CAMBRIDGE
UNIVERSITY PRESS

FOCAL ARTICLE

An ethical leadership assessment center pilot: Assessing and developing moral person and moral manager dimensions

Alexis d'Amato[1], Vignesh Murugavel[1], Kelsey Medeiros[1], and Logan L. Watts[2]

[1]University of Nebraska at Omaha, Omaha, NE, USA and [2]Psychology, University of Texas Arlington, Arlington, TX, USA
Corresponding author: Kelsey Medeiros; Email: kmedeiros@unomaha.edu

Abstract
Ethical leaders are those who exemplify moral behavior personally, as well as those who facilitate follower ethical behavior. Although recent attention has been given to the ethical leadership construct, there remains a lack of innovation regarding the assessment and development of ethical leaders in organizations. To address these issues, a pilot study was conducted to examine the convergent validity of an ethical leadership assessment center, as well as the efficacy of using assessment center feedback to foster ethical leadership. Assessees completed a battery of pre tests, a virtual business simulation with a novel exercise, and a set of post tests. Half of the assessees were randomly assigned to a feedback condition, whereas the other half did not receive feedback until after the post tests were completed. Due to low statistical power, quantitative results were inconclusive. Nevertheless, qualitative insights were gained that point to implications for validating assessment center methodologies when assessing and developing ethical leadership.

Keywords: Assessment center; ethical leadership; moral manager; moral person

Introduction

The 2018 indictments of multiple Volkswagen executives (Glinton & Gotbaum, 2018), and the arrest of Audi CEO (Romo, 2018) in connection with the Volkswagen emission scandal serve as reminders of the importance of understanding how to identify and develop ethical leaders. A more recent example includes Toyota Motor failing to report defects that interfere with cars-controlled tailpipe emissions, thus violating public health and the environment (Tabuchi, 2021). Knowing this violation was occurring, Toyota managers and leadership ignored the noncompliance in an effort to sell millions of vehicles. These examples serve to highlight that, despite increasing attention, ethical dilemmas and scandals continue to plague organizations. Recognizing that these organizations have high standards, standards stemming from organizational leadership, the employees within the organizations behaved unethically - failing reporting protocols and ultimately harming public health and safety. Certainly, unethical behavior is not isolated to the auto industry. Further examples can be seen in tech industry with Elizabeth Holmes' deceitful practices in the development of new healthcare tech (BBC, 2022), the UK government with Boris Johnson's COVID-19 rule violation (CNN, 2022), and Balenciaga's 2022 advertising campaign featuring children in bondage (NYTimes, 2022). In addition to these examples, research on ethical behavior in the workplace emphasizes the important role that leaders play in fostering an ethical environment (e.g., Brown & Treviño, 2006). As such, to foster more ethical workplaces, we must also understand how to foster ethical leadership.

The goal of this article is to examine a practical intervention focused on the development of ethical persons and managers using a virtual ethical leadership assessment center.

Brown et al. (2005) defines ethical leadership (EL) as the demonstration and promotion of ethical conduct. Specifically, the demonstration of normative behaviors through personal action and interpersonal relationships and the promotion of such conduct via communication and reinforcement. Normative behaviors include conduct such as honesty, trustworthiness, fairness, and care. The definition is purposefully vague, as it is context dependent. The other component of ethical conduct is promoting ethical behaviors via three elements. The first element is communication, in that leaders draw attention to and explicitly speak about ethics in the social environment while providing their followers with voice (Bass & Steidlmeier, 1999). The second element is reinforcement, in which leaders set standards and reward and discipline followers (Gini, 1998; Treviño et al., 2003). Finally, the third element of ethical conduct is related to decision making, in which ethical leaders consider the consequences of their decisions and make fair choices that can be observed by others (Bass & Avolio, 2000; Howell & Avolio, 1992).

The inclusion of demonstration *and* promotion is key, implying that ethical leaders are moral persons, acting ethically themselves, as well as moral managers, encouraging their followers to engage in ethical behavior (Brown & Treviño, 2006; Treviño et al., 2000). Although this is a burgeoning area of theoretical and empirical research, the lack of innovation in measurement hinders the growth of the field and limits the application of findings to organizational settings. Further, there is scant research on whether EL can be developed (Brown & Treviño, 2006). The present effort aims to address these issues by employing a longitudinal experimental design to assess the efficacy of receiving feedback based on an EL diagnostic assessment center (EL-DAC) for assessing and developing EL with respect to both moral person and moral manager dimensions.

Ethics assessment

Since its rise to popularity in the 1980s, the predominant means for assessing ethics in the workplace has been, and remains, integrity testing (Sackett & Wanek, 1996). Integrity testing focuses on traits relevant to EL by assessing self-reports of overt behaviors, attitudes, or personality traits, which predict counterproductive workplace behaviors and unethical behavior (Sackett et al., 1989; Ones & Viswesvaran, 1998). Testing initially began with polygraphs, but the focus later turned toward oral and written integrity tests (Berry et al., 2007; Sackett & Wanek, 1996), which displayed moderate predictive validity coefficients for managerial performance ($r = .41$; Schnmidt et al., 1992). More recently, there have been efforts to assess integrity using situational judgment tests (SJTs; Becker, 2005; de Meijer et al., 2010). Although validity evidence has been provided, integrity tests primarily focus on the traits associated with the moral person, and thus, fail to fully capture the EL construct.

Other efforts have been made to examine the decision making dimension of EL. For instance, Mumford et al. (2006) developed a series of ethical decision making SJTs in the sciences. These measures focus on meta-cognitive strategies (e.g., anticipating consequences) associated with ethical judgments. More centrally, Loviscky et al. (2007) developed a measure of managerial moral judgment (MMJ). The MMJ was modeled after the Defining Issues Test which is grounded in Kohlberg's (1976) theory of moral development. The authors provide evidence bearing on the construct validity of the measure. More recently, Watts et al. (2020) developed a measure of cognitive biases, providing predictive validity evidence for their role in managerial ethical decision making. These measures primarily focus on the moral person, opposed to a moral manager.

There have also been efforts to expand this measurement to include dimensions of the moral manger. For example, Kalshoven et al. (2011) aimed to assess both the moral person and moral manager components in their EL at Work Questionnaire (ELW). The ELW asked subordinates to respond to a series of questions regarding perceptions of a leader's integrity, ethical guidance, and

fairness, among others. More recently, Yukl et al. (2013) developed the EL Questionnaire (ELQ). Similar to the ELW, the ELQ consists of 15 items such as, "My boss opposed the use of unethical practices to increase performance." Taken together, these measures are effective in assessing one's ethicality and morality as a person or manager, but fail to illuminate how organizations can develop ethical thinking and action among their current workforce.

Although these measures mark progress, measures of EL that can be used for selection and development purposes remain sparse. Some have argued that other, more innovative methods for assessing ethics may not be feasible. For instance, Loviscky et al. (2007) argued that the measurement of moral judgement vis-à-vis assessment centers was impractical due to findings that scores may represent exercise ratings rather than traits (e.g., Guion, 1998; Sackett & Dreher, 1982). However, Loviscky et al. (2007) focused on moral judgment rather than behaviors and decision making strategies associated with EL. If these dimensions can be approximated with SJTs, it stands to reason that an assessee may similarly demonstrate related behaviors in ethically charged assessment center simulations. In another critique of assessment centers, Loviscky et al., argued that assessment centers are costly. Although we do not disagree with this point, assessment centers may be worth the high cost when organizations are interested in assessing and developing high-level talent (Thorton & Rupp, 2006.). The high costs, financial and otherwise, associated with unethical behavior may very well warrant investment in procedures that more finely pinpoint leader's strengths and weaknesses with regard to ethical behavior and that may open the door for developmental interventions that prevent costly scandals.

Expanding EL measurement to assessment centers offers an additional advantage. Although ethical behavior has traditionally been studied from the individual, or "bad apple" perspective, research suggests that the environment, or the "barrel," also plays a key role in how leaders behave (e.g., Kish-Gephart et al., 2010). An assessment center allows for the manipulation of environmental characteristics that approximate the culture and norms of a specific organization and may therefore provide a better estimation of how an assessee may behave under the unique circumstances of one organization or role. Despite the potential benefits, no research has explored if EL can, in fact, be assessed using an assessment center. Based on the preceding discussion, we propose the following hypothesis:

> *Hypothesis 1: Measurement of EL in an assessment center will demonstrate convergent validity with traditional measures of ethics-related individual differences, EL behaviors, decision making, and integrity.*

Moral person and moral manager dimensions

The moral person dimension consists of three pillars: traits, behaviors, and decision making (Brown et al., 2005). Moving beyond traits, the literature suggests several processes associated with ethical behavior and decision making. For instance, Rest (1986) argued that ethical behavior results from recognizing an issue, making a judgment, and establishing intentions to act. This model has been generally accepted and incorporated with other models and findings (e.g., Jones, 1991) to provide a framework for understanding how ethical decisions are made and actions are taken. Although foundational, this model leaves to question, what decision making processes and strategies result in ethical judgments? Several processes have been proposed including considering actions and implementation planning (Janis & Mann, 1977), information processing and evaluation (Bommer et al., 1987), and perspective taking and information seeking (Pizarro & Bloom, 2003).

The moral manager dimension includes three pillars: communicating ethical standards and values to followers, rewarding ethical behavior and punishing unethical behavior, and role-modeling by making sound ethical decisions (Brown et al., 2005; Treviño et al., 2003). Thus, the moral manager dimension focuses on how leaders encourage ethical decision making and

behavior in their followers. In this instance, the third pillar, role-modeling, can be subsumed into the moral person category as it focuses on the leader's own ethical decision making. The other two dimensions, communication and reward/punishment, may both be directly observed in an assessment center context. Although there are measures of EL available, they are based on subordinate ratings, which may be problematic because subordinates may have limited opportunities to observe supervisor unethical conduct (Wexley & Youtz, 1985). Thus, an assessment center may allow for more directly observable accounts of EL. Communication of ethical values and standards and leader discipline regarding ethical or unethical behavior are the final two dimensions assessed.

Although there have been several studies examining predictors (e.g., Mayer et al., 2012), outcomes (e.g., Toor & Ofori, 2009), moderators (e.g., Kacmar et al., 2010), and mediators (e.g., Walumbwa & Schaubroeck, 2009) associated with EL, there has been little work examining how to effectively develop ethical leaders. Specifically, there is limited research exploring how to develop moral manager skills. This is a critical need as developing manager traits, such as integrity and honesty, is not feasible (Kirkpatrick & Locke, 1991). Thus, developing ethical leaders through altering traits is not a viable option. A more effective path forward may be through the decision making and skill-based training efforts, which have shown some efficacy in the sciences and business (e.g., Medeiros et al., 2017; Watts et al., 2017). An assessment center may be particularly beneficial for developing ethical leaders as it may provide rich sources of behavioral information, makes use of relevant cases, and allows for feedback – key components of successful ethics training programs (Medeiros et al., 2017; Watts et al., 2017). Specifically, in a training context, feedback allows trainees to know what they are doing appropriately and should continue, and what behaviors or decisions need to be changed or improved upon (Goldstein & Ford, 2002). Providing feedback to assessees after an assessment center may work similarly, providing them with actionable information regarding their EL. This leads to our second hypothesis:

Hypotheses 2: Providing assessees with feedback on EL-DAC performance can improve scores on subsequent EL assessments.

In addition to convergent and predictive validity, research has also demonstrated the importance of face validity, or the assessee's perception of job relatedness regarding the assessment center procedures (Kluger & Rothstein, 1993; Smither et al., 1993). Put simply, face validity in this context refers to the extent to which assessees perceive the content of the assessment center as relevant to developing their EL skills. A commonly cited strength of assessment centers is their face validity, as it is an important dimension in testing attitudes and reactions (Kluger & Rothstein, 1993; Macan et al., 1994).

To our knowledge, no research has explored the development of EL in assessment centers or the use of online mediums to conduct assessment centers, thus highlighting the need to consider assessee reactions in this particular context as a marker of face validity. The examination of face validity is especially critical when studying a new application of assessment centers, as the underlying motive of assessment center research is to produce a practical product or procedure. Additionally, it is important to account for assessee reactions in an assessment center, as perceptions can affect subsequent performance in future assessments.

Reactions in favor of the EL-DAC may be expected due to the strong face validity of assessment centers in general; however, it is important to note that the EL-DAC has two unique features that set it apart from traditional assessment centers. The first is rather obvious: the construct of interest, EL. Traditional assessment centers seek to measure complex multidimensional competencies that, together, compose performance in a specific job. Although ethical leadership is indeed a complex multidimensional component of some jobs, the construct is especially difficult to define due to its context dependence. If the context of scenarios in the EL-DAC is unclear to assessees, assessees may not be able to evaluate and track their own efficacy. The right way of

responding may not be apparent; therefore, the assessee may feel the EL-DAC lacks objectivity. As such, we explore the following research question:

Research Question: To what extent do assessees positively or negatively respond to the EL-DAC?

Method

Participant info

Twenty-four assesses were recruited over a two-year time period. Assessees included undergraduate and graduate business students at a Midwestern university. On average, assessees had seven years of work experience. The sample consisted of eight men and nine women, with seven assessees selecting not to report their gender. Each assessee was offered $75 for their participation. Ethics board approval was obtained.

Ethical dimensions explication

We used an iterative approach to the explication of our behavioral dimensions. First, the primary investigators identified sub-dimensions of Treviño et al.'s (2000) moral manager and moral person dimensions. Each sub-dimension was defined and behavioral markers were identified. Although Treviño et al. (2000) provide clear theoretical definitions of the moral person and moral manager dimensions, as well as sub-dimensions, these definitions were not always well-suited for developing behavioral markers observable in an assessment center context. This was especially true for the moral person dimension. For example, "personal morality" and "doing the right thing" are theoretically important, and often touted as critical to EL, but "morality" and "right" remain vague terms that are difficult to define and observe in an applied setting. Indeed, whole fields of philosophy are dedicated to arguing about what is right and what is moral. As such, without clear consensus on what is right, wrong, and moral, these sub-dimensions are difficult to define and contextualize within an assessment center context. Further, traits, such as trustworthiness, are difficult to assess given their subjective nature.

Additionally, some of the sub-dimensions explicated by Treviño et al. (2000) used socially loaded labels and definitions (e.g., "personal morality"). Given the potential sensitivity around some of the phrasing, the primary investigators viewed these labels as inappropriate for the developmental nature of the assessment center. It is possible that retaining the original language could limit an assessee's responsiveness to feedback. For example, research on biases demonstrates that most people view themselves as moral and having integrity (e.g., Dong et al., 2019). Receiving feedback that one received a low score on personal morality may unnecessarily limit one's willingness to hear that, or other, feedback. As such, we deemed it necessary to change some of the original language of the sub-dimensions to encourage assessee reflection.

The remaining dimensions were reviewed by three graduate students familiar with the EL literature and were revised for content, specificity, and clarity. The final list contained eight dimensions and can be viewed in Tables 1 and 2.

Exercise design

The primary investigators reviewed the assessment center literature for exercises relevant to EL to develop two exercises that were eventually used in the assessment center. The first exercise identified was a case analysis. Cases from "Arthur Anderson Case Studies in Business Ethics" hosted by Carnegie Melon's Tepper School of Business (Tepper School of Business, 2022) were reviewed and pilot tested by a small sample of undergraduates. During the pilot test, students read the cases and were asked to respond to three prompts: 1) describe the ethical dilemma in the scenario, 2) describe the steps you would take to resolve the dilemma and 3) describe why you would choose to take these steps. The pilot test was conducted to ensure that the selected cases

Table 1. Markers of a Moral Person (Brown et al., 2005)

Marker	Definition	Behavioral Anchors
Integrity	Caring about preserving one's personal or professional reputation	• Pursues short-term consequences relevant to one's self • Pursues long-term consequences relevant to one's self • Gives thought to how one's actions will impact his or her reputation
Self-Awareness	Recognizing one's personal values/motives, emotions, limitations, and factors that might bias decision making	• Reviews personal motives, values, emotions, and influences • Considers avenues to mitigate potential negative impact of personal motives, values, emotions, and influences • Seeks additional information including asking for help • Understands they may not have all of the information
Concern for Norms	Attending to rules, professional guidelines, policies, and norms	• Identifies relevant rules, guidelines, policies, and norms • Seeks out information regarding relevant professional rules, guidelines, policies, and norms • Demonstrates concern for making decisions that honor societal expectations
Concern for Others	Caring about others' opinions, feelings, dignity, and welfare; anticipating consequences for others	• Identifies key stakeholders and their concerns • Asks relevant parties for their perspective or input • Decides if actions will be viewed as fair by all parties • Evaluates downstream consequences for others

were eliciting responses relevant to EL and that were captured by the explicated dimensions. The first case involved a dilemma about whether to cover up or expose potentially damaging information about one's firm related to a worker injury. The second case dilemma involved a conflict between protecting the environment versus protecting an employer's reputation. Both cases asked assessees to take on a leadership role. Based on a review of the responses, two cases and the three questions used in the pilot test were retained for the EL-DAC. Although all participants responded to both cases over the course of the study, the order in which cases were presented was randomized.

The second exercise borrowed components of Treviño et al.'s (1985) inbox exercise focused on EL. We selected a combination of both relevant and irrelevant prompts to minimize the threat of demand characteristics. Using this exercise as a foundation, we then incorporated a role-play focused on sexual harassment claims. We chose to include this role play, as sexual harassment as it is an ethical issue per the APA Ethics Code (APA, 2022). Specifically, while working through the inbox exercise, assessees were interrupted two times by the fictional administrative assistant who suggested that there had been complaints about sexual harassment from one of the managers in the team. The role play was conducted by one of the assessors who took on the role of the administrative assistant. A copy of the revolving door script can be found in the Appendix.

The EL-DAC

Due to the COVID-19 pandemic, the assessment center was conducted online. Unique to other assessment centers, communication was one-on-one with the assessor and the assessee via email,

Table 2. Markers of a Moral Manager (Treviño et al., 2000)

Marker	Definition	Behavioral Anchors
Communication	Informing and reminding others of expected ethical standards to be displayed and followed; others may include subordinates, superiors, peers, customers, suppliers, stakeholders, etc.	• Formally communicates about the importance of organizational values and standards related to ethics through public displays (e.g., speeches, emails, memos, physical artifacts) • Informally discusses organizational values and standards related to ethics (e.g., team meetings) • Educates employees about resources pertaining to ethical issues • Frames ethical initiatives in a non-cynical way • Invests in ethics training or relevant initiatives aimed at educating followers
Reward/Discipline	Rewarding and encouraging ethical behavior; punishing and discouraging unethical behavior	• Provides positive support through recognition, praise, awards, or other tangible rewards to followers who behave ethically • Appropriately punishes employees who engage in unethical behavior • Reports unethical behavior • Tracks ethical behavior via performance evaluations • Praises employees for speaking up about difficult issues
Role Modeling	Setting a good example for followers by taking a stand against unethical behavior and following appropriate procedures	• Demonstrates ethical behavior in front of employees, supervisors, customers, etc. • Engages in ethical behavior/practices • Participates in ethic training programs or initiatives
Giving Voice	Encouraging opportunities for others to express concerns and ideas pertaining to ethical issues	• Provides formal and informal opportunities for others to express issues and concerns (e.g., meetings, emails, anonymous suggestions) • Checks up on and takes action on employees' concerns • Encourages employees to speak up regarding their concerns • Establishes psychologically safe environment/relationship with subordinates • Is available to hear employee concerns

Qualtrics surveys, and Zoom. One week prior to the assessment center, assessees received an email with expectations for their participation, as well as a link to complete the IRB consent form and pre-test. The pre-test included several measures of ethicality including Becker's (2005) Integrity Scale (α = .90–.95) and Watts et al.'s (2020) biased attitudes scale (BiAS; α = .73–.85), as well as measures of Machiavellianism (Christie & Geis, 1970), Narcissism (Crowe et al., 2016), and Moral Identity (Aquino & Reed, 2002). Assessees were also asked to send a link to an EL survey including the Ethical Leadership at Work Questionnaire (Kalshoven et al., 2011; α = .81) and the Ethical Leadership Questionnaire (Yukl et al., 2013; α = .74) to 3–5 members of their professional team.

When assessees logged in for their assessment center using Zoom, they were greeted by one assessor. Assessees first completed one of the two selected case studies with no time limit. They were asked to complete three open-ended questions in response to each case study. Next, assessees began their revolving door exercise using the Qualtrics survey tool. Each assessee was allotted 45 minutes to complete the exercise. Again, they were asked to provide open-ended responses to

the exercise prompts. The assessee was interrupted twice at the 10- and 20-minute time mark. After completing the revolving door exercise, the assessee was asked to complete an online interview survey which asked questions regarding the EL- specific components of the exercise.

Two weeks after completing the assessment center, assessees were sent a post test survey. The post test included all the original measures included in the pre-test survey. Additionally, assessees were once again asked to send the EL survey to 3–5 team members.

Assessment center scoring and feedback

Exercises were scored independently by three organizational psychology graduate students using EL dimensions (Treviño et al., 2000). Raters were trained using frame-of-reference training on samples of responses until sufficient rater agreement was observed. The raters were retrained until there was no more than 1.0 variance in agreement. Thus, using a 5-point rating scale, the raters evidenced an average interrater agreement coefficient (*rwg*) above .90. In instances of substantial disagreement among raters, additional training was provided in the form of calibration meetings. Raters were provided participant responses shortly following participant completion of exercises and would individually rate the EL dimensions for each exercise (i.e., cases, inbox, role play). The dimensions used were the moral person and moral manager behavioral markers outlined in above sections (e.g., integrity, self-awareness, role-modeling; Brown et al., 2005; Treviño et al., 2000).

Rater scores were provided within 1–3 weeks of completion. Aggregated (i.e., averaged) scores from these raters and exercises were used to populate the EL-DAC feedback form provided to assessees. The feedback form included the assessee's score on each dimension, a rating of low, medium, or high for each score, and an example of how this dimension was (or was not) demonstrated during the exercises. For example, one assessee scored a high score of Concern for Others for their response *"this is obviously very concerning becaye of the safety of all the employees and individuals of all parties involved"* and another scored high on Giving Voice, as they stated *"I would be willing to listen to the voices of my team and hopefully bring some wisdom to the tough decision"*. Whereas other assessees scored low on Reward and Discipline by stating *"I appreciate you for reaching out"* and Concern for Norms by defaulting decision-making power and responding, *"I will inform upper management of the problem."*

Conditions

To test the efficacy of assessment center feedback at developing EL, assessees were randomly assigned to one of two conditions. In condition 1 (i.e., feedback), assessees received their post-test within 1 week of completing the assessment center and prior to completing the post survey. In condition 2 (i.e., no feedback), assessees received their post test 3-4 weeks after their assessment center and after completing the post survey. This design allowed for the comparison of post-test scores between groups who received feedback and those who did not receive feedback. However, all participants received feedback within 2 weeks of completing the assessment center to ensure fairness across assessees.

Results

EL-DAC convergent validity and feedback efficacy

Hypothesis 1: Convergent validity

Correlations were examined to assess convergent validity. To better observe the effectiveness of the revolving door exercise, the inbox, and role-play components were examined separately. As such, three correlation matrices are presented in Tables 3 (case analysis), Table 4 (inbox), and Table 5 (role play). Given the low statistical power, none of the relationships were statistically significant. Nevertheless, patterns in the direction of correlation coefficients (positive vs. negative)

Table 3. Correlations for case analysis dimensions

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. BiAS (Pre)	1.00														
2. BiAS (Post)	.26	1.00													
3. Integrity (Pre)	.36	.63	1.00												
4. Integrity (Post)	−.37	.05	−.08	1.00											
5. Machiavelli	.09	−.22	.04	.24	1.00										
6. Moral Identity	.13	−.56	.06	−.29	−.12	1.00									
7. Narcissism	−.03	.18	.53	−.33	.02	.42	1.00								
Case Analysis EL Dimensions															
8. Concern for Reputation	.46	−.19	−.08	.21	.24	.02	−.52	1.00							
9. Self-Awareness	.29	−.18	−.02	.31	−.15	.44	−.12	.70	1.00						
10. Concern for Norms	−.58	−.50	−.17	.30	−.16	.35	−.04	.24	.43	1.00					
11. Concern for Others	.68	−.17	.05	.12	.38	.41	−.00	.50	.59	−.27	1.00				
12. Communication	.13	.28	.46	.18	−.43	.21	−.03	−.16	.11	−.02	.05	1.00			
13. Reward and Discipline	.09	.58	.87	−.00	−.31	.03	.32	−.20	−.06	.03	−.24	.73	1.00		
14. Role Modeling	.54	.24	.46	.28	.28	.21	.11	.08	.23	−.44	.73	.56	.29	1.00	
15. Voice	.30	.31	.54	.07	−.43	.42	.24	−.17	.28	.10	.29	.91	.67	.68	1.00

Table 4. Correlations for Inbox Dimensions

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. BiAS (Pre)	1.00														
2. BiAS (Post)	.26	1.00													
3. Integrity (Pre)	.36	.63	1.00												
4. Integrity (Post)	−.37	.05	−.08	1.00											
5. Machiavelli	.09	−.22	.04	.24	1.00										
6. Moral Identity	.13	−.56	.06	−.29	−.12	1.00									
7. Narcissism	−.03	.18	.53	−.33	.02	.42	1.00								
Inbox EL Dimensions															
8. Concern for Reputation	.25	−.11	.06	−.06	−.04	.44	.32	1.00							
9. Self-Awareness	.64	−.05	.05	.23	.39	.15	−.17	.63	1.00						
10. Concern for Norms	.53	.19	.29	−.08	−.01	.20	.20	.90	.74	1.00					
11. Concern for Others	.61	−.13	−.07	−.20	−.18	.49	.04	.81	.74	.79	1.00				
12. Communication	.59	−.06	.07	−.11	.31	.17	.03	.81	.86	.90	.76	1.00			
13. Reward and Discipline	.35	.04	−.08	−.00	−.66	.40	−.13	.25	.24	.22	.63	.03	1.00		
14. Role Modeling	.80	.35	.43	−.25	.18	.01	.15	.64	.74	.88	.66	.87	.07	1.00	
15. Voice	.57	.26	.16	−.02	−.24	.12	−.15	−.27	.17	−.12	.24	−.17	.70	.06	1.00

Table 5. Correlations for Role Play Dimensions

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. BiAS (Pre)	1.00														
2. BiAS (Post)	.29	1.00													
3. Integrity (Pre)	.46	.62	1.00												
4. Integrity (Post)	−.34	.01	−.22	1.00											
5. Machiavelli	.07	−.17	.21	.32	1.00										
6. Moral Identity	.13	−.57	.07	−.29	−.13	1.00									
7. Narcissism	−.05	.22	.70	−.31	−.02	.42	1.00								
Role Play EL Dimensions															
8. Concern for Reputation	−.43	−.06	−.66	.05	−.81	−.14	−.40	1.00							
9. Self-Awareness	.35	−.10	−.30	.30	.11	−.22	−.83	.14	1.00						
10. Concern for Norms	.48	.45	−.02	.27	−.11	−.50	−.66	.20	.81	1.00					
11. Concern for Others	−.10	−.01	−.09	.67	−.15	.14	−.23	.29	.49	.42	1.00				
12. Communication	−.55	−.50	−.51	−.04	−.60	.53	.15	.63	−.36	−.51	.18	1.00			
13. Reward & Discipline	.09	.39	.30	−.29	−.75	.32	.56	.31	−.51	−.14	.07	.42	1.00		
14. Role Modeling	.51	.19	.07	.13	−.48	.28	−.17	.27	.44	.58	.66	.11	.48	1.00	
15. Voice	.31	.16	.07	.10	−.60	.34	−.08	.40	.38	.48	.75	.25	.53	.94	1.00

provided some insight bearing on convergent validity. For example, correlations between moral identity and EL dimension scores for both the case analysis exercise (Table 3) and the inbox exercise (Table 4) trended in the positive direction for all eight EL dimensions. Similarly, depending on the exercise scored, Machiavellianism showed mostly negative correlational patterns with the EL dimensions. Finally, despite a few exceptions, the eight EL dimensions within each exercise tended to correlate in the positive direction with one another.

Hypothesis 2: Efficacy of Feedback
Hypothesis 2 focused on differences between those who received immediate feedback and those who received delayed feedback (after the post-test). Similar to the examination of Hypothesis 1, the shift in our data collection resulted in a sample that would likely be underpowered for a statistical mean difference test. Thus, it is not surprising that no statistically significant differences were observed between the feedback versus no feedback conditions for the integrity test or BiAS scores. Although most assessees had at least one team member successfully complete the pre-surveys of EL, only five assessees had at least one team member complete the post-surveys of EL. As such, no analyses could be conducted regarding others' perceptions of the assessee EL before and after the assessment center.

Qualitative reactions to revolving door exercise
The revolving door exercise consisted of both an inbox and a role-play. Although originally designed as a physical disruption to the assessee's inbox exercise, COVID-19 circumstances necessitated moving the role-play to an online chat context. To our knowledge, the chat function

has not been used in a formal assessment center in the literature. Reactions to the revolving door exercise are threefold: (a) the chat elicited responses relevant to the dimensions, (b) using the chat function does represent how many people communicate at work, and (c) assessees sometimes did not notice and respond to the chat potentially creating construct irrelevant variance.

On the positive side, when assessees did notice the chat messages, their responses were related to the dimensions at hand. For instance, several assessees noted following up with formal investigations into the remarks, whereas others specifically noted punishments. Others noted needing to follow up on company policy before responding further. Still, some pushed the issue aside and suggested that it may take a while to resolve due to "more pressing issues." The range of responses suggests that this exercise both tapped relevant dimensions and provided the expected variability across assessees.

Last, assessees were asked to share their experience with the revolving door exercise during the post exercise interview. One of the authors identified five themes in the 24 comments provided by assesses. The first two themes—positivity and challenge—emerged for the majority of assesses. Specifically, 75% of assessees commented that they had a positive reaction to the exercise (e.g., "I actually really enjoyed this activity"), and 87.5% commented about the challenging nature of the exercise. Constraints that were reported to increase the challenge of the exercise included limited time, lack of personal management experience, lack of organization-specific knowledge relevant to the scenario, and the difficulty of solving interpersonal and ethical issues. Nevertheless, many assessees (25%) noted the realism of experiencing such constraints in a management role—our third theme. For example, one assessee noted, "Since there was a time limit, it stimulates real life working condition to scramble and try to reply all emails in a short period of time." The fourth theme, noted by 20.8% assessees, was that the exercise stimulated self-reflection and learning (e.g., "I feel l learned about myself during this process and how I make consequential decision under pressure"). The fifth, and final, theme was ideas for improving the exercise (29.2% of assessees). For example, one assessee noted they could have completed the exercise more efficiently if they had the ability to "toggle between emails." A full list of assessee comments can be viewed in Table 6.

Discussion

Experiential and quantitative results from this effort reveal several challenges and opportunities for assessing and developing EL using the assessment center approach. Regarding dimensions, the primary investigators were able to develop eight dimensions with behavioral anchors (see Brown et al., 2005) observable during the assessment center. This is an important outcome as it was unclear whether observable behavioral markers of EL in this context could successfully be developed. However, the dimensions also presented several challenges. Specifically, the principle investigators and the graduate research assistants noted the difficulty in observing moral person dimensions (e.g., integrity, self-awareness, concern for others) during the exercises. In contrast, both groups noted that the moral manager dimensions were displayed more frequently. As such, although the moral person and moral manager dimensions make appropriate theoretical distinctions, the moral person dimension appears less appropriate for assessing observable behaviors. Finally, it is likely more useful to assessees to receive feedback on the moral manager dimensions compared to the moral person dimensions. From a practical perspective, it is more appropriate and useful to provide feedback such as "Communicate about ethics with your team more" rather than "Have more integrity."

Regarding assessee reactions to the revolving door exercise, it is noteworthy that chat functions such as Microsoft Teams, Slack, and other communication channels have become common methods of communication in the workplace. As such, the introduction of an interactive chat function adds a level of face validity to the assessment center. Although the present effort used the

Table 6. Assessee Reactions to Revolving Door Exercise

#	Assessee Comment
1	The inbox activity was very good and provides fast responses as well as the ability to prioritize. The hierarchy of importance is tougher to make because of the time crunch. Additionally, I felt that as a leader, taking one hour on Friday to read and respond to email on in the digital age we live in demonstrates that the employee may not be properly motivated to do the job well. Throwing in the chat while responding to the email and determining proper responses was good. It would be cool to add another piece where you can spend the hour placing calls as well. The dilemma of some of the ethical items were more challenging, because in real life these would require a phone call and not just email and urgent reactions.
2	I think it was a fun activity to be a manager and see the responsibility for a day. For me, it was easier since I often receive 30 emails a day I need to quickly reply to. Some were more challenging when it was time sensitive.
3	The inbox activity was a lot easier than I thought it would be, especially with the time constraints. I think for a lot people it may have caused stress, but with my jobs and involvements I knew how to prioritize and respond. I saved a couple things for Monday that required meetings, but I am very proud of how I did.
4	My overall impression was that it considered complex, nuanced workplace scenarios that involve logical and ethical decision making by leaders. The activities were realistic and adequately challenging. I feel I learned about myself during this process and how I make consequential decision under pressure. The most challenging aspect of the activity was quickly thinking through a plan on how to handle the situation. That is, how much information do you give to the recipient as you had Shirley plan steps behind the scenes, such as having her contact HR, etc. It was also difficult to find political correct ways to say someone's suggestions were totally inappropriate.
5	I thought it was good. It was hard to prioritize the messages due to the fact I don't actually work for the company. However, the tree diagram did help. Anything e-mail having to do with numbers were more challenging because I didn't know the scenario besides the fact that sales have been down
6	I thought it was good and challenging. There were a few emails I didn't have enough time to read and it did take a lot more time than I realized to be able to respond and think of the importance of each. Some emails needed more than just a quick response and those required more time as a result.
7	My overall impression was a great exercise to see how one reacts to stress. It had both complicated scenarios, and simple ones with you still having to check your chat function like you would at work. Yes some aspects were more challenging, the emails that were just stating that a report needs to be completed or to take a training course were simple. It was the complicated issues where I had to figure out who they were referring to, position, and how I should respond to them that were more challenging.
8	I really enjoyed this exercise. While I had to respond to emails in my current workplace, there were none quite like the ones in the exercise. It got me thinking and evaluating everything I was saying which made it more enjoyable. It presented hard decisions as well as realistic scenarios.
9	Overall, it went good but I thought it was kind of challenging because you had to position yourself to different management level and think like them with zero experience.
10	Overall, I thought it was a fun and interesting exercise. Many of the questions dealt with morality and whether or not to put profits first or what the "Right Choice" was. The most challenging aspect was the time of course and deciding what needed to be addressed immediately. Not knowing how many emails were left was also daunting.
11	I found this exercise to be a bit difficult due to not understanding the culture of the firm or business itself. In my personal job, I know when I'm speaking to other coworkers, we are informal but when speaking to clients I am more formal. So I was not sure which one was appropriate here. I found myself in the first email spending extra time trying to remember the name of the person I was pretending to be as well as the date I was writing these emails and even the name of my assistant.
12	I thought it was fun. I liked the quick-fire nature of having to get through my inbox in a short period of time. It was challenging just navigating in a non-outlook format. I feel I would have been more efficient if I could have toggled between emails like the native application allows.

(*Continued*)

Table 6. (*Continued*)

#	Assessee Comment
13	My overall impression of the inbox activity was that it was fun! I think I could have answered all of the emails in the 45 minute time frame if I were really Pat Sneed. I had to re-read some information to really grasp who the email/situation was about and then scroll backwards in the survey to the list of who is in the company at what job descriptions. Some of the situations presented were fairly easy to determine the solution. Others required multiple steps and personnel to complete. There were a few emails that I explained I was busy but would get back to them within the next few days. It was challenging to handle a vast array of situations from customer complaints to employee concerns and knowing which ones to take to HR and which ones to handle myself. The most important thing in reading and replying to emails is to assume positive regard in all cases.
14	It would be easier if I could preview all emails in a list to help determine which emails are more important to respond first (e.g. from higher ups, or title with issues). Since there was a time limit, it stimulates real life working condition to scramble and try to reply all emails in a short period of time.
15	I think it gave a very detailed insight into the challenges and responsibilities of senior management. Some aspects were very challenging and required a lot of critical thinking to ensure that the right decisions are being made for the employees and the company while keeping in mind the consequences everyone would have to face because of it.
16	There were a lot of ethical decisions to be made which made it more challenging. I tried to address all of those in the most detail which left less time for other issues. It seemed like Pat was going to have to have a lot of meetings with other people on the team to address all of the problems that were going on. I would not want to be in that position because of all the concerns. Overall, I felt this was a good activity to challenge my ability to think quickly and come up with solutions or at least a response saying I would come back to it later.
17	I thought the inbox activity was good, it really made me think of the kinds of hard decisions a leader has to make quickly. Some aspects were a lot more challenging than others. Some of the situations revolved around a really big decision that would affect a lot of company things. Making these decisions so far with no additional information was hard. It was also hard to make decisions when one decision is the right thing to do but will hurt the company and the other is the bad decision but doesnt hurt the company. Your instinct is to protect the company and your job and make these bad decisions. But then you need to consider your personal ethics and the ethics codes of the company and make the right choice even if it will result in consequences for the company.
18	It was good, however, I felt some items may have needed to be addressed by other departments. And some may have been tasks that simply wouldn't need addressed. If I was really on a time crunch, I may not send a note until Monday saying I was still planning to complete ethics training.
19	It kept me on my toes and elevated my stress level. Dealing with HR related issues were the most difficult in the given time. I felt like reading all the emails before responding would have helped me prioritze better.
20	The inbox activity provided an effective level of "real-world" applicability/relatability, particularly with the inclusion of real-time chats with "Shirley". The overall effect was more of a simulation than an assessment, and I had a favorable experience in that regard. Responding to certain emails requesting decisions on high-stake items such as department budgets, personnel issues, and hiring decisions under the time constraints was challenging. Normally, I would prefer to give these items significantly more time to consider, but I appreciate that, in the real world, this is not always a possibility.
21	Overall was a real good exercise to review and prioritize the tasks with the time constraint. Prioritizing the tasks based on the intensity and magnitude of the issue or task is very critical for Leaders and is one of the challenge we run into every day. As I worked through this activity, after the fact felt like I could have improvised on my responses. The challenging part was I wasn't able to review 3 of the tasks that were there. So I should have spent more time prioritizing the tasks, rather than addressing and responding to each of these.
22	I think it's one of the better leadership assessment exercises I have experienced so far. I think trying to remember the people in the organization and their roles is the most challenging aspect of the exercise. I think it would be beneficial to provide the organizational chart as a separate attachment so it can be viewed at any time. I think the time constraint makes it more realistic.

(*Continued*)

Table 6. (*Continued*)

#	Assessee Comment
23	I actually really enjoyed this activity. This was a little more challenging since I could not move files or forward / flag emails. I use my inbox as a TO DO inbox so it is usually cleared daily. I think it was hard to prioritize the emails. I think it is important to immediately address the HR-related issues and the customer and/or client complaints. I did try to think about the personal issues I was having to address in an hour so I did rush myself a little bit.
24	The inbox activity has too many emails to respond to. Each email had a lengthy text to read. Reading, accessing the scenario, and replying to all emails in 45 mins was not enough. Zoom chat never signaled me of an incoming text.

Zoom chat function, more realistic revolving door tasks should include more high-tech, face valid chat functions such as Teams and Slack. Future research might also consider how to better represent the modern work experience in an inbox exercise. Just as the in-basket moved to inbox, perhaps now is the time to consider a more dynamic exercise that better reflects the multiple different types of communication tools (e.g., email, chat, text) available to workers today.

It is important to note a major drawback to the exercise. Although participants were informed that some colleagues may use the chat function, some assessees did not notice the chat, or did not notice the chat until later in the exercise. Failing to notice the chat likely represents construct irrelevant variance as it is distinct from an assessee who reads the chat and chooses not to respond. In this instance, we were unable to parse out these two types of responses and as such, were unable to accurately capture the assessee's response to the text-based role play. Using practice chats, or other chat methods, may be fruitful in future efforts using this approach.

Additionally, variability in assessee responses to the revolving door exercise suggests that the exercise was effective for eliciting a range of responses. This is particularly true for the role-play component of the exercise. As such, we encourage future researchers to continue to expand the arsenal of available assessment center exercises that incorporate modern technology and better reflect the way people live and work. For example, COVID-19 made remote and hybrid working a mainstay for many. Creating remote (such as this) or hybrid assessment centers that allow assessees to use common platforms and technologies could create a more face valid experience. However, it is still important to bear in mind potential downsides of using technology in the assessment center exercise. Specifically, technology should not be used casually as it can create construct irrelevant variance. As such, technology and modern platforms should only be used to the extent that that technology is important on the job and required for performance.

In line with the point above, we note that our study design focuses primarily on work and leadership in a virtual context. Therefore, our observations may not generalize to assessing EL in assessment centers in all work contexts. Mesmer-Magnus & DeChurch (2009) and De Guinea et al. (2012) identified key differences in social interactions in virtual versus in-person work. Namely, openness behaviors are expected to be less frequent, communication and knowledge sharing are expected to decrease, and conflict is expected to increase in virtual work. Such social variables may affect how EL behaviors are enacted and detected as part of a virtual assessment center process. A related concern emerges when considering the predictive validity of EL assessment centers. Specifically, our approach was limited in examining others' ratings of assessee EL. Given that peers and coworkers may have limited opportunities to observe an assessee's unethical conduct, particularly in virtual work settings, coworker ratings of EL may be too contaminated or insufficient as a criterion for establishing the predictive validity of EL assessment centers.

The lack of empirical results regarding the efficacy of the assessment center feedback is disappointing. However, given the complexity of EL, developing this approach may be difficult and

require a longer period of time than may be observed in the present study. As such, future research should consider a multi-wave longitudinal approach that tracks the development of EL over time.

Despite these results, correlational relationships suggest some promising patterns. For example, the largely negative relationships between the assessment center dimensions and Machiavellianism and positive relationships with moral identity suggests that the dimensions may be appropriately assessing important ethics constructs. However, the results were inconsistent across constructs. As such, more research is needed to fully understand the potential of assessment centers for developing EL.

Of course, it is important to note that these correlations may not be generalizable. For example, it is plausible that different results may have been observed in a face-to-face setting. As methodological choices can introduce construct irrelevant variance, it is possible that additional factors related to our use of a digital medium may have impacted results. Although this approach was necessary due to the COVID-19 pandemic, and, in fact, presented a unique opportunity to conduct the assessment center in a virtual environment, future research is needed to understand and compare the differences between an online and a face-to-face EL assessment center.

Our research offers an initial examination of the potential for using the assessment center technique to develop EL. Our experience and initial results suggest that it is possible to use this technique to assess EL. Specifically, assessment centers may be particularly useful for evaluating the "moral manager" dimension of ethical leadership. Assessing the "moral person" component proved much more difficult and as such, may be better suited for other assessment techniques. This notion was supported by empirical evidence, which demonstrated some patterns, albeit inconsistently, of relationships across ethics-related variables and the dimensions assessed.

Although this work presents preliminary findings and opens the way for future research, some suggestions for practice can be discussed. Organizations may seek to quantify EL to preempt damaging scandals. This study demonstrates that some critical qualities that may seem abstract (e.g., EL), can indeed be measured. Specifically, assessment centers can capture the moral managing components of EL; however, the moral component of ethical leadership may be better evaluated though other assessments (e.g., survey response).

Our study also revealed that EL assessment center approaches can be adapted to incorporate interventions or pre/post tests. This point may be especially useful for designing EL development strategies. Moreover, our review of assessee feedback suggests that managers may even respond positively to live, high-fidelity evaluation of their ethical behavior. An example of the high-fidelity nature of our inbox/role-play exercise was the organizational chart provided to each assessee that illustrated their position in the organization relative to other characters in the simulation. This organizational chart helped to simulate the power differences between positions might occur in real-world work relationships—power differences each assessee had to navigate to successfully demonstrate their EL.

Although conclusions are limited due to a small sample, initial empirical results suggest that the assessment center technique may produce a backlash effect seen in other training contexts if attempting to develop attitudes. As such, future research is needed to understand the trajectory of development for EL using this and other techniques. As illustrated in the Toyota Motors example, ethical dilemmas continue to arise in modern organizations, as the integration of profit goals, public health, and environmental issues become salient. This research offers preliminary solutions to these relationships in that virtual EL assessment centers may contribute to the future of ethical leader development.

References

APA. Ethical principles of psychologists and code of conduct 2022. https://www.apa.org/ethics/code#:~:text=Sexual%20harassment%20is%20sexual%20solicitation,workplace%20or%20educational%20environment%2C%20and.

Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, **83**, 6, 1423–1440. DOI: 10.1037/0022-3514.83.6.1423.

Bass, B. M., & Avolio, B. J. (2000). *MLQ: Multifactor leadership questionnaire*. Palo Alto, CA: Mind Garden.

Bass, B. M., & Steidlmeier, P. (1999). Ethics, character, and authentic transformational leadership behavior. *The Leadership Quarterly*, **10**(2), 181–217. DOI: 10.1016/S1048-9843(99)00016-8.

BBC News, (2022). Theranos founder Elizabeth Holmes jailed for fraud. https://www.bbc.co.uk/news/world-us-canada-63685131.

Becker, T. E. (2005). Development and validation of a situational judgment test of employee integrity. *International Journal of Selection and Assessment*, **13**(3), 225–232. DOI: 10.1111/j.1468-2389.2005.00319.x.

Berry, C. M., Sackett, P. R., & Wiemann, S. (2007). A review of recent developments in integrity test research. *Personnel Psychology*, **60**, 271–301. DOI: 10.111/j.1744-6570.2007.00074.x.

Bommer, M., Gratto, C., Gravander, J., & Tuttle, M. (1987). A behavioral model of ethical and unethical decision making. *Journal of Business Ethics*, **6**, 265–280. DOI: 10.07/BF00382936.

Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *The Leadership quarterly*, **17**, 6, 595–616. DOI: 10.1016/j.leaqua.2006.10.004.

Brown, M. E., Treviño, L. K., & Harrison, D. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, **97**, 117–134. DOI: 10.1016/j.obhdp.2005.03.002.

Christie, R., & Geis, F. L. (1970). *Machiavellianism*. Academic Press.

Crowe, M., Carter, N. T., Campbell, K., & Miller, J. D. (2016). Validation of the Narcissistic Grandiosity Scale and creation of reduced item variants. *Psychological Assessment*, **28**(12), 1550–1560. DOI: 10.1037/pas0000281.

De Guinea, A. O., Webster, J., & Staples, D. S. (2012). A meta-analysis of the consequences of virtualness on team functioning. *Information & Management*, **49**(6), 301–308. DOI: 10.1016/j.im.2012.08.003.

de Meijer, L. A. L., Born, M. P., van Zielst, J., & van der Molen, H. T. (2010). Construct-driven development of a video-based situational judgment test for integrity: A study in a multiethnic police setting. *European Psychologist*, **15**, 229–236. DOI: 10.1027/1016-9040/a000027.

Dong, M., van Prooijen, J. W., & van Lange, P. A. (2019). Self-enhancement in moral hypocrisy: Moral superiority and moral identity are about better appearances. *PloS One*, **14**(7).

Gini, A. (1998). Work, identity, and self: How we are formed by the work we do. *Journal of Business Ethics*, **17**, 707–714. DOI: 10.1023/A:1017967009252.

Glinton, S., & Gotbaum, R. Former Volkswagen CEO indicted over emissions testing scandal (2018). https://www.npr.org/sections/thetwo-way/2018/05/04/608374639/former-volkswagen-ceo-indicted-over-emission-testing-scandal.

Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th edn). Wadsworth/Thomson Learning.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Lawrence Erlbaum Associates Publishers.

Howell, J. M., & Avolio, B. (1992). The ethics of charismatic leadership: Submission or liberation? *Academy of Management Perspectives*, **6**(2). DOI: 10.5465/ame.1992.4274395.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict choice and commitment*. The Free Press.

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue- contingent model. *Academy of Management Review*, **16**, 366–395. DOI: 10.5465/amr.1991.4278958.

Kacmar, K. M., Bachrach, D. G., Harris, K. J., & Zivnuska, S. (2010). Fostering good citizenship through ethical leadership: Exploring the moderating role of gender and organizational politics. *Journal of Applied Psychology*, **96**, 633–642. DOI: 10.1037/a0021872.

Kalshoven, K., Den Hartog, D. N., & De Hoogh, A. H. B. (2011). Ethical leadership at work questionnaire (ELW): Development and validation of a multidimensional measure. *The Leadership Quarterly*, **22**, 51–69. DOI: 10.1016/j.leaqua.2010.12.007.

Kirkpatrick, S. A., & Locke, E. A. (1991). Leadership: Do traits matter? *The Executive*, **5**, 48–60. DOI: 10.5465/ame.19991.4274679.

Kish-Gephart, J. J., Harrison, D. A., & Treviño, L. K. (2010). Bad apples, bad cases, and bad apples: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology*, **95**(1), 1–31. DOI: 10.1037/a0017103.

Kluger, A. N., & Rothstein, H. R. (1993). The influence of selection test type on applicant reactions to employment testing. *Journal of Business and Psychology*, **8**, 3–25.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive developmental approach. In T. Lickona (Eds.), *Moral Development and Behavior: Theory, Research, and Social Issues*. Holy, Rinehart, & Winston.

Loviscky, G. E., Treviño, L. K., & Jacobs, R. R. (2007). Assessing managers' ethical decision- making: An objective measure of managerial moral judgment. *Journal of Business Ethics*, **73**(3), 263–285.

Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. (1994). The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology*, **47**(4), 715–738. DOI: 10.1111/j.1744-6570.1994.tb01573.x.

Mayer, D. M., Aquino, K., Greenbaum, R. L., & Kuenzi, M. (2012). Who displays ethical leadership, and why does it matter? An examination of antecedents and consequences of ethical leadership. *Academy of Management Journal*, **55**, 151–171. DOI: 10.5465/amj.2008.0276.

Medeiros, K. E., Watts, L. L., Mulhearn, T. J., Steele, L. M., Mumford, M. D., & Connelly, S. (2017). What is working, what is not, and what we need to know: A meta-analytic review of business ethics instruction. *Journal of Academic Ethics*, **15**, 245–275. DOI: 10.1007/s10805-017-9281-2.

Mesmer-Magnus, J., & DeChurch, L. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology*, **94**, 535–546. DOI: 10.1109/emr.2012.6172774.

Mumford, M. D., Devenport, L. D., Brown, R. P., Connelly, S., Murphy, S. T., Hill, J. H., & Antes, A. L. (2006). Validation of ethical decision making measures: Evidence for a new set of measures. *Ethics & Behavior*, **16**(4), 319–345. DOI: 10.1207/s15327019eb1604_4.

Ones, D. S., & Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology*, **83**(1), 35–42. DOI: 10.1037/0021-9010.83.1.35.

Paton, E., Friedman, V., & Testa, J. (2022). What to know about Balenciaga's Campaign Controversy. *The New York Times*. https://www.nytimes.com/2022/11/28/style/balenciaga-campaign-controversy.html.

Picheta, R. Boris Johnson's staff got drunk, brawled and abused cleaners during Covid lockdowns, damning report finds. CNN (2022). https://edition.cnn.com/2022/05/25/uk/boris-johnson-sue-gray-partygate-report-gbr-intl/index.html.

Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: Comment on Haidt, 2001. *Psychological Review*, **110**(1), 193–196. DOI: 10.1037/0033-295X.110.1.193.

Rest, J. R. (1986). *Moral development: Advances in research and theory*. Praeger.

Romo, V. Audi CEO arrested in connection with Volkswagen emissions scandal (2018). https://www.npr.org/2018/06/18/620953719/audi-ceo-arrested-in-connection-with-volkswagen-emissions-scandal.

Sackett, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology*, **42**, 491–529. DOI: 10.1111/j.1744-6570.1989.tb00666.x.

Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology*, **67**(4), 401–410. DOI: 10.1037/0021-9010.67.4.401.

Sackett, P. R., & Wanek, J. E. (1996). New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel Psychology*, **49**, 787–829. DOI: 10.1111/j.1744-6570.1996.tb02450.x.

Schnmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel Selection. *Annual Review of Psychology*, **43**, 627–670. DOI: 10.1146/annurev.ps.43.020192.003211.

Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, **46**, 49–76.

Tabuchi, H. (2021, January 14). Toyota to pay a record fine for a decade of clean air act violations. *The New York Times*. https://www.nytimes.com/2021/01/14/climate/toyota-emissions-fine.html.

Tepper School of Business. Arthur Anderson Case Studies in Business Ethics (2022). http://public.tepper.cmu.edu/ethics/aa/arthurandersen.htm.

Thorton, G. C., & Rupp, D. E. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development* (pp. –17). Taylor & Francis.

Toor, S., & Ofori, G. (2009). Ethical leadership: Examining the relationships with full range leadership model, employee outcomes, and organizational culture. *Journal of Business Ethics*, **90**, 533–547. DOI: 10.1007/s10551-009-0059-3.

Treviño, L. K., Brown, M., & Hartman, L. P. (2003). A qualitative investigation of perceived executive EL: Perceptions from inside and outside the executive suite. *Human Relations*, **56**(1), 5–37. DOI: 10.1177/0018726703056001448.

Treviño, L. K., Hartman, L. P., & Brown, M. (2000). Moral person and moral manager: How executives develop a reputation for ethical leadership. *California Management Review*, **42**(4), 128–142. DOI: 10.2307/41166057.

Treviño, L. K., Sutton, C. D., & Woodman, R. W. Effects of cognitive moral development on reinforcement contingencies on ethical decision making: An experiment. In: *45th Annual Meeting of the Academy of Management, San Diego, CA* (1985).

Walumbwa, F. O., & Schaubroeck, J. (2009). Leader personality traits and employee voice behavior: Mediating roles of EL and work group psychological safety. *Journal of Applied Psychology*, **94**(5), 1275–1286. DOI: 10.1037/a0015848.

Watts, L. L., Medeiros, K. E., McIntosh, T. J., & Mulhearn, T. J. (2020). Decision biases in the context of ethics: Initial development and validation. *Personality and individual Differences*, **153**(15). DOI: 10.1016/j.paid.2019.109609.

Watts, L. L., Medeiros, K. E., Mulhearn, T. J., Steele, L. M., Connelly, S., & Mumford, M. D. (2017). Are ethics training programs improving? A meta-analytic review of past and present ethics instruction in the sciences. *Ethics & Behavior*, **27**(5), 351–384.

Wexley, K. N., & Youtz, M. A. (1985). Rater beliefs about others: Their effects on rating errors and rater accuracy. *Journal of Occupational and Organizational Psychology*, **58**(4), 265–275. DOI: 10.1111/j.2044-8325.1985.tb00200.x.s.

Yukl, G., Mahsud, R., Hassan, S., & Prussia, G. E. (2013). An improved measure of ethical leadership. *Journal of Leadership & Organizational Studies*, **20**(1), 38–48. DOI: 10.1177/1548051811429352.