric is an effective instrument of control when coupled with impression management: the mentees have little option but to trust their mentors, who project organizations' values on the surface, despite the exploitation behind the scenes. Finally, impression management would enable the leaders in question, being unethical, to capitalize on the fact that there is a certain degree of ambiguity between image and reality and keep the power intact by dressing the self-interest in a language of integrity.

The problem of ethical leadership branding also makes the description of camouflage and impression management more complex. The model of ethical leadership applied by Brown and Treviño (2006) sees leaders as models of ethical leadership who take the role of fair, imbued moral leaders and care about the stakeholders. In organizations, it is frequently operationalized in the form of mentorships, the set of code of conduct, or symbolic communication within which the adherence of a leader to ethical principles is emphasized. Nevertheless, it is found that ethical leadership may also be acted instead of being genuine (Ng & Feldman, 2015). As in the case of some leaders, presenting the image of an honest nurturer may yield legitimacy and loyalty that can be exploited in the future to serve selfish interests.

Such strain is seen especially in ethical signaling. Moral authority may be displayed through speeches, policy endorsements, or through high-profile mentorship programs, where, in warfare,

 doi:10.20944/preprints202602.0850.v1

strong-suited leaders are promoting personal ambitions behind the scenes (Tourish, 2019). It makes ethical branding not so ethical, but it works on the image of the guardianship. These reputation defenses hinder the quest to inquire, as anyone who questions a leader perceived as highly ethical does face the risk of social and organizational repercussions (Neves & Story, 2015).

This phenomenon is epitomized in the metaphor of the development of the wolf in a scholar's robe. Leaders clothe themselves with the attributional clothes of ethics and professional status in a bid to neutralize doubt to ensure trust. This presents a paradoxical situation for mentees in that they find comfort with the moral standing height of the mentor. However, there is the possibility of a dependent, indebted, and exploitative relationship. In the long term, this inconsistency will undermine credibility upon striking an observation of how the branding of ethical leadership could be used to cover the opportunistic trait instead of assuring ethical behaviors.

2.4. Mentee Experience: Gratitude vs. Exploitation

Mentees often feel torn between gratitude and exploitation, a tension illuminated by Jung's Shadow Theory, where leaders' hidden motives distort authentic relationships (Jung, 1959/1990). The Toxic Triangle explains how susceptible followers may internalize manipulation when mentorship occurs in permissive settings (Padilla, Hogan, & Kaiser, 2007). Trait Activation Theory adds that situational cues of dependency or career advancement trigger Machiavellian leaders to exploit mentees' sense of obligation (Tett & Burnett, 2003).

Mentorship is often positioned as a mutually advantageous cooperation, and reciprocity in most cases has an overly evolved proportion with mentees' perceptions and reactions towards their mentor. Based on the foundational norm of reciprocity of Gouldner (1960), the inherent behavior in an individual is the compulsion to pay back what is provided to one. This norm is extreme in the mentoring setting since the assistance that mentors provide career advocacy, sponsorship, and psychosocial encouragement that has a life-altering dimension in most cases for the mentees. Consequently, the mentees might internalize a superior sense of indebtedness, which is well beyond the norm of workplace transactions (Steadman, 2017). Although in health-related relations, reciprocity leads to trust and loyalty, in exploitative relations, reciprocity is used as an adequate control. Machiavellian tutors might undermine their generosity to guarantee compliance, silence, or further loyalty.

In recent studies, using reciprocity to justify exploitative practices by mentees has visible effects (Shore et., 2015). Stocks (2022) discovered that there are occasions when mentees rationalize that they are overworked or sidelined because the mentor invested in their careers, and the mentee pays them back. This reasoning shows gratitude becomes an enormous psychological debt that mentees cannot settle. This brings about states of asymmetric dependency, with the mentees being of the notion that they need to always provide services of compliance, even when the mentors act in ethically dubious ways.

The need to stay devoted even when the relationship is exploitative can easily initiate the concept of cognitive dissonance, which is a mental struggle when members of social establishments have conflicting ideas or realities (Festinger, 1957). In the case of mentees, Dissonance arises when the perceived real rewards of the mentorship, in the form of promotions, skill expansion, or a vaster network, exist alongside manipulative or self-regarding practices on the part of the mentor. This conflict compels mentees to deal with the conflict between gratitude and exploitation. Studies indicate that this kind of dissonance may also make them feel stressed and reduce job satisfaction, not to mention that the type of stress may steer the person to rationalize the unethical behavior to sustain psychological harmony (Balamurugan & Divyabharathi, 2021).

This dissonance has a high-value risk, especially in an organizational environment dominated by power asymmetries. Most mentees fear uttering grievances because their speaking might hurt their credibility or lead to application of withholding tactics (Cotton et al., 2011). Silence then becomes a strategy of coping with others, but it bolsters the superiority of the mentor. In the long run, the

 doi:10.20944/preprints202602.0850.v1

mentees can even rationalize exploitation and expect themselves to believe that the gains exceed the suffering, hence creating steps of addictive dependency and obedience.

Unfortunately, such debt often becomes perpetual, symbolizing the lasting psychological burden of reciprocity norms within exploitative mentorship relationships. In contrast to any financial or contractual debt, which is limited and quantifiable, principled mentoring debt tends to be emotional and moral regarding obligation due to loyalty, gratitude, or personal loyalty. The effect of this framing is that even when formal mentoring relations have end, mentees in question find themselves endlessly obliged (Balamurugan & Divyabharathi, 2021). The Machiavellian mentors take advantage of this indeterminacy; they portray themselves as life-long benefactors and thus remain structurally influential. Such an interplay disturbs the line between professional duty and personal indebtedness. It can also make mentees feel guilty for not breaking the relationship, even when it turns out to be partially harmful, since it will be seen as a form of betrayal to the mentors who helped them in the relationship. In this way, gratitude transforms into a psychological burden, and mentorship evolves from a developmental relationship into a control mechanism. The metaphor of debts that cannot be cleared highlights that the ideas of reciprocity, acted upon, put the mentees in loops of reciprocation that may continue to exist throughout their careers, making it challenging to understand when someone really cares and when they are taking advantage of you.

2.5. Organizational Consequences

Organizational consequences emerge when mentorship is weaponized, as Jung's Shadow Theory shows how concealed destructive traits can shape collective norms (Jung, 1959/1990). The Toxic Triangle demonstrates how toxic mentors, compliant followers, and permissive institutions together normalize manipulation across systems (Padilla, Hogan, & Kaiser, 2007). Through Trait Activation Theory, pressures like performance reviews or crises activate dark traits that corrode trust, weaken morale, and destabilize culture (Tett & Burnett, 2003).

On an organizational scale, manipulative mentorship overwhelmingly affects employees' trust and psychological safety, the two key components of a well-functioning workplace culture. According to Edmondson (1999), psychological safety can be referred to as the shared feeling that one can share ideas, concerns, and mistakes without the fear of being adversely affected. However, this kind of safety is usually destroyed in work environments where they are immersed in manipulative mentorship. Additionally, mentees who maintain perverse unwritten debts or other forms of subtle threat are not inclined to be utterly frank because they are concerned that their sincerity could lead to retaliation, loss of face, or being left out of the processes of promotion or advancement. The immediate mentor pair does not just experience this silencing influence. However, it follows the teams where others are likely to observe the dangers of protest and internalize caution on their part (Fernandez, 2008).

Innovation and collaboration decrease with the level of psychological safety. When introduced into teams that work in environments of mistrust and suspicion, creative thinking and problem-solving are frequently suppressed by teams that have acquired defensive communication styles. Furthermore, when mentors take advantage of the language of ethics but are not so keen to cover their manipulative intent, they misrepresent the climate within the organization, leaving people uncertain as to which actions are accepted, rewarded, or not (Ensher & Murphy, 2008). Such a grey area does not encourage workers to refer to their full extent, as organizational promotion regulations are geared to individual loyalty rather than organizational principles.

Trust, when broken, is infamously hard to mend. According to Furnham and Taylor (2004), workers who discover manipulation behind a moral facade often feel betrayed, a feeling that can lead to cynicism. This kind of betrayal is especially harmful as ethical leaders are supposed to be an example; when their integrity shows that it is a performance rather than genuine, employees tend to direct the level of mistrust away from the single, individual mentor and towards the leadership framework in general. Such trust fissures reverberate through the organizational webs, compelling workers to reevaluate the sincerity of corporate ethics tone and leadership pledges.

 doi:10.20944/preprints202602.0850.v1

Manipulative mentorship has organizational implications significantly beyond damage to immediate relations. According to Loi et al. (2014) workers who are disatisfied often check out of their jobs, resulting in low morale and diminished organizational identification. With the proliferation of cynicism, organizational initiatives may be met with skepticism, especially those conceptualized based on ethics, mentorship, or social responsibility. When the truth behind a time-saving facade is revealed, employees doubt further efforts, perceiving them as tokenism instead of genuine attempts to save time, leading to distrust and a lack of engagement (Ehrich et al., 2018).

Another critical impact is turnover. Workplaces often become intolerable to employees who are exploited or led down the wrong path and would rather move to an organization that better suits their personal values. High turnover not only impacts the team cohesiveness but also creates financial and cultural drain on the organization. Notably, the resignation of disappointed workers can have a silent testimony effect on others, telling them that checking ethically oriented branding will not work and lead to more departures. Systemically, manipulative mentorship undermines wider corporate social responsibility (CSR) and ethics initiatives. Credibility and trust are essential pillars of CSR initiatives, and whenever employees sense that leadership is somehow manipulative, the ethical command of these initiatives is lost. As a result, even good intentions become discounted as strategic PR efforts, compromising their ability to bring any significant cultural change into existence (Ensher & Murphy, 2008).

The image of trust fractures summarizes the underlying organizational harm of manipulative mentorship. Similar to inclusion cracks in a purely physical system, breaches of trust can initially appear isolated to specific relationships. However, these rifts propagate over time across teams and departments, undermining organizational culture. Trust dislocations occur subtly: withholding information, encroaching communication, unwillingness to work together, and increased turnover intentions. Superficial reforms like codes of ethics or training programs cannot mend these fractures. These actions can appease employees in the short term, but they do not necessarily penetrate the cultural causes of manipulation. Sustainability in repairing needs governance through open leadership approaches, accountability systems, and a culture that enables employees to express complaints without fear of reprisal. Organizations must look beyond ethical branding and focus on policies that prevent exploiting workers in the guise of mentorship (Brown & Treviño, 2014). Through these underlying cultural problems, organizations can restore the integrity upon which trust flourishes.

3. Methodology

Understanding the dynamics of dark traits required moving beyond numerical indicators and survey-based measures. As argued by Alowais and Suliman (2025), qualitative approaches were necessary to capture subtle forms of manipulation, concealed motives, and lived experiences that quantitative methods could not adequately reveal. Guided by this perspective, the study adopted a qualitative design using semi-structured interviews to examine how mentorship, ethics, and Machiavellian strategies were experienced by employees. This approach enabled depth, nuance, and theory-building that would not have been achievable through quantitative techniques alone.

3.1. Research Approach

The study followed a qualitative, exploratory research design to examine mentorship as both an ethical obligation and a potential mechanism for manipulation. A qualitative approach was particularly appropriate because it enabled an in-depth understanding of lived experiences, subjective perceptions, and subtle power dynamics embedded within mentor–mentee relationships (Creswell & Poth, 2018). While quantitative surveys may capture prevalence, they would not have revealed the emotional complexity associated with gratitude, exploitation, and impression management that characterized the phenomenon under investigation. An exploratory design was therefore adopted to allow emergent themes to surface in an area that remains insufficiently theorized, particularly in relation to manipulative mentorship framed as ethical guidance.

 doi:10.20944/preprints202602.0850.v1

3.2. Participants

The study engaged sixteen participants drawn from a range of Higher Education Institutions (HEIs) in the UAE. Participants included lecturers, senior lecturers, assistant professors, associate professors, and academic administrators, representing both teaching and leadership roles across business, social sciences, and applied disciplines. This diversity ensured that perspectives reflected mentoring practices across different levels of academic and institutional responsibility. A sample size of sixteen participants was considered methodologically sufficient, as thematic saturation was reached when subsequent interviews generated overlapping insights without producing new themes. In qualitative research, depth and richness are prioritized over breadth, and saturation commonly occurs with samples of 12–15 participants (Guest, Bunce, & Johnson, 2006). Expanding the sample further was neither feasible, given the limited pool of Emirati faculty within HEIs, nor analytically justified, as additional interviews were unlikely to yield novel insights. Participants included eight lecturers, three senior lecturers, three assistant professors, and two associate professors.

3.3. Sampling Strategy

Purposive sampling was employed to identify participants who had experienced both supportive and problematic mentoring relationships. This strategy was appropriate because the study aimed to contrast genuine mentorship with forms of exploitation occurring under ethical pretenses (Etikan et al., 2016). Inclusion criteria were limited to employees with a minimum of three years of professional experience to ensure sufficient exposure to formal or informal mentorship relationships. Snowball sampling was also used selectively to reach participants, willing to share sensitive experiences that may not have been accessible through direct recruitment.

3.4. Data Collection: Semi-Structured Interviews

Data were collected through semi-structured interviews, which balanced consistency with flexibility. An interview guide was developed around core questions addressing perceptions of mentorship, experiences of support versus constraint, and interpretations of organizational ethical messaging. This format allowed participants to elaborate on personal experiences, introduce unanticipated themes, and articulate complex emotional states such as gratitude, indebtedness, and betrayal (Kallio et al., 2016). Interviews lasted between 20 and 40 minutes and were conducted either in person or via secure video conferencing platforms, depending on logistical considerations. Guiding questions explored factors contributing to the deterioration of mentorship relationships, experiences of simultaneous support and constraint, organizational cues shaping mentorship perceptions, and the tension between gratitude and perceived manipulation. These questions were designed to elicit reflective accounts of mentorship paradoxes.

3.5. Data Analysis

Data were analyzed using thematic analysis as outlined by Braun and Clarke (2006), following six stages: familiarization with the data, generation of initial codes, theme development, theme review, theme definition, and report production. Thematic analysis was well suited to the study as it supported both inductive and deductive coding. Inductively, it allowed patterns to emerge from participants' narratives, while deductively it enabled engagement with established theoretical constructs such as impression management, norms of reciprocity, and cognitive dissonance. Although NVivo can provide structure and comparability in large qualitative datasets (Mortelmans, 2019), exclusive reliance on software risks reducing interpretation to mechanical coding. As demonstrated by Alowais and Suliman (2025), manual thematic analysis offers greater contextual sensitivity, particularly in studies of dark traits where tone, contradiction, and implicit meaning are critical. NVivo was therefore used as a supporting tool, while primary analysis and interpretation remained manual.

 doi:10.20944/preprints202602.0850.v1

3.6. Ethical Considerations

Ethical rigor was particularly important given the study's focus on power, exploitation, and organizational culture. Participants provided informed consent and were informed of their right to withdraw at any time. Pseudonyms were used in transcripts and reporting, and identifying information, including institutional names, was anonymized. Given the potential risk of retaliation, particular care was taken to ensure participant anonymity. Interviews were conducted in secure environments, either in private physical settings or through encrypted online platforms. All data were securely stored and encrypted to protect recordings and transcripts (British Psychological Society, 2021). Ethical approval was granted by the British University in Dubai (BUiD) Ethics Committee under a low-risk self-assessment procedure, as the study formed part of an ongoing PhD thesis.

3.7. Reflexivity

Researcher reflexivity was treated as an integral component of the study. As qualitative scholars note, researchers inevitably bring assumptions, values, and professional experiences into the interpretive process (Lichterman, 2017). Reflexivity was addressed through the maintenance of a research journal documenting analytic decisions, emerging biases, and interpretive reflections throughout the study. This process helped mitigate the risk of overemphasizing either positive or negative mentoring experiences. Member checking was also employed to enhance credibility, with thematic summaries shared with participants to confirm resonance with their experiences. This process strengthened interpretive validity and ensured that findings reflected participants' perspectives rather than solely the researcher's interpretations (Birt et al., 2016).

4. Themes from Interviews

The semi-structured interviews generated a rich set of narratives that reveal the complexity of mentorship experiences. Although many participants mentioned that mentorship is a healthy experience, in their narratives, there was an innate mention of the darker side, whereby gratitude, ethical branding, and rituals of the organizations were wielded as weapons. Thematic analysis yielded five major motifs: The Wolf in a Scholar's Robe, Debts That Never End, Trust Fractures, Ambivalence of Gratitude, and Signals of Dual Image. These themes described how the participants managed to navigate through mentorship relations based on both sincere assistance and manipulative undertones. In the results section, participants' direct quotations will first be presented in their original form to illustrate lived experiences. Following each set of quotations, the thematic analysis will be provided as a separate commentary, rather than embedding interpretation within the narrative. This structure ensures transparency, allows participants' voices to stand on their own, and enhances the credibility of the thematic coding process.

4.1. The Wolf in a Scholar's Robe

1. "He quoted the ethics charter in every meeting but then asked me to pick up his children after class." Lecturer 1
2. "In public panels, she introduced me as a success story, but in reality, I was doing background work she took credit for." Assistant Professor 1
3. "He would praise CSR values to the press, but in the office, he only cared about building his own empire." Senior Lecturer 2
4. "The contradiction was painful; how can someone preach responsibility so loudly and yet behave so irresponsibly behind closed doors?" Lecturer 3
5. "I was told I was the 'future of the institution,' but my tasks looked more like a secretary's than a scholar's." Associate Professor 1
6. "Whenever there was a photo opportunity with CSR, he made sure I was in the frame, but when it came to actual research, I was excluded." Lecturer 4

 doi:10.20944/preprints202602.0850.v1

7. "It felt like being part of a stage play; ethics was the costume, but manipulation was the script." Assistant Professor 2
8. "In official speeches he praised mentorship, but in practice, he used my work to push his projects forward." Lecturer 5
9. "We were constantly told about 'integrity' in emails, but I saw integrity disappear in the daily treatment of staff." Senior Lecturer 3
10. "I started to feel I was there only as proof that he was a good mentor, not because my growth mattered." Lecturer 6
11. "He called himself an ethical leader, but the only ethics I saw were for reputation management, not for people." Associate Professor 2
12. "At first I admired his talks on values, but then I realized they were shields to cover his self-interest." Lecturer 7

The participants recounted experiences with mentors who projected principled, ethical identities while privately pursuing self-interest. These people might have put forward their personalities as role models, often with reference to business values or corporate social responsibility ideologies, but the mentees have found discrepancies between what was said and what was done.

They mentioned a participant that their mentor was the type of leader who could mention the ethics code in a meeting and then abuse you to run personal errands. This depicts the allegory *of the wolf in a scholarly robe*: a person who disguises opportunism as the guise of decency. Such leaders proved to build trust by undertaking symbolic activities like participating in CSR activities, supporting mentorship activities, or even praising mentees during press conferences, but utilized this credibility to hide disagreeable activities.

The theme brings out how challenging the mentees found it to deal with an authoritative figure who so successfully identified their persona with that of ethics within the organization. Moral camouflaging rendered the mentees unable to differentiate between actual support and manipulative purpose, and they frequently became lost and confused about their right to voice their concern.

The participants' accounts consistently revealed the paradox of leaders who *performed ethics* while simultaneously exploiting those they mentored. By quoting codes of conduct, referencing CSR, or showcasing mentees in public forums, these mentors cultivated a polished moral persona. Yet, behind this façade, they delegated personal errands, withheld recognition, and appropriated mentees' work for self-promotion.

The dissonance between the ethical image and the manipulative reality created what many described as a *stage play*, where ethics was the costume and manipulation the script. This aligns with Jung's Shadow Theory (1959/1990), which suggests hidden destructive impulses are projected as virtue, and with the Toxic Triangle (Padilla, Hogan, & Kaiser, 2007), showing how destructive leaders exploit vulnerable followers in permissive contexts. Similarly, Trait Activation Theory (Tett & Burnett, 2003) explains how organizational crises or expectations can trigger these dark behaviors. The result was a climate of mistrust and confusion: gratitude for symbolic gestures conflicted with resentment at daily treatment, eroding trust in mentorship as a meaningful institutional practice.

4.2. Debts That Never End

1. *At first, I was grateful for the advice, but later it felt like every Favor had strings attached." Lecturer 2*
2. *"He reminded me often that I owed my promotion to him, so saying no was never an option." Senior Lecturer 1*
3. *"What began as guidance slowly turned into unpaid obligations such as committee work, errands, tasks outside my role." Assistant Professor 2*
4. *"She always stressed loyalty, and if I hesitated, I was made to feel disloyal or ungrateful." Lecturer 4*
5. *"I started to feel my debt was endless, like I would never stop repaying for the 'opportunities' I got." Associate Professor 1*
6. *"When I tried to set boundaries, he said: 'Don't forget who opened doors for you.'" Lecturer 6*
7. *"It wasn't mentorship anymore; it was moral blackmail disguised as gratitude." Senior Lecturer 3*

 doi:10.20944/preprints202602.0850.v1

8. *"They framed it like a family bond, you don't refuse family, but really it was control." Lecturer 5*
9. *"Even after the mentorship ended, I carried that sense of obligation, like a shadow I couldn't shake." Assistant Professor 1*

Debts that never *end are the second theme*, which talks of how norms of reciprocity were manipulated to create an open-ended obligation. Participants often commented that mentors stressed the value of loyalty and gratitude and presented in their guidance the idea that it was a gift that had to be repaid over time. Many of the mentees believed this dynamic grew stronger with time. An example is that one participant stated that it began with me feeling grateful for their guidance, but it became me having to do extra work that was not my actual job. What began as mentorship shifted toward an open-ended sense of indebtedness? These ever-present duties offered some sort of restraint. Mentors would remind mentees about previous chances in their lives in making ways of being directed at promotion, access to networks, or professional sponsorship, making resistance to requests almost equivalent to betrayal. It was through this framing that the distinction between voluntary reciprocity and forced compliance became blurred. Mentees often internalized the belief that refusing would jeopardize their careers, producing a long-term psychological tether that extended even after formal mentorship ended.

This theme highlights how mentorship was reframed as a cycle of perpetual indebtedness, where gratitude evolved into restraint and compliance. What began as voluntary reciprocity was transformed into coerced loyalty, with mentees pressured to equate refusal with betrayal. The process resonates with Jung's Shadow Theory (1959/1990), as leaders project generosity outward while concealing manipulation. It also reflects the Toxic Triangle (Padilla, Hogan, & Kaiser, 2007), where destructive leaders exploit followers' vulnerability through moral leverage. Finally, Trait Activation Theory (Tett & Burnett, 2003) explains how situational cues—such as promotions or access to networks—activate manipulative strategies that bind mentees psychologically. The outcome was a persistent tether of obligation that blurred the line between mentorship and exploitation, eroding autonomy and trust.

4.3. Trust Fractures

1. *"The day I realized my mentor's guidance was just for his own gain, something broke inside me." Lecturer 3*
2. *"It wasn't just him I stopped trusting; it was the whole system that called him an ethical leader." Senior Lecturer 2*
3. *"I used to admire him, but once the mask fell, I felt fooled and naïve." Assistant Professor 2*
4. *"After that experience, I could never look at mentorship programs the same way again." Lecturer 6*
5. *"The betrayal wasn't personal only felt institutional." Associate Professor 1*
6. *"Once you see the hidden agenda, you start questioning every kind word as manipulation." Lecturer 7*
7. *"I lost faith in the mentorship scheme; it was just a façade to polish the university's image." Assistant Professor 3*
8. *"Even genuine leaders now face my scepticism I can't tell who is authentic anymore." Lecturer 2*
9. *"The fracture spreads fast—colleagues avoid programs, nobody trusts leadership workshops now." Senior Lecturer 1*
10. *"Psychological safety is gone; every act of care feels like a trap." Lecturer 5*

The powerful disappointment which followed the realization of mentees of discrepancies between the ethical image of a figure they had emulated and the manipulative methods the latter enacted was a recurrent theme. These "trust fractures" often emerged suddenly when a hidden agenda surfaced or gradually, as mentees pieced together patterns of self-serving behavior.

Some respondents explained their experiences of betrayal not to the specific mentor but to the organizational system as a whole that had recognized them as ethical leaders. To others, the discovery that ethical branding was a sham elicited profound cynicism. One interviewee commented, I lost faith in the entire mentorship program. It was more of a front to put the company in a better light.

 doi:10.20944/preprints202602.0850.v1

These fractures spread beyond the mentor–mentee dyad, undermining collective trust in organizational culture. The workers were more reserved, reluctant to participate in mentoring programs, and distrustful of leadership programs. The long-term effect was erosion of psychological safety, as individuals questioned whether apparent care and guidance were authentic or instrumental.

This theme captures the deep rupture mentees felt when the ethical image of mentors was unmasked as manipulation, leading to enduring trust fractures. The dissonance generated profound cynicism, extending from individuals to the wider organizational system. This aligns with Jung's Shadow Theory (1959/1990), where the hidden destructive side of leaders is eventually revealed, shattering projections of virtue. The Toxic Triangle (Padilla, Hogan, & Kaiser, 2007) illustrates how such leaders flourish in permissive systems that brand them as ethical while enabling harm. Finally, Trait Activation Theory (Tett & Burnett, 2003) explains how crises or organizational opportunities activate deceptive strategies, which when exposed, erode psychological safety and collective trust. These fractures undermined not only mentorship relationships but also institutional credibility, discouraging participation in future leadership programs.

4.4. Ambivalence of Gratitude

1. *"I can't deny he opened doors for me, but the price I paid was my peace of mind." Lecturer 2*
2. *"I was thankful for her support, yet every task felt like repayment, not growth." Senior Lecturer 1*
3. *"Gratitude kept me silent; I didn't want to seem unfaithful even when I felt exploited." Assistant Professor 2*
4. *"It's confusing, you feel lucky to be supported but guilty for doubting their motives." Lecturer 4*
5. *"Sometimes I justified the manipulation as the cost of advancing my career." Associate Professor 1*
6. *"I told myself I should be grateful, but deep down I was resentful." Lecturer 6*
7. *"They always reminded me of what they had done for me, so questioning them felt wrong." Senior Lecturer 3*
8. *"I felt conflicted, was it betrayal to dislike someone who gave me opportunities?" Lecturer 7*
9. *"My gratitude became a chain; it stopped me from speaking up about unfair treatment." Assistant Professor 1*
10. *"I lived with cognitive dissonance: respect for their help, anger at their control." Lecturer 3*
11. *"Even now, I wrestle with guilt when I recall how I resented my mentor." Senior Lecturer 2*

Gratitude emerged as a double-edged emotion within mentorship. The participants continuously reported that they valued career opportunities, knowledge transfer, and advocacy carried out by the mentors. However, with gratitude, most also talked of resentment when they felt manipulated or overstretched. This strain gave rise to a higher order of gratitude, in which affirmative feelings existed with unease and deception.

Other mentees were trying to justify dubious conduct as a reasonable price of entry into a higher ladder. Still others are haunted by feelings of guilt because they have negative feelings about a mentor, as they suppose they ought to have positive feelings. This struggle resonates with the concept suggested by Festinger (1957) called cognitive dissonance, in which individuals are trying to find a solution to conflicting beliefs. Indicatively, one interviewee confessed*:" I had to give them credit on their part because they had opened doors in my career; however, I had to do nothi*ng *about how confined I was*". This ambivalence even influenced the silence of mentees because the ethical pressures of gratitude complicated their ability to express grievances without appearing ungrateful or unfaithful. The emotional burden of this tension often extended beyond the mentorship, affecting mentees' broader sense of workplace trust and self-worth.

This theme highlights the paradox of gratitude as both empowering and constraining. While mentees valued career opportunities and advocacy, gratitude became entangled with resentment and guilt, creating a climate of cognitive dissonance (Festinger, 1957). Drawing on Jung's Shadow Theory, these emotions reveal the hidden manipulation behind mentors' apparent benevolence. The Toxic Triangle (Padilla, Hogan, & Kaiser, 2007) also explains how gratitude can be exploited as a lever for

 doi:10.20944/preprints202602.0850.v1

compliance, while Trait Activation Theory (Tett & Burnett, 2003) shows how situational cues like promotions intensify this emotional ambivalence. The outcome was silence, strained trust, and reduced well-being, as mentees felt unable to voice concerns without betraying their sense of loyalty.

4.5. Signals of Dual Image

1. *"He never missed a chance to talk about integrity in meetings, but in practice he bent rules for his own benefit." Lecturer 2*
2. *"In public, she was the champion of values; in private, she was the biggest violator of them." Senior Lecturer 1*
3. *"It felt like a performance—ritualizing ethics speeches while acting otherwise behind closed doors." Assistant Professor 1*
4. *"He supported a few chosen people as proof of generosity, but the rest of us were ignored." Lecturer 4*
5. *"Selective kindness was the strongest tool—it built loyalty but also division." Senior Lecturer 2*
6. *"The institution hailed him as an ethical leader, which made it impossible for us to question him." Lecturer 5*
7. *"Her reputation created pressure to interpret my bad experiences positively, as if I was the problem." Assistant Professor 2*
8. *"When the company itself calls them the face of ethics, who dares to challenge the dual image?" Lecturer 6*
9. *"The endorsement from above made us second-guess our instincts was gaslighting in slow motion." Senior Lecturer 3*
10. *"I often wondered: is this genuine care or just another act for compliance?" Lecturer 7*
11. *"They had mastered impression management—it was branding, not mentorship." Associate Professor 1*
12. *"The ethical rituals were like camouflage; they provided cover for exploitative practices." Lecturer 3*
13. *"To us, it felt like the wolf hiding in a scholar's robe—respected outside, manipulative inside." Assistant Professor 3*

Lastly, the interviews showed signals that a mentor was always taking to stay both principled and unimpeded in their leadership. Patterns included: ritualizing ethics talk, where mentors often spoke about organizational values in front of people, yet acted contrary to such behaviors behind the scenes. Just like mentors, the selective generosity being shown was another aspect of mentors providing visible support to some mentees and denying support to others, usually as a method of building loyalty. This dual image was further negatively stimulated by organizational endorsement. Mentees felt an extra burden to interpret experiences positively when companies hailed mentors as ethical leaders. The surrounding context of the mentor's reputation thereby complicated the consistency of recognizing or questioning inconsistency. This theme brings out the functionality of the impression management strategy, not just at the personal level but also at the organizational level. Mentors established legitimacy, therefore covering exploitative practices by repositioning personal branding with institutional rituals. To mentees, these cues made it hard to tell whether particular acts were manifestations of genuine care or performances meant to get them to comply.

In summary, the themes captured during these interviews give a complicated picture of mentorship as both enabling and limiting. The moral masquerade of the grades who sham ethical labels via their academic hoodie is the wolf caught in the veil of the scholar. Debts that never end explains how a reciprocity norm can bond mentees in an everlasting sense of debt. The fractures of trust demonstrate the spread of disillusionment throughout organizations, dismantling morale and safety. The psychological dilemma of appreciating assistance and detesting being exploited consists in the ambiguity of gratitude. Lastly, there are the indicators of two pictures, highlighting the organizational practices that ground the legitimacy of duplicity. These results, combined, imply that mentorship is not just a developmental practice: it is also a location where ethics, power, and impressions have contact. These stories have spoken of the significance of critically examining the process of institutionalizing mentorship and how Madeleines have been guided through its contradictions.

 doi:10.20944/preprints202602.0850.v1

The accounts of mentors sustaining a dual image such as ethical in public, manipulative in private may illustrate how impression management functions as moral camouflage. This aligns with Jung's Shadow Theory (1959/1990), where the shadow persona hides destructive impulses behind socially acceptable masks. The Toxic Triangle (Padilla, Hogan, & Kaiser, 2007) also explains how organizational systems enable such duplicity by celebrating leaders as ethical exemplars, thereby legitimizing exploitation. Finally, Trait Activation Theory (Tett & Burnett, 2003) clarifies how institutional rituals and recognition cues trigger Machiavellian traits, turning mentorship into a staged performance of ethics. Collectively, these findings highlight that mentorship in such contexts becomes less a developmental practice and more a site where ethics, power, and institutional branding collide.

5. Discussion

The results suggest mentorship as an ambivalent environment in which empowerment and manipulation exist. The wolf in a scholar's robe and debts that go nowhere show that mentorship can be applied to career development and invidious influence. This is consistent with the new research that ethical leadership is a practice that can be borrowed as an impression management strategy, not a mode of authentic morality (Tourish, 2020s). Via self-branding, leaders who position themselves as ethical mentors develop reputational walls that compound the inability of mentees to detect or curb exploitation. The weaponization of indebtedness is accomplished through the persistent nature of reciprocity norms. Articles prove maladaptive reciprocity exerts enduring power effects (Cropanzano et al., 2017). The sense of duty intensified by gratitude among the mentees makes exploitative mentorship last undetected. The interviews pay an addition to this literature by demonstrating how such indefinite debts may turn into a defining choice of mentees even after the official engagement is terminated.

One significant organizational impact was trusting fractures such as: cheating by a professionally branded mentor severely depressed individual trust, as well as overall perception of organizational integrity. Recent investigations underline that once trust has been destroyed, it cannot be restored by a single act and rather the system cultural improvements (Ng & Feldman, 2015; Schyns & Schilling, 2019). The discovery implies that mentorship programs must not be based on merely evaluating the results but also challenging the moral ethics of the mentors. Opposing gratitude takes a deeper turn into the concept of cognitive dissonance in mentorship. Although past literature shows that mentees remain silent whenever exploitation is brought up (Nye et al., 2021), the present study reveals that gratitude increases dissonance and generates guilt and self-blame. Such ambivalence is why mentees might justify unethical conduct, making it harder to detect. Lastly, the cue signals dual image reveals how organizational rituals, namely, ethics talk, selective generosity, and public endorsement, reinforce manipulative practices. Recent research invoking organizational hypocrisy says that when symbolic ethics conflict with behaviors, workers shun cynical effects that harm engagement and output (Ehrich et al., 2018; Newman et al., 2017). The current results confirm this argument and further demonstrate how hypocrisy is played out between people through mentorship. The resulting discussion stresses that mentorship cannot be measured in specific and independent outcomes like career progression. Instead, it should be reviewed as a place of moral ambiguity, which is decided by force, reciprocation, and organizational encouragement. This research offers support that mentorship is not an ethical notion on its own but instead on the integrity of exemplars and the institutional systems that support them.

The findings across the five themes extend the explanatory power of major leadership theories. First, Jung's Shadow Theory (1990) is advanced by demonstrating how ethical façades function not merely as hidden impulses but as *institutionally rewarded performances*. The dual image of mentors ("wolf in a scholar's robe") shows that organizational recognition can amplify the shadow self, making ethical camouflage sustainable over time. Second, the Toxic Triangle (Padilla, Hogan, & Kaiser, 2007) is deepened by highlighting how "debts that never end" and "trust fractures" reveal not only destructive leaders but also the complicit environments that normalize manipulation. This

 doi:10.20944/preprints202602.0850.v1

study shows that when mentorship is endorsed as an ethical practice, mentees feel pressured into silence, extending the model beyond susceptible followers to *structurally silenced followers*. Third, Trait Activation Theory (Tett & Burnett, 2003) is refined by showing that cues such as institutional rituals, ethics talk, and selective recognition activate Machiavellian tendencies in leaders. In the case of "signals of dual image," it is not only situational demands but also organizational *branding strategies* that trigger dark trait expression, reinforcing manipulation behind care. Collectively, these contributions show that mentorship is not simply a dyadic exchange but a systemic site where dark traits are activated, shadows are institutionalized, and toxic environments are reproduced.

This study provides practical guidance for institutions seeking to safeguard mentorship from manipulation and ensure that ethical support is genuine rather than performative. By revealing how trust fractures and "debts that never end" undermine psychological safety, the findings urge universities to adopt transparent mentorship policies, routine feedback mechanisms, and explicit safeguards against exploitative reciprocity.

The results also resonate with broader concerns in higher education identified by Alowais and Suliman (2025), who demonstrated how dark traits in leadership can cascade downward and erode institutional trust. Similarly, Alowais (2024) argued that ethical philosophies must anchor organizational strategies if sustainability and fairness are to be realized. Together, these studies reinforce the urgency of developing accountability structures in academic mentorship, ranging from formal recognition for authentic mentorship to risk reviews that decouple institutional branding from ethical gatekeeping.

Practically, this means that higher education institutions must move beyond slogans of "supportive mentorship" and invest in systems that ensure ethical guidance is consistent, transparent, and resistant to manipulation. Doing so not only protects mentees but also strengthens organizational integrity, aligning with national goals for academic Emiratization and sustainable development.

6. Conclusion

That paper examined the darker aspects of mentoring, including how ethical appearances and norms of reciprocity can be gamed to perpetuate power. Five motifs have been identified through qualitative interviews of mid-level employe*es: the wolf in a scholar's robe, debts that cannot be resolved, fractures of trust, ambivalence* of g*ratitude, and signs of* dual image. These themes reveal that mentorship is not only a developmental instrument but also an instrument used to wield power and influence impression management.

The results have three main contributions. First, they contribute to the body of research in ethics by demonstrating how ethical branding can cover up abuse. Although ethical leadership is typically hailed as a source of trust and fairness, this research establishes that this form of branding can be turned into a weapon. Second, the findings reiterate the role of reciprocity norms in developing long-term commitments. The feeling of indebtedness perceived by mentees shows that exploitative mentorship has become common and is seldom met with resistance. Third, the paper draws attention to the organizational-level effects, especially the loss of trust and credibility in cases where manipulative activities are exposed.

In practical terms, the study recommends that organizations go beyond empty ethics programs to protect the integrity of mentorship. These involve defining confidential reporting processes, critically reviewing mentorship programs, and completing the cultural conditions that facilitate duplicity. To mentees, the results confirm that gratitude and disillusionment may not always be conflicting, and that expressing an imagined form of self-preservation via challenges of exploitative mentorship is no longer a place to be seen as betrayal. Finally, mentorship is still good but not morally neutral. The dangers of manipulation, silence, and loss of trust will continue to go hand in hand with their good sides unless the organizations realize their hidden dynamics and solve them.

References

 doi:10.20944/preprints202602.0850.v1

Allen, T. D., & O'Brien, K. E. (2019). Formal mentoring programs and organizational attraction. *Human Resource Development Quarterly*, *17*(1), 43-58.

Allen, T. D., Eby, L. T., O'Brien, K. E., & Lentz, E. (2017). The state of mentoring research: A qualitative review of current research methods and future research implications. *Journal of vocational behavior*, *73*(3), 343-357.

Alowais, A. A. (2024). The ethical dilemma of profit: Evaluating the Triple Bottom Line and the role of moral conscience in business decisions. *Journal of Ecohumanism*, *3*(8), 7143–7150.

Alowais, A. A., & Suliman, A. (2025). The influence of leader Dark Triad on employee Dark Triad in higher education institutions. *TPM – Testing, Psychometrics, Methodology in Applied Psychology*, *32*(S4), 523–541. Cises srl.

Baker, H. K., Pandey, N., Kumar, S., & Haldar, A. (2020). A bibliometric analysis of board diversity: Current status, development, and future research directions. *Journal of business research*, *108*, 232-246.

Balamurugan, G., & Divyabharathi, S. (2021). Emotional Dissonance and Job Satisfaction–A Review of Empirical studies. *International Journal of Scientific Development and Research*, *6*(4), 126-130.

Belschak, F. D., Den Hartog, D. N., & De Hoogh, A. H. (2018). Angels and demons: The effect of ethical leadership on Machiavellian employees' work behaviors. *Frontiers in psychology*, *9*, 1082.

Belschak, F. D., Den Hartog, D. N., & De Hoogh, A. H. (2018). Angels and demons: The effect of ethical leadership on Machiavellian employees' work behaviors. *Frontiers in psychology*, *9*, 1082.

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: a tool to enhance trustworthiness or merely a nod to validation?. *Qualitative health research*, *26*(13), 1802-1811.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

British Psychological Society. (2021). *Code of Ethics and Conduct*.

Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *The leadership quarterly*, *17*(6), 595-616. sciencedirect.com/science/article/pii/S104898430600110X

Cai, H., Wang, L., & Jin, X. (2024). Leader's Machiavellianism and employees' counterproductive work behavior: Testing a moderated mediation model. *Frontiers in Psychology*, *14*, 1283509.

Christie, R., & Geis, F. L. (2013). *Studies in machiavellianism*. Academic Press.

Collison, K. L., Vize, C. E., Miller, J. D., & Lynam, D. R. (2018). Development and preliminary validation of a five factor model measure of Machiavellianism. *Psychological assessment*, *30*(10), 1401.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Cropanzano, R., Anthony, E. L., Daniels, S. R., & Hall, A. V. (2017). Social exchange theory: A critical review with theoretical remedies. *Academy of management annals*, *11*(1), 479-516.

Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., & DuBois, D. L. (2018). Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of vocational behavior*, *72*(2), 254-267.

Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., & DuBois, D. L. (2013). Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of vocational behavior*, *72*(2), 254-267.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative science quarterly*, *44*(2), 350-383.

Ehrich, L. C., Kimber, M., Millwater, J., & Cranston, N. (2018). Ethical dilemmas: A model to understand teacher practice. *Teachers and Teaching: theory and practice*, *17*(2), 173-185.

Ensher, E. A., & Murphy, S. E. (2005). *Power mentoring: How successful mentors and protégés get the most out of their relationships*. John Wiley & Sons.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics*, *5*(1), 1-4.

Fernandez, L. (2008). *Policing dissent: Social control and the anti-globalization movement*. Rutgers University Press.

Festiger, L. (1957). A theory of cognitive dissonance. *Row, Peterson, New York*.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Furnham, A., & Taylor, J. (2004). *The dark side of behaviour at work*. Palgrave Macmillan UK.

 doi:10.20944/preprints202602.0850.v1

Goffman, E. (2023). The presentation of self in everyday life. In *Social theory re-wired* (pp. 450-459). Routledge.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American sociological review*, 161-178.

Greenbaum, R. L., Hill, A., Mawritz, M. B., & Quade, M. J. (2017). Employee Machiavellianism to unethical behavior: The role of abusive supervision as a trait activator. *Journal of Management*, *43*(2), 585-609.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59–82.

Humphrey, R. H., Ashforth, B. E., & Diefendorff, J. M. (2015). The bright side of emotional labor. *Journal of organizational behavior*, *36*(6), 749-769.

Jaiswal, P., & Bhal, K. T. (2014). Behavioural flexibility: The use of upward impression management tactics by subordinates for good performance rating from leader and impact of organizational & leader's Machiavellianism. *Global Journal of Flexible Systems Management*, *15*(4), 313-326.

Jung, C. G. (1959/1990). *The archetypes and the collective unconscious* (R. F. C. Hull, Trans., 2nd ed.). Princeton University Press. (Original work published 1959)

Jung, C. G. (1990). *The archetypes and the collective unconscious* (R. F. C. Hull, Trans., 2nd ed., Vol. 9, Part 1). Princeton University Press. (Original work published 1959)

Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing*, *72*(12), 2954-2965.

Lichterman, P. (2017). Interpretive reflexivity in ethnography. *Ethnography*, *18*(1), 35-45.

Loi, R., Chan, K. W., & Lam, L. W. (2014). Leader–member exchange, organizational identification, and job satisfaction: A social identity perspective. *Journal of Occupational and Organizational psychology*, *87*(1), 42-61.

Macassa, G., McGrath, C., Tomaselli, G., & Buttigieg, S. C. (2021). Corporate social responsibility and internal stakeholders' health and well-being in Europe: a systematic descriptive review. *Health Promotion International*, *36*(3), 866-883.

Mann, K. V., & Gaufberg, E. (2016). Role modeling and mentoring in the formation of professional identity. *Teaching medical professionalism. Supporting the development of a professional identity. Cambridge: Cambridge University Press. pp*, 84-96.

Mortelmans, D. (2019). Analyzing qualitative data using NVivo. In *The Palgrave handbook of methods for media policy research* (pp. 435-450). Cham: Springer International Publishing.

Neves, P., & Story, J. (2015). Ethical leadership and reputation: Combined indirect effects on organizational deviance. *Journal of Business Ethics*, *127*(1), 165-176.

Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. *Human resource management review*, *27*(3), 521-535.

Ng, T. W., & Feldman, D. C. (2015). Ethical leadership: meta-analytic evidence of criterion-related and incremental validity. *Journal of applied psychology*, *100*(3), 948.

Nye, C. D., Prasad, J., & Rounds, J. (2021). The effects of vocational interests on motivation, satisfaction, and academic performance: Test of a mediated model. *Journal of Vocational Behavior*, *127*, 103583.

Padilla, A., Hogan, R., & Kaiser, R. B. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly*, *18*(3), 176–194.

Padilla, A., Hogan, R., & Kaiser, R. B. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly*, *18*(3), 176–194.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.

Pizzolato, D., & Dierickx, K. (2023). The mentor's role in fostering research integrity standards among new generations of researchers: A review of empirical studies. *Science and Engineering Ethics*, *29*(3), 19.

Scandura, T. A. (1998). Dysfunctional mentoring relationships and outcomes. *Journal of management*, *24*(3), 449-467.

Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The leadership quarterly*, *24*(1), 138-158.

Shore, W. J., Toyokawa, T., & Anderson, D. D. (2008). Context-specific effects on reciprocity in mentoring relationships: ethical implications. *Mentoring & Tutoring: Partnership in Learning*, *16*(1), 17-29.

 doi:10.20944/preprints202602.0850.v1

Steadman, J. R. (2017). *The effects of over-indebtedness on interpersonal relationships* (Doctoral dissertation, Memorial University of Newfoundland).

Stocks, C. A. (2022). *The 'Imperative'of Informal Mentoring to Subvert Gender Role Incongruence Among Women in Higher Education Leadership: A Qualitative Study* (Doctoral dissertation, American College of Education).

Swargiary, K. (2024). *Techniques for manipulation and mind control*. ERA, US. books.google.com/books?hl=en&lr=&id=_G0pEQAAQBAJ&oi=fnd&pg=PA5&dq=At+least+formally,+mentorship+is+consistent+with+organizational+ethics+and+CSR%3B+at+the+same+time,+in+practice,+this+relationship+may+become+the+vehicle+of+control+and+self-serving+manipulation+&ots=cC6BmhePv1&sig=rmGLiK_EDfDKLvKir1aj8QqYwEc

Tepper, B. J. (2007). Abusive supervision in work organizations: Review, synthesis, and research agenda. *Journal of management*, *33*(3), 261-289.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*(3), 500–517.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*(3), 500–517. https://doi.org/10.1037/0021-9010.88.3.500

Tourish, D. (2020). The triumph of nonsense in management studies. *Academy of Management Learning & Education*, *19*(1), 99-109.

Trevino, L. K., & Brown, M. E. (2004). Managing to be ethical: Debunking five business ethics myths. *Academy of management perspectives*, *18*(2), 69-81.

Walter, C. E. (2025). *You and Your Team's Effectiveness: Understanding the Impact of Machiavellianism in Supervisors and Team Members on Team Effectiveness* (Master's thesis, Azusa Pacific University).