

# Predicting CEO misbehavior from observables: comparative evaluation of two major personality models

Adam Szekeres<sup>1</sup> and Einar Arthur Snekkenes<sup>1</sup>

Department of Information Security and Communication Technology, Norwegian University of  
Science and Technology, Gjøvik, Norway  
{adam.szekeres,einar.snekkenes}@ntnu.no

**Abstract.** The primary purpose of this study is to demonstrate how publicly observable pieces of information can be used to build various psychological profiles that can be utilized for the prediction of behavior within a risk analysis framework<sup>1</sup>. In order to evaluate the feasibility of the proposed method, publicly available interview data is processed from a sample of chief executive officers (CEOs) using the IBM Watson Personality Insights service. The hypothesis-that group membership gives rise to a specific selection bias-is investigated by analyzing the IBM Watson-derived personality profiles at the aggregate level. The profiles are represented by two major theories of motivation and personality: the Basic Human Values and the Big Five models. Both theories are evaluated in terms of their utility for predicting adverse behavioral outcomes. The results show that both models are useful for identifying group-level differences between (1) the sample of CEOs and the general population, and (2) between two groups of CEOs, when a history of rule-breaking behavior is considered. The predictive performance evaluation conducted on the current sample shows that the binary logistic regression model built from the Basic Human Values outperforms the Big Five model, and that it provides a practically more useful measurement of individual differences. These results contribute to the development of a risk analysis method within the domain of information security, which addresses human-related risks.

**Keywords:** CEO · Psychological Profiling · Unobtrusive Measures · Basic Human Values · Big Five · Behavior Prediction.

## 1 Introduction

Strategic decisions are long-term plans produced by a small number of senior managers aimed at achieving well-defined organizational objectives, with significant impact (positive or negative) on the safety and security of organizations and information systems spanning across the entire range of the corporate hierarchy. Such decisions affect a wide range of stakeholders, thus a certain level of friction is unavoidable [4,37,32]. The principal-agent problem within the economics and management literature addresses the tension between management interests and governance objectives. The principal-agent problem arises in agency theory and describes a situation in which one party (principal)

---

<sup>1</sup> The final authenticated version is available online at [https://doi.org/10.1007/978-3-030-34866-3\\_7](https://doi.org/10.1007/978-3-030-34866-3_7).

delegates work to another party (agent) who is responsible for performing that work on behalf of the principal. The theory is concerned with resolving two problems that may arise in any agency relationship [8]. The first problem relates to the possibility that the agent's and the principal's desires or goals are in conflict, and it is difficult or expensive for the principal to verify what the agent is actually doing (i.e. hidden actions). The second problem arises from the difference between the parties' attitude towards risk, where the principal and the agent might prefer different actions due to different risk preferences and due to information asymmetry (i.e. hidden information).

Information security is a domain where negative externalities (e.g. principal-agent problem) may be present at various levels of abstraction. The highly complex threat landscape is characterized by misaligned stakeholder incentives (e.g. cost of developing sufficiently secure hardware and software vs. being first on the market, etc.), asymmetric knowledge about vulnerabilities (hidden information) and various other factors [1]. Internet of Things (IoT)-enabled critical infrastructures are becoming more and more prevalent due to their economic benefits. While they offer increased levels of automation, crucial strategic decisions are still the responsibility of people in leading positions. This may lead to situations in which the safety, security and stability of societies is increasingly dependent on the motivation of fewer and fewer key decision-makers [9].

Most information security risk analysis frameworks focus on the technological aspects and neglect the strategic decision-making perspective. The Conflicting Incentives Risk Analysis (CIRA) method developed by Rajbandhari and Snekkenes aims to bridge this gap by focusing on human motivation when addressing information security risks [24]. The method's applicability to real-world cases is limited by the lack of psychological theories that would enable the prediction of stakeholder behavior. Therefore, this study aims at evaluating two major psychological models of personality in terms of their performance for predicting undesirable stakeholder actions without direct access to subjects. The necessity for using unobtrusive profiling methods arises from the assumption that real-world stakeholders would be reluctant to explicitly reveal their motivations and they would be inclined to provide socially desirable answers when traditional assessment methods (i.e. questionnaire, interview, etc.) are utilized, which would confound the validity of the whole risk analysis process. While this study focuses on the misconduct of CEOs, the analysis is applicable to any other class of stakeholders.

## 1.1 Problem Statement

The CIRA method focuses on the misalignment between stakeholder motivations for risk identification [25]. To improve the method, it is necessary to incorporate psychological theories that enable the characterization of individual stakeholders and the prediction of their future behavior without requiring direct interaction between the analyst and the subjects. Based on these requirements, the objectives of this study are as follows:

- compare two personality models that can be used to characterize individual stakeholders,
- assess an unobtrusive profiling method's suitability for the purpose of risk analysis,
- analyze how a specific group membership gives rise to a selection bias, manifested in the psychological profiles,

- compare the predictive performance of the personality models with regard to undesirable behavioral outcomes.

## 1.2 Research Questions

Based on the aforementioned requirements and goals, the primary research question is as follows: *can publicly observable variables reflecting individual choice be used to construct psychological profiles suitable for predicting behavior in the context of risk analysis [35]?*

The following sub-questions were constructed in order to answer the main research question:

- **RQ 1:** To what extent is it feasible to use an unobtrusive profiling method to derive stakeholder characteristics?
- **RQ 2:** Is it feasible to detect a potential selection bias by analyzing personality profiles at the group-level?
- **RQ 3:** How does the *Basic Human Values* model compare to the *Big Five* model in terms of predicting stakeholder misbehavior?

This work contributes to the literature of information security risk analysis by presenting how publicly observable stakeholder data (i.e. recorded interviews) can be utilized for the purpose of risk analysis. The method relies on an existing application (IBM Watson Personality Insights), while the purpose of the analysis differs significantly from its established use cases. To assess the method's feasibility this study focuses on organizational leaders due to the fact that other classes of stakeholders might not be allowed to interact officially with the public, however the approach can be applicable to any other classes of stakeholders (e.g. CFO, COO, CIO, CISO). This study extends on previous work [35], by including an additional psychological model, and by comparatively evaluating the two personality models in terms of their capabilities for predicting real-world behavior. The paper is structured as follows: Section 2 introduces relevant theories and the IBM Watson application, Section 3 provides an overview about the methods used in the study. Results of the conducted analyses are presented in Section 4. Section 5 provides an overview about the results and their relevance, including limitations and plans for further work. Section 6 summarizes and concludes the present study.

## 2 Related work

This section provides an overview about the psychological theories, constructs and the application that served as the foundations of this study.

### 2.1 Sources of Bias

There are several research perspectives that aim to provide an explanation about the processes that guide people with certain traits or characteristics into various work positions. Extensive research investigates how different characteristics are desirable on one

hand, and how they might have a negative impact on organizational or societal objectives. Several disastrous outcomes have been linked to the decision-maker's psychological attributes, which explains the increased research interest into the ethical aspects of high-impact decision-making [34,38]. This section introduces two main mechanisms that contribute to a selection bias in executive roles (i.e. personal attraction to a specific role and selection of candidates by the board of directors).

**Selection Bias by Personal Motivations** Need for power, prestige and money are assumed to be key motivators that draw individuals to the highly competitive corporate world. Various decisions which contribute to undesirable social outcomes (e.g. exploiting sweatshop labor, environmental pollution, etc.) have been attributed to key decision-maker's psychological features. Furthermore, several organizational risks (e.g. embezzlement, bribery, etc.) can be enumerated which represent a conflict between the self-interested individual and the overall organizational objectives. One explanation for such incidents is proposed by Boddy, who discusses the over-representation of corporate psychopaths in key decision-maker positions. According to his definition corporate psychopaths are "people working in corporations who are self-serving, opportunistic, ego-centric, ruthless and shameless who can be charming, manipulative and ambitious" who are drawn to corporations since they can provide individuals with highly valued resources [3]. Corporate psychopaths are outwardly charming, and engaging, skillful at manipulating others to their own advantage, with a lack of concern for the consequences of their actions, and give a high priority for their own goals and ambitions. Their ability to demonstrate desirable traits that the organization values for a certain position is easily exploited by such individuals when presenting a charming facade, which distinguishes them from the commonly held perception of the insane psychopath.

The authors of [2] set out to investigate the prevalence and consequences of psychopathic tendencies in a sample of 203 corporate professionals taking part in a management development program. The study was motivated by the "growing public and media interest in learning more about the types of person who violate their positions of influence and trust, defraud customers, investors, friends, and family, successfully elude regulators, and appear indifferent to the financial chaos and personal suffering they create" [2]. The findings revealed the complex association between situation-congruent self-presentation and how psychopathic traits (although not classified as Antisocial Personality Disorder) can be beneficial in corporate environments. The results showed that the highest psychopathy scores were obtained from high-potential candidates in senior management positions. A noteworthy finding of the study is how the corporation evaluated individuals with several psychopathic traits. High psychopathy scores were associated with perceptions of good communication skills, strategic thinking, and creative/innovative abilities and simultaneously, with poor management style, failure to act as a team player, and poor performance appraisals (as rated by immediate bosses).

Another empirical study investigated the association between the Dark Triad personality traits and the basic human values structure [13]. The Dark Triad (Machiavellianism, Narcissism, and Psychopathy) is a popular grouping of individual differences that represent antisocial personality traits below clinical threshold. The antisocial aspect of the triad comes from the shared underlying attitudes and modes of behavior that

characterize these traits. Entitlement, superiority, dominance, manipulativeness, lack of remorse, impulsivity are the common features of the Triad [13]. The study found in two different cultures (i.e. Swedish and American) that Hedonism, Stimulation, Achievement and Power values were the highest ranking values for individuals high on Dark Triad traits. The authors claim that those characterized by high scores on the Dark Triad traits, hold values that promote Self-enhancement at the expense of others, thus treating other people as means toward their gains. The association between Self-enhancement values and the Dark Triad traits is referred to as dark value system which has further moral implications.

**Selection Bias by Role Requirements** The match between certain personality features and various organizational settings is investigated by the Person-Organization (P-O) fit theories. Morley [20] discusses a shift in recent recruitment practices in which the traditional focus on knowledge, skills and abilities (KSAs), has moved toward seeking an optimal fit between the candidate's personality, beliefs and values and the organization's espoused culture, norms and values. In a similar vein, the Attraction-Selection-Attrition (ASA) framework seeks a fit at the personal level between the candidate and the organization's work values. According to the ASA model, candidates are attracted to organizations that exhibit characteristics similar to their own, and organizations tend to select employees who are similar to the organization in key aspects [28]. Value congruence has become a widely accepted operationalization of P-O fit [16].

Role requirements vary a lot even within the same organization (e.g. managerial role requirements are different from the requirements of a production line worker). A large-sample study aimed at identifying a distinctive managerial profile in terms of the Big Five model of personality. Managers reached significantly higher scores on the following nine personality traits and facet when compared to members of other occupations: Extraversion, Assertiveness, Conscientiousness, Emotional Stability, Agreeableness, Optimism, Work Drive, Customer Service Orientation, Openness. The results can be practically useful during the personnel selection process to increase the P-O fit required for specific job types [18].

Another investigation was conducted to test the hypothesis that different work environments can be differentiated by analyzing the value structures of the workers [14]. The enterprising environment (e.g. manager, banker, financial advisor) is characterized by material and concrete goals, and requires one to lead, convince or manipulate others in order to achieve desired organizational and financial goals. According to the hypothesis Power and Achievement values are most compatible with these requirements, while the enterprising environment would inhibit the expression of Benevolence and Universalism values. The results revealed a strong positive correlation between the enterprising occupations and Power and Achievement values, while a negative correlation was observed in relation to Universalism values. This study successfully differentiated occupations based on the dominant human values that are present in each particular field, providing further evidence about a selection bias in action.

The surveyed research results highlight some of the ways through which selection bias is introduced to different work roles and occupations. First, individuals with certain traits or characteristics are attracted to specific jobs, then the active selection process by

the recruiters produces the final set of employees. Analyzing the risks to an organization largely depends on understanding the nature of these biases.

## 2.2 Conflicting Incentives Risk Analysis

The relevance of focusing on the stakeholder motivation is recognized in the Conflicting Incentives Risk Analysis (CIRA) method [25]. It identifies stakeholders (i.e. individuals), the actions that can be taken by these stakeholders and the consequences of the actions. A stakeholder is an individual who has interest in taking a certain action within the scope of the analysis. The procedure distinguishes between two types of stakeholders: *Strategy owner* (the person who is capable of executing an action) and the *Risk owner* (whose perspective is taken - the person at risk). At the core of the method is the economic concept of utility, which captures the benefit of implementing a strategy for each stakeholder. The cumulative utility encompasses several utility factors, each representing valuable aspects for the corresponding stakeholders, thus modelling an individual's motivation. Two types of risks are identified in the method: *Threat risk* relates to the perceived decrease in the total utility for the risk owner, and *Opportunity risk* relates to missed utility gains due to the strategy owner's lack of motivation (i.e. costs associated with a beneficial action). Thus, risk is conceptualized as a misalignment of incentives between these two classes of stakeholders, and risk identification focuses on uncovering activities that would be beneficial for the Strategy owner while potentially harmful for the Risk owner [31]. Threat risk closely resembles the concept of moral hazard as it captures a wide range of behaviors that are beneficial for one party and detrimental for the other who has to suffer the consequences [6]. This study focuses on Threat risks that can be attributed to the motivation of organizational leaders.

## 2.3 Theory of Basic Human Values

The theory of *Basic Human Values* by Shalom Schwartz [29] identifies 10 distinct values that are universally recognized across various cultures and provides a unified and comprehensive view on the motivation of individuals. Values both represent desirable end goals and prescribe desirable ways of acting. Six key features characterize all values:

- "Values are beliefs linked to affect.
- Values refer to desirable goals that motivate action.
- Values transcend specific actions and situations.
- Values serve as standards or criteria.
- Values are ordered by importance.
- The relative importance of multiple values guide action." [29]

Furthermore, all 10 values capture one of the three key motivational aspects that are grounded in universal requirements of human existence: needs of individuals as biological organisms, requisites of coordinated social interaction, and survival and welfare needs of groups. Values guide behavior, given that the context or situation activates the relevant values. The values form a circular structure which represents a motivational

continuum, where adjacent values are compatible with each other and opposing values are in conflict. The ten values are grouped under 4 higher dimensions as represented by Fig 1. The theory acknowledges that most actions are expressive of more than one value, and that a person’s specific value-hierarchy modifies his/her perceptions about the relevant aspects of a situation. This may give rise to different interpretations of the same situation across individuals.

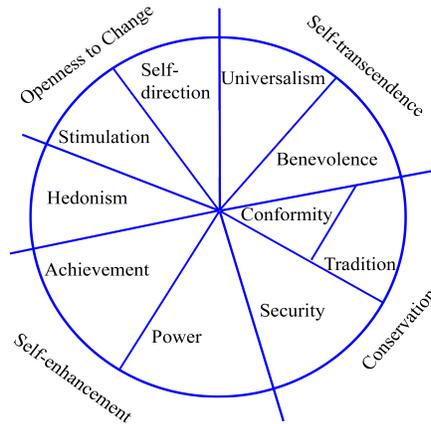


Fig. 1: Circular value structure, with 4 higher dimensions. Source: [29]

## 2.4 Big Five Personality Traits

The five factor model of personality or the *Big Five* defines five broad, distinct dimensions, that capture individual differences in terms of emotional, interpersonal experiences, recurring ways of behavior, and motivational styles [19]. The model is the result of several decades of extensive research in the domain of personality psychology, and represents one of the most widely accepted and utilized conceptualizations of personality. The five factors emerged from lexicographic investigations and are regarded as fundamental and stable dimensions of human personality, recognized across cultures. The large-scale acceptance of the model, and the consensus in relation to the utility of the Big Five provided researchers with a common framework from different traditions, which enabled productive investigations in a wide range of domains. It’s practical applicability has been demonstrated extensively in industrial/organizational, educational, clinical and other (e.g. [11]) settings. According to trait theory, individuals can be placed on a continuum along the five main dimensions, which comprise of six facets (narrower, more specific aspects of personality [19]) as shown in Table 1.

Table 1: The Big Five dimensions and narrow facets of personality, based on [19].

Openness to experience	Conscientiousness	Extraversion	Agreeableness	Neuroticism
fantasy	competence	warmth	trust	anxiety
aesthetics	order	gregariousness	straightforwardness	hostility
feelings	dutifulness	assertiveness	altruism	depression
actions	achievement striving	activity	compliance	self-consciousness
ideas	self-discipline	excitement-seeking	modesty	impulsiveness
values	deliberation	positive emotions	tender-mindedness	vulnerability

## 2.5 IBM Watson Personality Insights

Personality Insights (PI) is part of IBM’s artificial intelligence platform called Watson. Previously known for defeating the top human players in Jeopardy, the service these days is a comprehensive set of artificial intelligence solutions available for the consumer market. The service is utilized in a wide range of fields including health care, weather forecast, electric load optimization, etc. The PI utilizes machine learning solutions to uncover an individual’s psychological characteristics based on texts produced by the person. The PI service’s main use cases involve targeted marketing, customer care services, automated personalized interactions, among several others. The service produces profiles based on four different models of individual differences [35]:

1. Big Five personality model - these characteristics describe relatively stable behavioral tendencies and modes of experiences.
2. Needs - based on the earliest investigations into human motivation capturing an individual’s high-level desires.
3. Basic Human Values - values capture both desirable goals that people pursue and standards of acting, thus providing a summary about the underlying motivations behind one’s actions.
4. Consumption preferences - optimized for predicting the user’s likelihood for buying a certain product or engaging in different activities.

In terms of the Basic Human Values, the service calculates scores for five high-level dimensions: Conservation, Openness to change, Self-enhancement, Self-transcendence and Hedonism separately, whereas the original formulation identifies only four dimensions, and places Hedonism in either Openness to change or Self-enhancement. The service provides scores on all the Big Five dimensions as well as scores for each facet. For each personality model the PI computes two scores: percentile scores and raw scores. “To compute the percentile scores, IBM collected a very large data set of Twitter users (one million users for English, ...) and computed their personality portraits. IBM then compared the raw scores of each computed profile to the distribution of profiles from those data sets to determine the percentiles. The service computes normalized scores by comparing the raw score for the author’s text with results from a sample population” [12]. While the percentile scores can provide insights about an individual’s position on a trait compared to PI’s original sample, it is not well-suited to characterize an individual’s profile for the purpose of choice predictions, since the value structure relative

to a sample population does not necessarily correspond to the individual's own value priorities. To allow comparison between different populations and scenarios the service also provides raw scores which resemble scores the person would get when completing a corresponding personality inventory. Thus raw scores are more useful for making comparisons to results derived from other studies.

### 3 Methods

#### 3.1 Participants

The convenience sampling method produced a sample which consisted of 116 CEOs (105 male, 11 female), aged between 34-95 years ( $M = 59.41$ ,  $SD = 9.23$ ) with sufficient amount of texts for running accurate analysis by the IBM Watson service. The amount of text available for the individuals ranged between 264-11384 words ( $M = 3830.98$ ,  $SD = 1672.28$ ). The majority of the subjects were born in the USA ( $N = 52.6\%$ ), followed by India ( $N = 12.9\%$ ), United Kingdom ( $N = 6.9\%$ ) and 21 other countries ( $N = 27.6\%$ ). 84.4% of the sample had at least bachelor or equivalent level degrees. The total compensation for the CEOs in year 2016 ranged between \$45,936 - \$46,968,924 ( $M = \$15,988,276.78$ ,  $SD = \$10,600,982.56$ ) according to publicly available sources [27].

#### 3.2 Data Collection

The data collection and production activities (i.e. interview source identification, pre-processing, Watson analysis) are identical to those explained in [35]. In order to answer the Research Questions it was necessary to run an initial pilot study to assess the feasibility of the data collection activity. During the pilot study the first step involved the identification of relevant sources of data. To this end the Wikipedia article on the List of chief executive officers of notable companies was used that contains CEOs with diverse national and industrial backgrounds [39]. At the time of the start of the data collection the list consisted of 174 subjects. The second step involved the identification of suitable sources of information that could be linked to the individual and provided sufficient input to the Watson service for achieving its maximum precision (3000 words/subject is recommended by the service description). In this phase we relied on video interviews, interviews published in online newspapers, news articles, company communications and social media profiles. Although it was possible to collect the necessary amount of data from the individuals, the procedure was not feasible due to high diversity of contexts, the uncertainty related to the actual author of the texts and the time needed to collect the data, so in the final data collection phase this procedure was modified in the following way:

- The search was restricted to videos published on YouTube that (a) were in English, (b) the subject could be clearly identified while providing his thoughts, and (c) were supplemented with captions.
- The search then was executed by using the subject's name with the following additional terms (in the same order): - interview, talk, presentation. In case the first search term did not provide sufficient amount of text the next one was used.

- In order to achieve as high validity as possible for the analysis we aimed at collecting mainly interviews and discussions that are more spontaneous and reflective in content (thus we aimed at minimizing the reliance on well-rehearsed communications or texts written by other parties for presentation purposes).
- Each video was carefully observed in real time to check the accuracy of the captions and to ensure that only the subject’s utterances are extracted for analysis, while omitting any noise (interviewer/audience questions, false transcriptions, etc.)
- A fresh install of Google Chrome was utilized in incognito mode, to keep personalized search results to a minimum and to maximize the reproducibility of the search results.

After a sufficient amount of text was collected from the subjects, the texts were submitted to the Watson PI service producing the psychological profiles for each individual [35].

For the purpose of a more fine grained analysis, CEOs that have been associated with various forms of rule breaking behavior leading to moral hazard have been identified in the current sample. To this end extensive web searches were conducted with the name of the individual and the additional search term (e.g. fraud, scandal, corruption). The first 20 search results were screened for each subject in order to identify possible associations with moral hazard. Using a broad sense of the moral hazard concept, any behavior was eligible for inclusion which had a negative effect on the reputation of the organization by drawing public attention to the underlying misconduct (irrespective of the nature of the misconduct) and the actions were conducted under the administration of the CEO in focus. The activities included: bribery of public officials, tax evasion, accounting fraud, insider deals, ethical misconduct, etc. The procedure resulted in the identification of 31 CEOs (26.7% of the sample) associated with undesirable behavior, and enabled profile comparisons between the two CEO groups [35].

### 3.3 The Concept of Difference

To characterize group differences several approaches were considered. In the first approach the percentile scores derived from the Watson PI service were used, that inherently contain a comparison between the subject’s results and the original sample’s distribution, on which the service was validated ( $N \sim 1$  million users) [12]. This approach provides an understanding about the CEO sample’s overall position across each personality dimension. Since the parameters are not publicly available for the original sample, a reference distribution was used to test differences between the current and the hypothesized original sample.

The second approach utilizes the raw scores derived from the PI service, which are equivalent to the scores one would get when completing an actual psychometric test (as suggested by the Watson manual [12]). These scores can be compared to results obtained from different populations, therefore are more suitable for validation. The second procedure followed this line of reasoning, and aimed at identifying differences between the profiles of CEOs and the general population.

However, rank orders in isolation do not provide all the necessary information about and individual’s trade-off decisions, since a preference reversal (i.e. choosing different

strategies with the same value orders among individuals) is possible. Considering this fact and in accordance with the theory's formulation, the relative importance of values should be analyzed when certain strategies are evaluated. Furthermore, since several studies use different instruments and methodologies for assessing the personality models or use different levels of analysis, it was necessary to enhance the compatibility and comparability of research findings [17]. To this end, in the third procedure the raw scores were summed across all dimensions, and each score was multiplied by the  $\text{Sum}^{-1}$ , to quantify each value's contribution to the overall utility ( $=1$ ). The same procedure was carried out for research results that served as reference for the comparisons. This approach provides an assessment of an individual's personality profile independent of the instrument used for conducting the profiling. All analyses were conducted with SPSS 25 by IBM.

## 4 Results

### 4.1 Percentile score comparisons with Watson PI Sample

The first procedure aimed at detecting the existence of a selection bias using the percentile scores of each personality model. Percentile scores from the Basic Human Values and the Big Five scores were transformed by mapping them to a standard normal distribution, then for each dimension One-Sample t-tests were conducted with a reference standard normal distribution ( $M = 0$ ) to assess whether the scores were drawn from the specific hypothesized distribution.

**Basic Human Values** The results indicate that the group means for Conservation ( $M = -1.57$ ),  $t(115) = -29.30$ , Hedonism ( $M = -1.95$ ),  $t(115) = -81.24$ , Self-enhancement ( $M = -1.24$ ),  $t(115) = -30.06$ , and Self-transcendence ( $M = -0.84$ ),  $t(115) = -21.19$ , were significantly different from the reference distribution's mean scores,  $p \leq 0.001$  for each. The group mean score of Openness to change ( $M = 0.06$ ),  $t(115) = 1.14$ ,  $p = 0.25$  was not significantly different from the hypothesized population mean. Fig 2 shows the distribution of all the values based on the transformed percentile scores.

**Big Five** The same procedure was conducted for the Big Five dimensions and the results indicate that mean scores for the Big Five dimensions Openness to experience ( $M = 1.94$ ),  $t(115) = 51.80$ , Conscientiousness ( $M = 0.62$ ),  $t(115) = 14.58$ , Agreeableness ( $M = -0.79$ ),  $t(115) = -11.33$ , and Neuroticism ( $M = 0.79$ ),  $t(115) = 23.90$  were significantly different from the reference distribution's mean scores,  $p \leq 0.001$  for each. The mean score for Extraversion ( $M = -0.03$ ),  $t(115) = -0.62$ ,  $p = 0.54$  was not significantly different from the hypothesized population mean. Fig 3 shows the distribution of scores on all the dimensions of the Big Five personality model using the transformed percentile scores.

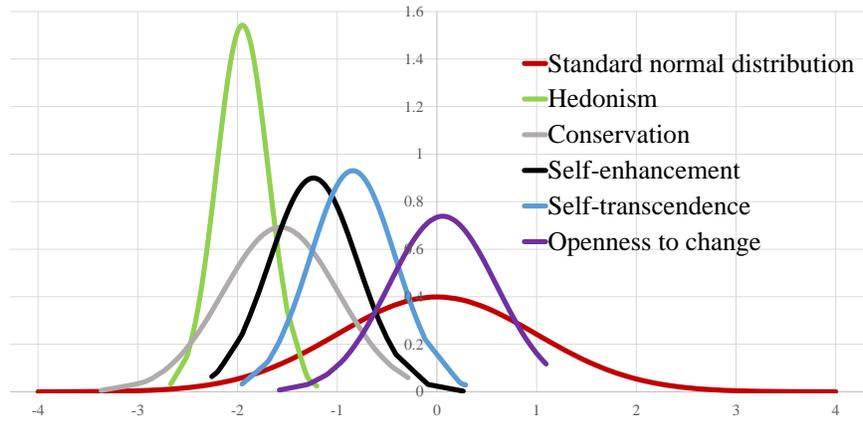


Fig. 2: Basic Human Values percentile score distributions. [35]

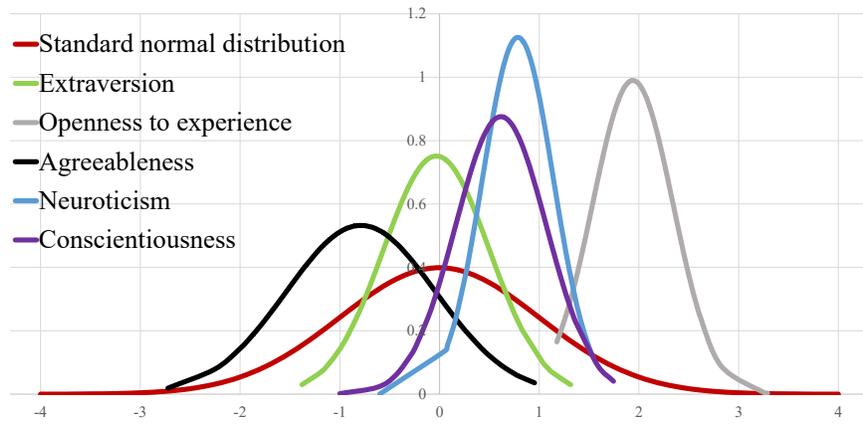


Fig. 3: Big Five percentile score distributions.

## 4.2 Raw score comparison with samples from other studies

Raw scores provide information on how an individual would be scored when providing answers on the related personality inventory. Therefore, raw scores are more suitable for performing comparisons with results obtained from other published research studies.

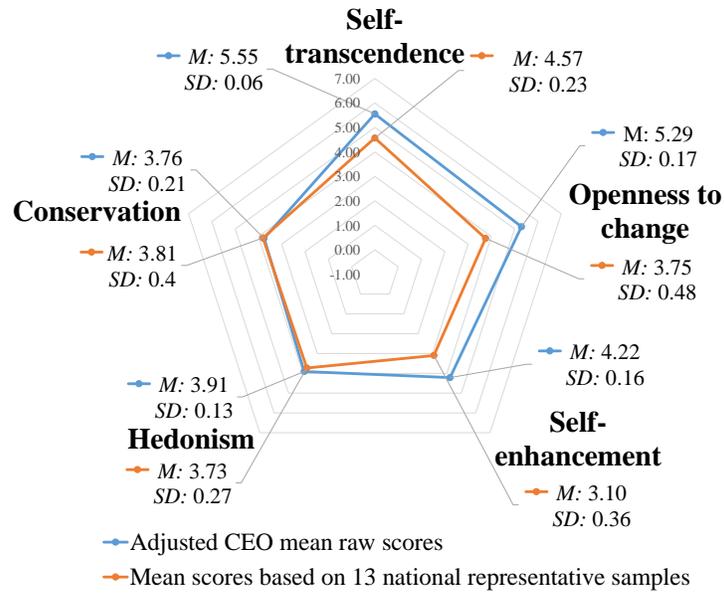
**Basic Human Values** In the following procedure the raw scores have been transformed to match with the original scale's scoring system used in the study by Schwartz and Bardi [30]. The representative or near-representative samples provide the necessary comparison that allows for a more detailed description of the value profiles. Fig 4a shows the general population's value priorities compared with the CEO value priorities based on the raw scores.

**Big Five** The Big Five profile scores were compared to a large-scale study, which gathered personality profiles from a sample of 132,515 American and Canadian internet users aged between 21-60 years [33]. The scores are reported using the percentage of maximum possible (POMP) scoring method, which is a metric constructed by a linear transformation of raw metric scores into a 0 to 100 scale, where 0 represents the minimum possible score and 100 represents the maximum possible score [5]. Therefore these scores are directly comparable to the raw scores derived from the IBM PI service (range 0-1). Fig 4b shows the mean score comparison between the large scale sample and the current CEO sample.

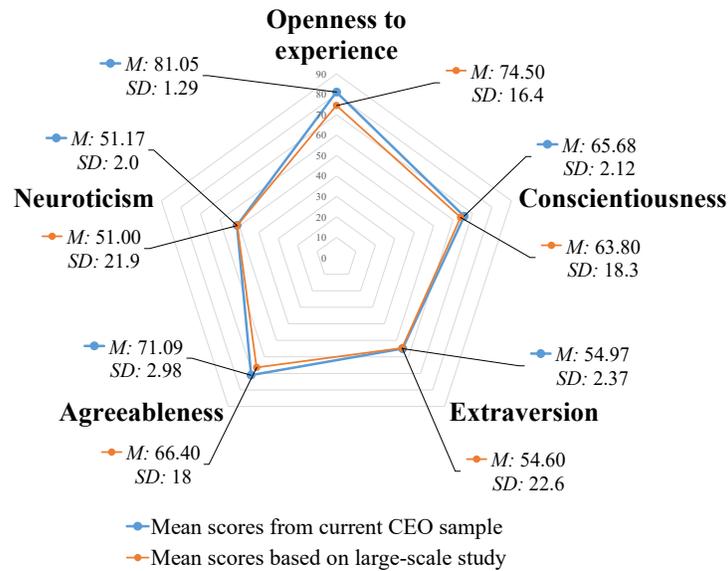
## 4.3 Comparison between CEO sub-groups

The following procedures aimed at analyzing differences among the two groups in the present CEO sample, based on a classification that identified a track record of rule-breaking behavior.

**Basic Human Values** For the purpose of individual level choice prediction, the relative importance among the values has to be considered according to the original formulation of the theory. To this end, the profiles from the two CEO groups were converted to reflect relative importance among the Basic Human Values as described in 3.3, and five independent samples t-tests were performed on the raw scores to compare each value's importance across the two classes of CEOs to detect differences in the value profiles. Fig 5 illustrates the relative importance of values among the two CEO groups and the general population. Rank order of the values is marked above the bars where the CEO sample's ranking is followed by the general population's rank on each value. Table 2 shows the results of the performed t-tests.



(a) Basic Human Values profile.



(b) Big Five personality dimensions.

Fig. 4: Comparison of CEO raw profile scores from the IBM Watson PI service to research results obtained from representative samples.

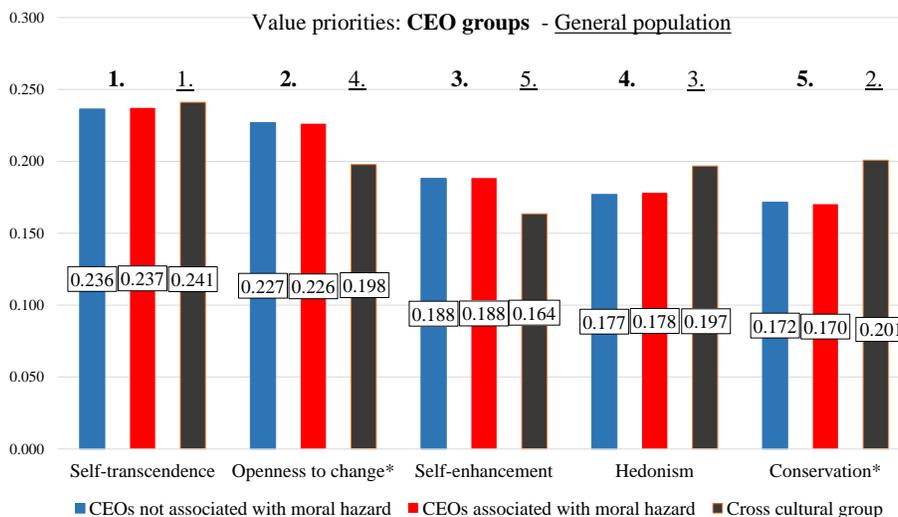


Fig. 5: Comparison between the relative importance of the Basic Human Values among two groups of CEOs and general population. \* marks a significant difference between the two CEO groups in terms of the importance of corresponding values [35].

Table 2: Results of the independent samples t-tests among two CEO groups using the Basic Human Values model [35].

Values	CEO raw scores associated with moral hazard (n = 31)		CEO raw scores not associated with moral hazard (n = 85)		t-test
	M	SD	M	SD	
Self-transcendence	0.82	0.01	0.82	0.01	n.s.
Openness to change	0.78	0.02	0.79	0.02	2.20*
Self-enhancement	0.65	0.02	0.65	0.02	n.s.
Hedonism	0.61	0.01	0.61	0.02	n.s.
Conservation	0.59	0.02	0.60	0.03	2.07*

Note. \*p < .05; two-tailed.  
M = Mean. SD = Standard Deviation

**Big Five** The same grouping was used when running five independent samples t-tests to analyze which dimensions of the Big Five personality model indicate group-level differences among the two classes of CEOs. Table 3 presents results of the tests. Extraversion was the only dimension with significant difference between CEOs who have been linked to moral hazard, and those who have not, while the other dimensions are statically indistinguishable from each other between these sub-groups.

Table 3: Independent samples t-tests among two CEO groups with the Big Five model.

Big Five dimensions	CEO raw scores associated with moral hazard (n = 31)		CEO raw scores not associated with moral hazard (n = 85)		t-test
	M	SD	M	SD	
Openness to experience	0.81	0.01	0.82	0.01	n.s.
Conscientiousness	0.65	0.02	0.66	0.02	n.s.
Extraversion	0.54	0.02	0.55	0.02	1.98*
Agreeableness	0.71	0.03	0.71	0.03	n.s.
Neuroticism	0.51	0.02	0.51	0.02	n.s.

*Note.* \*p = .05; two-tailed.

M = Mean. SD = Standard Deviation

#### 4.4 Predictive performance comparison of the Basic Human Values and Big Five models

The final set of analyses focused on comparing the predictive capabilities of the two different personality models. Raw scores were transformed to z-scores and the guidelines provided by [23] were followed when conducting the analyses and presenting the results. Binary logistic regression models were built separately and the variables were entered in a single step in order to assess the overall predictive performance of the two theories. The dependent variable had two levels (i.e. clean track record vs evidence of rule-breaking, coded as 0 and 1). In case of the Basic Human Values model, the overall model evaluation proved that the model provided a significant improvement over the intercept only model, and the inferential goodness-of-fit test (Hosmer–Lemeshow) was insignificant ( $p > .05$ ), suggesting that the model was fit to the data well. In case of the Big Five model, the overall model evaluation was not significantly better than the null-model.

Table 4 presents the overall model using the Basic Human values as predictors and Table 5 shows the details of the predictive performance evaluation of the model. For the Big Five personality dimensions, Table 6 shows the overall model and Table 7 shows the performance metrics related to this conceptualization of personality. Sensitivity and specificity were computed according to the guidelines provided by [10].

A final model was built, to test whether a combination of predictors from the two different theories could yield improved predictive performance. Predictors were entered

by using the conditional forward stepwise selection method with entry testing based on the significance of the score statistic, and removal testing based on the probability of a likelihood-ratio statistic based on conditional parameter estimates. The first block contained all Basic Human Values as predictors, and the next block contained all the Big Five dimensions. The resulting final model is shown in Table 8.

Table 4: Logistic regression model using the Basic Human Values profiles.

Predictor	$\beta$	SE $\beta$	Wald's $\chi^2$	df	p	Odds ratio
Constant	-1.15	0.24	23.68	1	0.00*	0.32
Conservation	-0.50	0.27	3.47	1	0.06	0.61
Openness to change	-0.74	0.29	6.38	1	0.01*	0.48
Hedonism	-0.05	0.29	0.03	1	0.87	0.87
Self-enhancement	0.22	0.32	0.47	1	0.49	1.24
Self-transcendence	-0.24	0.28	0.78	1	0.38	0.78
Test			$\chi^2$	df	p	
<b>Overall model evaluation</b>			<b>12.82</b>	<b>5</b>	<b>0.02*</b>	
Goodness-of-fit-test:						
Hosmer & Lemeshow			12.34	8	0.14	

*Note.* \*p < 0.05. Cox and Snell R<sup>2</sup> = .105. Nagelkerke R<sup>2</sup> = .152.

Table 5: Predictive performance evaluation of the Basic Human Values model.

Observed	Predicted		% Correct
	Yes	No	
Yes	7	24	22.6
No	4	81	95.3
Overall %			75.9

*Note.* TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative, Sensitivity = TP / (TP + FN) = 22.6%. Specificity = TN / (TN + FP) = 95.3%.

Table 6: Logistic regression model using the Big Five profiles.

Predictor	$\beta$	$SE \beta$	Wald's $\chi^2$	$df$	$p$	Odds ratio
Constant	-1.11	0.23	23.66	1	0.00*	0.33
Openness to experience	-0.09	0.25	0.13	1	0.71	0.91
Conscientiousness	-0.40	0.30	1.75	1	0.19	0.67
Extraversion	-0.61	0.27	5.12	1	0.02*	0.54
Agreeableness	-0.08	0.26	0.09	1	0.77	0.93
Neuroticism	0.70	0.31	4.97	1	0.03*	2.01
Test			$\chi^2$	$df$	$p$	
<b>Overall model evaluation</b>			<b>10.76</b>	<b>5</b>	<b>0.06</b>	
Goodness-of-fit-test:						
Hosmer & Lemeshow			13.65	8	0.09	

Note. \* $p < 0.05$  Cox and Snell  $R^2 = .089$ . Nagelkerke  $R^2 = .129$ .

Table 7: Predictive performance evaluation of the Big Five Model.

Observed	Predicted		% Correct
	Yes	No	
Yes	4	27	12.9
No	2	83	97.6
Overall %			75

Note. TP: True Positive, TN: True Negative,  
 FP: False Positive, FN: False Negative  
 Sensitivity =  $TP / (TP + FN) = 12.9\%$ .  
 Specificity =  $TN / (TN + FP) = 97.6\%$ .

## 5 Discussion

This study aimed at analyzing two different models of personality to detect a selection bias among chief executive officers by using text-based personality inferences provided by the IBM Watson PI service. Our results suggest that a selection bias can be detected by the Basic Human Values and the Big Five models as well. According to the results there are clearly identifiable differences among the universally established value structures in the general population and the sample of CEOs. Furthermore, differences can be identified in the Big Five profiles between these groups. This marked difference is interpreted as an evidence of a selection bias among organizational leaders. The importance of these differences in the motivational and personality structures is discussed in this section with directions for further work.

The analyses based on percentile scores revealed that both the Basic Human Value structure and the Big Five profile of the current sample of CEOs shows significant differences from the Watson Personality Insight service's hypothesized sample. With the exception of Openness to change (Basic Human Values) and Extraversion (Big Five),

Table 8: Results of the logistic regression model by combining predictors from both theories.

Predictor	$\beta$	SE $\beta$	Wald's $\chi^2$	df	p	Odds ratio
Constant	-1.13	0.23	23.73	1	0.00**	0.32
Openness to change	-0.61	0.23	6.93	1	0.01**	0.54
Conservation	-0.59	0.23	6.34	1	0.01**	0.56
Test			$\chi^2$	df	p	
<b>Overall model evaluation</b>			<b>11.57</b>	<b>2</b>	<b>0.00**</b>	
Goodness-of-fit-test:						
Hosmer & Lemeshow			9.36	8	0.31	
<i>Note.</i> **p $\leq$ 0.01 Cox and Snell R <sup>2</sup> = .095. Nagelkerke R <sup>2</sup> = .138.						

all other dimensions of the corresponding models showed differences from the original sample's hypothesized distributions. Due to the large sample size used during the validation of the service, it can be regarded as an indicator of valid differences between these samples, however due to the lack of detailed information about the original sample it is not possible to draw further conclusions based on percentile scores.

The second set of analyses focused on the utility of raw scores and the comparisons relied on established results from other large-scale studies. In terms of the Basic Human Values, the investigations revealed that there are important differences between the rank order of values among CEOs and the general population. While Self-transcendence values (i.e. care for the welfare of closely related others, as well as care for all the people and for nature) are most important for both groups the similarities between CEOs and non-CEOs end at this point. Openness to change (i.e. self-direction, independence, creating, stimulation and seeking out challenges) ranks as the second most important value in case of corporate leaders, while it is the second least important motivational factor for the general population. Openness to change and Conservation values can be found at opposing sides of the motivational circumplex, which reflects that decisions that promote the obtaining of a particular goal inhibit the simultaneous fulfillment of the competing need. Therefore a high priority given to Openness to change values would result in choices increasing novelty and chances for expressions of independent action at the expense of maintaining stability and stability. Self-enhancement values (i.e. expression of competence, achievement of status and control over others) rank at the third position for CEOs, while it is the least important motivational value in the general population. Although one might expect that leaders of world-leading organizations (expressing power and achievement values) would be mainly motivated by Self-enhancement values at the expense of Self-transcendence values, these results contradict this expectation. The rank order difference of Self-enhancement values between non-CEOs (5.) and CEOs (3.) however clearly expresses their preference for high social status and prestige. While for non-CEOs, the second most important motivational tendencies relate to Conservation values (i.e. security, safety of self and of society, restraint of actions likely to harm others, respect for customs), these goals are less important to leaders, as it ranks the lowest

on their motivational hierarchy, indicating that actions promoting Conservation values have a much lower intrinsic motivational effect (e.g. in order to make an action appear at least as rewarding as an action expressing Openness to change values it has to be incentivized much more externally). The relative importance of values matches closely with the various Enterprising value profiles as discussed in [14], placing CEOs close to other occupations characterized by material and concrete goals.

In terms of the Big Five model, raw scores are more closely matched with those of the general population. A higher mean score on Openness to experience indicates elevated preference for adventure, novel experiences, curiosity and intellectual challenges, which can be seen as a desirable attribute for organizational leaders promoting growth, and motivating employees. On the other hand, it is also related to risk-taking behavior. Higher scores on Agreeableness is surprising, since lower scores are associated with competitiveness and self-direction, which are considered important leader characteristics. A more detailed analysis of the facet scores on this dimension could reveal which aspects contribute to the elevated score.

The third set of analyses aimed at identifying between-group differences within the current CEO sample, when previous history of misbehavior is taken into account. Based on the Basic Human Values model a slight, but significantly lower relative importance attributed to Openness to change and Conservation values was associated with various undesirable behaviors that can be detrimental to the reputation of the organization led by the particular CEOs. Out of the Big Five dimensions only Extraversion showed a significant difference between groups, where lower Extraversion scores were associated with undesirable actions. This finding is similar to the results obtained by [26] which showed that self-reported computer criminal behavior was associated with higher levels of Introversion (i.e. lower levels on Extraversion) and similarly, no other significant differences were found between the two student groups in terms of the Big Five profiles.

The final evaluations were conducted to test the utility of the two major theories for the prediction of behavioral outcomes. Since both theories aim at providing a comprehensive view on the organization of the human psyche by identifying basic and necessary structures that are pervasive and relatively stable within individuals [19], they were used in two separate logistic regression models as a single unit. A third model was built to investigate whether a combination of the two theories could achieve improvements over any of the models in isolation. The model built from the Basic Human Values represented a significant improvement from the null-model, and achieved the highest score on the  $R^2$  metric ( $R^2 = 0.152$ ) out of the three models. The logistic regression analysis including all the Big Five dimensions resulted in a model that was not significantly better than a null-model, which purely guesses the majority class. This finding is surprising considering that the Big Five is the most widely accepted and utilized model of personality, and several studies claim that it has substantial predictive utility in a wide range of domains [22,21]. The final combined model contained no predictors from the Big Five (none of them reached the inclusion criteria), thus all variance explained by the model is attributed to Basic Human Values. The overall model reached a higher significance level (i.e. lower p value) at the expense of some explained variance (change from the model with all Basic Human Values in terms of  $R^2$  is:  $-0.014$ ). This results suggests that

the two models are to a great extent overlapping, but the Basic Human Values model might be more comprehensive.

A limitation of the present study is the relatively small sample size, which can be extended in future studies, since the method of analyzing personality profiles by using the Watson PI service is a feasible method for gathering information about the motivation of decision makers for the purpose of risk analysis. Sample size limitations may potentially hamper the performance of the binary logistic regression models, therefore it would be necessary to increase the number of observations for events and non-events for improved models. It would potentially lead to better sensitivity and specificity scores, and in order to compute positive and negative predictive values, the prevalence rates of offending behavior could be investigated in future work [10]. Furthermore, a more detailed description and classification of the various forms of rule-breaking behavior could clarify the connection between the particular strategy owner's profile and the nature of negative impact inflicted upon the organization, to achieve a better assessment of the risks relating to individuals.

In a risk analysis setting direct access to subjects is a major limitation. Since previous work has established the extent to which the most easily available pieces of information (i.e. demographic features) are useful for constructing stakeholder profiles [36], future work will focus on other classes of observable features (e.g. ownership of items [7], or various forms of online behavior with digital traces [15], etc.) for the construction of psychological profiles.

## 6 Conclusion

This exploratory study aimed at analyzing how publicly observable pieces of information (i.e. spoken texts, group membership) associated with individuals can be utilized to detect a selection bias among groups of people working in similar roles. A set of chief executive officers were selected for the purpose of testing the methods' usefulness, for two main reasons: the availability of relevant and necessary data, and due to the significance of the role they play in organizations. However, the principles presented in this study are applicable to other classes of stakeholders as well, and are not limited to the CEO role. The selection bias is revealed by patterns of specific psychological characteristics that distinguish CEOs from the general population. Furthermore, within the analyzed CEO sample, additional differences could be detected among two groups that were generated by considering available evidence about rule-breaking behavior (i.e. association with moral hazard).

The specific psychological differences were investigated through two major theories that account for stable individual differences among people. The Big Five personality model is evaluated against the Basic Human Values model in terms of group-level differences, and in terms of predictive capabilities. The results show that both models are useful in detecting a hypothesized selection bias, but the Basic Human Values model performs better in terms of predictive utility as a comprehensive model of individual differences and motivation. The unobtrusive nature of the text analysis combined with the procedures described in this study enables risk analysts to study human-related risks in various environments where adversarial stakeholder behavior is assumed and it is cru-

cial to be prepared against undesirable consequences of those actions (e.g. information security).

## ACKNOWLEDGEMENTS

This work was partially supported by the project IoTSec – Security in IoT for Smart Grids, with number 248113/O70 part of the IKTPLUS program funded by the Norwegian Research Council.

## References

1. Anderson, R., Moore, T.: Information security: where computer science, economics and psychology meet. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **367**, 2717–2727 (2009)
2. Babiak, P., Neumann, C.S., Hare, R.D.: Corporate psychopathy: Talking the walk. *Behavioral sciences & the law* **28**(2), 174–193 (2010)
3. Boddy, C.R.: The implications of corporate psychopaths for business and society: An initial examination and a call to arms. *Australasian Journal of Business and Behavioural Sciences* **1**(2), 30–40 (2005)
4. Cohan, J.A.: " i didn't know" and" i was only doing my job": Has corporate governance careened out of control? a case study of enron's information myopia. *Journal of Business Ethics* **40**(3), 275–299 (2002)
5. Cohen, P., Cohen, J., Aiken, L.S., West, S.G.: The problem of units and the circumstance for pomp. *Multivariate behavioral research* **34**(3), 315–346 (1999)
6. Dembe, A.E., Boden, L.I.: Moral hazard: a question of morality? *New Solutions: A Journal of Environmental and Occupational Health Policy* **10**(3), 257–279 (2000)
7. Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G.: Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **9**(3), 522–550 (2011)
8. Eisenhardt, K.M.: Agency theory: An assessment and review. *Academy of management review* **14**(1), 57–74 (1989)
9. Fosso, O.B., Molinas, M., Sand, K., Coldevin, G.H.: Moving towards the smart grid: The norwegian case. In: *Power Electronics Conference (IPEC-Hiroshima 2014-ECCE-ASIA), 2014 International*. pp. 1861–1867. IEEE (2014)
10. Glaros, A.G., Kline, R.B.: Understanding the accuracy of tests with cutting scores: The sensitivity, specificity, and predictive value model. *Journal of clinical psychology* **44**(6) (1988)
11. Grassegger, H., Krogerus, M.: The data that turned the world upside down (Jan 2017), [https://motherboard.vice.com/en\\_us/article/mg9vvn/how-our-likes-helped-trump-win](https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win)
12. IBM: The science behind the service (2017), <https://console.bluemix.net/docs/services/personality-insights/science.html#science>, [Online; accessed 14-February-2018]
13. Kajonius, P.J., Persson, B.N., Jonason, P.K.: Hedonism, achievement, and power: Universal values that characterize the dark triad. *Personality and Individual Differences* **77**, 173–178 (2015)
14. Knafo, A., Sagiv, L.: Values and work environment: Mapping 32 occupations. *European Journal of Psychology of Education* **19**(3), 255–273 (2004)
15. Kosinski, M., Stillwell, D., Graepel, T.: Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* **110**(15), 5802–5805 (2013)

16. Kristof-Brown, A.L., Zimmerman, R.D., Johnson, E.C.: Consequences of individuals' fit at work: A meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel psychology* **58**(2), 281–342 (2005)
17. Lindeman, M., Verkasalo, M.: Measuring values with the short schwartz's value survey. *Journal of personality assessment* **85**(2), 170–178 (2005)
18. Lounsbury, J.W., Sundstrom, E.D., Gibson, L.W., Loveland, J.M., Drost, A.W.: Core personality traits of managers. *Journal of Managerial Psychology* **31**(2), 434–450 (2016)
19. McCrae, R.R., John, O.P.: An introduction to the five-factor model and its applications. *Journal of personality* **60**(2), 175–215 (1992)
20. Morley, M.J.: Person-organization fit. *Journal of Managerial Psychology* **22**(2), 109–117 (2007)
21. Paunonen, S.V.: Big five factors of personality and replicated predictions of behavior. *Journal of personality and social psychology* **84**(2), 411 (2003)
22. Paunonen, S.V., Ashton, M.C.: Big five factors and facets and the prediction of behavior. *Journal of personality and social psychology* **81**(3), 524 (2001)
23. Peng, C.Y.J., Lee, K.L., Ingersoll, G.M.: An introduction to logistic regression analysis and reporting. *The journal of educational research* **96**(1), 3–14 (2002)
24. Rajbhandari, L., Snekkenes, E.: Intended actions: Risk is conflicting incentives. *Information Security* pp. 370–386 (2012)
25. Rajbhandari, L., Snekkenes, E.: Using the conflicting incentives risk analysis method. In: *IFIP International Information Security Conference*. pp. 315–329. Springer (2013)
26. Rogers, M.K., Seigfried, K., Tidke, K.: Self-reported computer criminal behavior: A psychological analysis. *digital investigation* **3**, 116–120 (2006)
27. Salary.com: Executive compensation it starts with the ceo (2004), <https://www.salary.com/executive-compensation-it-starts-with-the-ceo/>, [Online; accessed 14-February-2018]
28. Schneider, B., Goldstein, Harold, W., Smith, D.B.: The asa framework: An update. *Personnel psychology* **48**(4), 747–773 (1995)
29. Schwartz, S.H.: An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture* **2**(1), 11 (2012)
30. Schwartz, S.H., Bardi, A.: Value hierarchies across cultures: Taking a similarities perspective. *Journal of cross-cultural psychology* **32**(3), 268–290 (2001)
31. Snekkenes, E.: Position paper: Privacy risk analysis is about understanding conflicting incentives. In: *IFIP Working Conference on Policies and Research in Identity Management*. pp. 100–103. Springer (2013)
32. Soltani, B.: The anatomy of corporate fraud: A comparative analysis of high profile american and european corporate scandals. *Journal of business ethics* **120**(2), 251–274 (2014)
33. Srivastava, S., John, O.P., Gosling, S.D., Potter, J.: Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of personality and social psychology* **84**(5), 1041 (2003)
34. Stenmark, C.K., Mumford, M.D.: Situational impacts on leader ethical decision-making. *The Leadership Quarterly* **22**(5), 942–955 (2011)
35. Szekeres, A., Snekkenes, E.A.: Unobtrusive psychological profiling for risk analysis. In: *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications - Volume 1: SECRYPT*. pp. 210–220. SciTePress (2018)
36. Szekeres, A., Wasnik, P.S., Snekkenes, E.A.: Using demographic features for the prediction of basic human values underlying stakeholder motivation. In: *Proceedings of the 21st International Conference on Enterprise Information Systems - Volume 2: ICEIS*. pp. 377–389. INSTICC, SciTePress (2019)
37. Van Peurse, K.A., Zhou, M., Flood, T., Buttimore, J.: Three cases of corporate fraud: an audit perspective (2007)

38. Whyte, G.: Decision failures: Why they occur and how to prevent them. *Academy of Management Perspectives* **5**(3), 23–31 (1991)
39. Wikipedia: List of chief executive officers Wikipedia, the free encyclopedia (2004), [https://en.wikipedia.org/wiki/List\\_of\\_chief\\_executive\\_officers](https://en.wikipedia.org/wiki/List_of_chief_executive_officers), [Online; accessed 06-December-2017]