Toxic Workers*

Michael Housman
Cornerstone OnDemand

Dylan Minor
Kellogg School of Management, Northwestern University

November, 2015

Abstract

While there has been a strong focus in past research on discovering and developing top performers in the workplace, less attention has been paid to the question of how to manage those workers on the opposite side of the spectrum: those who are harmful to organizational performance. In extreme cases, aside from hurting performance, such workers can generate enormous regulatory and legal fees and liabilities for the firm. We explore a large novel dataset of over 50,000 workers across 11 different firms to document a variety of aspects of workers' characteristics and circumstances that lead them to engage in what we call "toxic" behavior. We also explore the relationship between toxicity and productivity, and the ripple effect that a toxic worker has on her peers. Finally, we find that avoiding a toxic worker (or converting him to an average worker) enhances performance to a much greater extent than replacing an average worker with a superstar worker.

Keywords: strategic human resource management, misconduct, worker productivity, ethics, superstar

*The authors thank Steve Morseman for his invaluable research support. Comments from Jennifer Brown and Lamar Pierce are gratefully acknowledged. A final thank you to the participants from the Academy of Management meeting (2015), Alliance for Research on Corporate Sustainability conference (2015), Goethe University, Harvard Business School I&I (2014), NBER Summer Institute (2015), Northwestern Law School, Queens University, and University of Cologne. This research was conducted in collaboration with the Workforce Science Project of the Searle Center for Law, Regulation and Economic Growth at Northwestern University.

1 Introduction

There is an abundance of work that explores how to find, develop, and incentivize top performers so as to enhance organizational performance (Lazear and Oyer (2007) and Gibbons and Roberts (2013)). What this work makes clear is that hiring the right people is very important. Finding the positive outliers—the "stars"—can substantially increase performance (e.g., Azoulay et al. (2010), Sauermann and Cohen (2010), and Oettl (2012)). However, there are outliers from the other side of the distribution that have yet to receive as much academic attention: those workers who are harmful to an organization's performance (Pierce and Balasubramanian (2015)). At their most harmless, these workers could simply be a bad fit, leading to premature termination and a costly search for and training of a new worker. However, more damaging to the firm is a worker who engages in behavior that adversely affects fellow workers or other company assets; we label this type of worker "toxic." *Thus, a toxic*[1] *worker is defined as a worker that engages in behavior that is harmful to an organization, including either its property or people.* In its most dramatic form, such worker misconduct can cost a firm billions of dollars, as evidenced by JP Morgan's "London Whale" incident with Bruno Iksil.[2] At another extreme, such workers can even mortally harm current or past colleagues, as tragically witnessed in the fatal shooting of Virginia WDBJ-TV reporters by their former colleague.[3] But even relatively modest levels of toxic behavior can cause major organizational cost, including customer loss, loss of employee morale, increased turnover, and loss of legitimacy among important external stakeholders (Robinson and Bennett (1995), Litzky et al. (2006), Ermongkonchai (2010) and MacLean et al. (2010)).

The antecedents of worker misconduct are varied. There is consistent evidence that incentives can play a very important role in causing adverse outcomes (e.g., see Oberholzer-Gee and Wulf (2012) and Larkin (2014)). There is also evidence suggesting that a worker's personal characteristics are important in determining his ethical behavior (e.g., see Ford and Richardson (1994) and Loe et al. (2000)). Lazear and Oyer (2007) suggest that the selection of workers plays a role at least as important, if not more important than incentives in generating outcomes. Thus, one approach

[1] We use the term toxic to capture both the basic definition of toxic as being something harmful and also the notion that toxic workers tend to infect others with such behavior, as is shown in our empirical analysis.

[2] See http://www.bloombergview.com/quicktake/the-london-whale. In this case, it was ultimately not Mr. Iksil himself who was charged (he cooperated with authorities), but rather his supervisor and junior trader.

[3] http://www.cbsnews.com/news/virginia-wdbj-station-shooting-alleged-gunman-posted-video-of-shooting-on-social-media/

to managing toxic workers—and the approach we focus on in this paper—is simply avoiding them. However, in order to do so, we must be able to identify them ahead of time. By exploring a novel dataset of the actual conduct and characteristics of many workers that are quasi-randomly placed across and within different organizations, we find that workers who are overconfident, self-regarding, and profess to the follow the rules are much more likely to be terminated for toxic behavior.

In addition to these individual predictors, we also find evidence that an employee's work environment contributes to the likelihood of him becoming a toxic worker (e.g., Vardi (2001) , Greve et al. (2010), and Pierce and Snyder (2008)). Our paper complements the work of Pierce and Snyder (2008) who show that in the setting of automobile emissions testing a worker's environment has significant effects on her individual ethical conduct: Alongside showing that this environmental effect is also present in a broader setting, we are able to compare the importance of an individual's characteristics and identify which individual characteristics matter in determining outcomes, which adds substantially to our explanatory power. Pierce and Snyder's (2008) findings about the impact of workplace environment, while important, explained only the minority of the variation of outcomes. Thus, we integrate both situational and individual traits to explore toxicity (Hegarty and Sims (1978), Trevino (1986), Vaughan (1999)).

We also document other important features of toxic workers. Specifically, we find that toxic workers are much more productive than the average worker. Thus, as in Gino and Ariely (2012) and Frank and Obloj (2014), we find that there is a potential trade-off when employing an unethical person: they are corrupt, but they excel in work performance. This might explain how a toxic worker can persist in an organization. To explore this tradeoff further, we explicitly examine the tradeoff in increased productivity and the propensity for toxicity. When doing so, setting aside justice and ethical motivations for avoiding toxic workers, we find that avoiding toxic workers is still better for the firm in terms of net profitability, despite losing out on a highly productive worker. We also identify which personal characteristics present a tradeoff (or not) in terms of influencing both productivity and toxicity.

Finally, we estimate the value of finding a "superstar," defined as workers in the top 1% of productivity, versus the value of avoiding a toxic worker. Succeeding in the latter generates returns of nearly two-to-one compared to those generated when firms hire a superstar. This suggests more broadly that "bad" workers may have a stronger effect on the firm than "good" workers. In many other fields and disciplines researchers have found that a negative has a stronger impact than a positive. For example, in the domain of finance, loss aversion recognizes that in terms of magnitude losses have more of an impact than gains (see Tversky & Kahneman (1992)). In the discipline of psychology it is a generally accepted principle that bad experiences have a stronger hold on our psyches than good ones (Baumeister et al. (2001)). Finally, in the field of linguistics, it has been found that humans preferentially attend to negative words over positive or neutral ones (Estes and Adelman (2008)). It is no surprise to us that these findings hold true in the field of human resource management, as well.

Much of the past research on unethical worker conduct has been based on surveys, self-reports, and intention-based outcomes (Weaver and Trevino (1999), Greenberg (2002), and Pierce and Snyder (2014)). Bertrand and Mullainathan (2001) suggest that the mixed results of this past work likely stem from the challenge of empirically examining subjective data. For our setting, we study toxic workers as those who are actually terminated for toxic behavior. Thus, this paper complements this important work by linking personal characteristics of workers quasi-randomly placed within organizations with objective conduct outcomes across a very large, novel data set. Further, due to the unique quasi-random placement, results suggest causal rather than simply correlational relationships.

An alternative to avoiding toxic workers altogether, is to reform those already in the organization. With resource constraints it may not be feasible for some, if not most, organizations to pursue this second path. However, since we find that a worker's environment is also important in influencing toxic outcomes, there is some hope that through judicious management of a worker's environment, toxicity can be reduced. Nonetheless, further exploring this channel is beyond the scope of the current paper.

2 Theoretical Considerations: The Person and the Situation

In principle, there are almost infinite factors of the person and the situation that could create toxic workers. Here, we outline several that we deemed important based on the extant literature. Following this, we develop several hypotheses to guide our empirical exploration. The empirical proxies for each of these factors are discussed in section 3.2.

It is well established that other-regarding preferences determine the kinds of actions people choose and that people have heterogeneous levels of other-regardingness (e.g., see Andreoni and Miller (2002) and Fisman, Kariv and Markovitz (2007)). All things equal, those that are less other-regarding should be more predisposed to toxicity, as they do not fully internalize the cost that their behavior imposes on others (Van Lange and Kuhlman (1994) and Folmer and De Cremer (2012)). A way to capture one's degree of other-regardingness is to identify how concerned one is about taking care of another's needs (De Cremer and Van Lange (2001)). The degree of caring for others should impact the choices one makes that affect others. Specifically, those that show little concern for another's interests are less likely to refrain from damaging others and their property. Thus, *ceterus paribus*, those less caring for others should be more likely to engage in toxic behavior. Here, we refer to such a person as Self-regarding, yielding our first hypothesis.

Hypothesis 1: Self-regarding workers are more likely to be toxic

A common manifestation of overconfidence is overestimating one's own abilities (Dunning et al. (2003) and Moore and Healy (2008)). For a formalization of such an outcome, consider an example where people may be good or bad at a future task, denoted as G or B, respectively. Someone has some belief P that represents the chance that they will be G instead of B. Thus, one's expectation of the ability on a task is $P \times G + (1 - P)\, B$. It then follows that someone who is overconfident will believe that $\widetilde{P} > P$. That is, he simply believes that he is more likely to do well on the task than he has reason to objectively believe, resulting in his overestimating his ability (Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991) and Campbell et al. (2004)). With such a conceptualization of overconfidence—believing that the probability of the better outcome is higher than one ought to believe—we can link overconfidence to the likelihood of engaging in misconduct. Now have G denote the payoff from engaging in misconduct and B denote the net payoff of engaging in misconduct when caught,[4] where $G > B$. Thus, someone that is overconfident believes the expected payoff from engaging in misconduct is higher than someone who is not overconfident, as they believe the likelihood of the better outcome is higher than it really is. Hence, all things equal, those that are overconfident should be more likely to engage in misconduct. This leads to our next hypothesis.

Hypothesis 2: Overconfident workers are more likely to be toxic

An apparently straightforward factor for measuring the propensity of misconduct is whether or

[4] That is, we can think of B as G minus some cost or penalty for being caught.

not a worker agrees that the rules should always be followed: It would seem that those who always follow rules are likely to follow ethical rules, as well. Indeed, those who embrace following the rules are often deemed conscientious, and it has been shown that this personality trait predicts lower incident rates of adverse behaviors (Salgado (2002), Lee, Ashton,and Shin (2005), and Mount, Ilies, and Johnson (2006)). However, in our setting, subjects are highly incentivized to respond to a rule following question in a job application in whichever way they believe will secure them a job. Consequently, it could also be the case that those who claim the rules should be followed are more Machiavellian in nature, purporting to embrace whatever rules, characteristics, or beliefs that they believe are most likely to obtain them a job. There is strong evidence that Machiavellianism leads to deviant behavior (Hegarty and Sims (1978), Moore et al. (2012), and O'Boyle et al. (2012)). Hence, our next hypothesis captures the possibility that workers respond to a job application survey question as either a genuine conscientious person or a more Machiavellian person. Thus, whether such a rule following application question identifies good or bad workers is an empirical question.

Hypothesis 3A: Workers that claim the rules should be followed are less likely to be toxic

Hypothesis 3B: Workers that claim the rules should be followed are more likely to be toxic

Possibly the most important factor in determining whether the work environment will increase the likelihood of misconduct is the likelihood that the worker's colleagues will engage in toxic behavior. Several important studies have found positive correlations between toxic-like behavior between coworkers (Robinson and O'Leary-Kelly (1998), Gino and Bazerman (2009), and Moore and Gino (2013)). Pierce and Snyder (2008), in particular, find strong evidence that there are ethical-worker peer effects, akin to productivity peer effects. Thus, using Pierce and Snyder's logic, increased exposure to toxicity should lead to more toxicity.

Hypothesis 4: Workers with increased exposure to other toxic workers are more likely to be toxic

In terms of exposure to toxicity, certain job positions are likely to lead to different levels of toxic behavior. For example, some positions involve more regular contact with other workers, which could increase or decrease the likelihood of toxicity based on the behavior of those other workers (Mayer et al. (2009)). Furthermore, some positions are easier to monitor than others; a highly un-monitored position may be more likely to breed toxicity. There is also evidence that a job position's degree of task diversity can influence misconduct (Derfler-Rozin et al. (2015)). Alternatively, different types of jobs can yield different levels of deviance based on the different types of customers with whom an employee must interact (Pierce and Snyder (2013)). Whatever the particular source, the type of position a worker has, *ceterus paribus*, should prove to be an important factor.

Hypothesis 5: A worker's type of job position should affect his likelihood of toxicity

3 Empirical Setting

3.1 Estimation Strategy

For our empirical analysis, we first identify the factors that lead to toxicity through testing our hypotheses. The natural setting to explore the likelihood of toxicity in our setting of employment terminations is a proportional hazard model. Here a "failure" is being terminated for toxicity. We will also later explore a linear model, with which we found results similar to those using a hazard model. After exploring antecedents of toxicity, we will study the consequences of toxicity: toxic spillover and productivity.

We assume that there is some base hazard rate $h(t)$ over time. That is, over time, given that a worker is still employed, there is some natural chance of a toxic termination that can evolve in a complex way over time. In addition to this base rate, we assume the likelihood of termination for toxicity changes at any given point as a function of both the person and situation. We denote this function as $f(t|X)$, where X captures the person and situation factors. Thus, the overall chance of someone being terminated for toxic behavior at time t, given they have not yet been terminated already, is then the product $h(t)\, f(t|X)$.

In terms of estimation, we utilize a proportional hazards model (see Cameron and Trivedi (2005)) so that we can avoid assumptions about the shape of the base hazard rate $h(t)$ over time. We then assume that this base hazard is modified by

$$f(t|X) \equiv e^{\beta_p \mathbf{x}_{p,t} + \beta_s \mathbf{x}_{s,t}},$$

where $\mathbf{x}_{p,t}$ is a vector of personal traits at time t and $\mathbf{x}_{s,t}$ is a vector of situation characteristics at time t. The role of e is simply to ensure that the composite hazard rate $h(t) f(t|X)$ is never negative.[5] In other words, the baseline hazard rate $h(t)$ can be an arbitrary shape over time, but it may be modified by the person and situation at time t by $f(t|X)$.

This setup then gives us the following partial log-likelihood function to maximize:

$$\log L = \sum_{j=1}^{D} \left[\sum_{i \in D_j} x_i \beta - d_j \log \left\{ \sum_{k \in R_j} e^{x_k \beta} \right\} \right],$$

where i indexes subjects, x_i is a vector of covariates representing the person and the situation, j indexes failure times in chronological order, D_j is the set of d_j failures at time j, and R_j is the set of all subjects that could potentially fail at time j.

In our empirical setting, we have both many types of workers across workgroups as well as quasi-random matching to workgroups. Company executives explain that the typical worker placement process is a function of periodic work flow and other forces, which are not predictable. We will show that first placements are approximately random and that our main effects persist when adding workgroup fixed effects over the first placement. However, since these robustness tests yield similar results to analysis using the full dataset with all placements, we will begin our analysis with the whole, and then turn to individual parts for our robustness tests.

Since we have a very large sample and find that our results are consistent when focusing on quasi-random placement, we also abstract away from separating the notion of engaging in toxic behavior and being terminated for toxic behavior. For expositional variety, we will use multiple phrases such as "a worker engages in toxic behavior," "he is a toxic worker," and "she is terminated for toxic behavior" interchangeably. However, strictly speaking, these phrases all mean that the worker is ultimately terminated for toxic behavior.

3.2 Data

The data were obtained from a company that builds and deploys job-testing software to large employers. Many of these companies are business-process outsourcers (BPOs) that themselves

[5] Precisely, we need the image of $f(t|X)$ to be in the set $\mathbb{R}^{+} \cup \{0\}$, since it must be that $h(t) f(t|X) \in [0, +\infty)$.

provide a variety of business services (e.g., customer care, outbound sales, etc.) to their clients. The employees included in the dataset are all engaged in frontline service positions and paid on an hourly basis. From these organizations, we were able to obtain and combine three separate datasets on the basis of employee IDs:

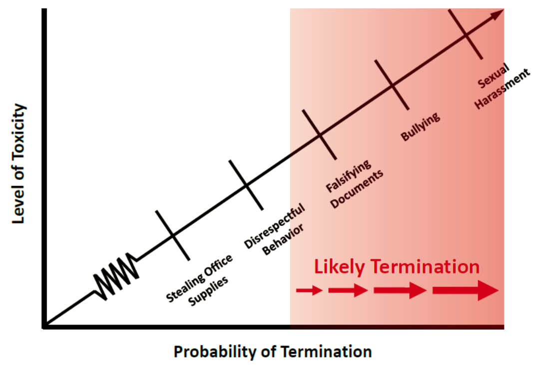
1) Job-testing data: The vendor supplying the data has developed a proprietary job test that assesses applicant fit for the position for which they are applying. We were able to obtain select questions that appeared on the test. We were unable to obtain sensitive hiring information such as gender, age, and ethnicity.

2) Attrition data: All of the companies with which the vendor engages provide an attrition feed that indicates (among other things) the employee's hire date, termination date (as applicable), reason for termination, their location, job title, and the supervisors to whom they reported while employed by the firm.

Since we can only observe toxicity by means of termination, we are generally studying the more extreme versions of toxicity, though there exists a whole continuum of toxicity. At the most benign end, we might include activities like taking home office supplies for personal use, which technically is stealing. One survey by Pendaflex suggested 75% of people admit to stealing pencils and pens from the office, whereas 38% admit to stealing company stationary.[6] On the other end of the spectrum in terms of harm and severity would be, for example, sexual harassment. In 2014, the EEOC oversaw 786 settled sexual harassment cases, which represented less than 1% of workers.[7] In our setting, we observe approximately 5% of workers terminated for toxic behavior over time. The chart below portrays how we should be capturing the more extreme versions of toxicity, which we also expect to occur less frequently. These more severe levels that we are capturing are important to understand, and where we focus on in this paper.

[6] Villano, M. (2006, May 2). The Workplace: Supplying corporate raiders. The New York Times.

[7] See http://www.eeoc.gov/eeoc/statistics/enforcement/sexual_harassment_new.cfm



3) Performance data: For a subset of employees included in our analysis, we were able to obtain daily performance data that represent productivity by measuring the average amount of time an employee required to handle a transaction and customer satisfaction scores indicating how well she served the customer.

Common employee IDs across all three of these datasets allowed us to merge them together in order to look at relationships between assessment responses and an employee's likelihood of engaging in toxic behavior. In total, the dataset covers 11 firms, 184 sub-firms (end clients of BPOs), 2,882 workgroups, each reporting to a particular supervisor, and 58,542 workers. Table 1 provides a summary of our main variables of interest.

From the assessment data, we were able to obtain several different measures of worker quality and predicted performance.

Each employment assessment is designed by an industrial-organizational psychologist and attempts to measure an employee's knowledge, skills, and abilities. We have a prediction of how self-regarding versus other-regarding (or prosocial) a given worker might be. This assessment is based on some questions that could be construed as measuring the degree of "other-regardingness" of a worker. Here is a sample set of choices presented to applicants:

1. I like to ask about other people's well-being

   OR

2. I let the past stay in the past

Choosing Statement 1 would give subjects a greater other-regarding score, whereas choosing

Statement 2 predicts a subject to be self-regarding. The overall assessment is derived from a whole collection of questions in line with the above style used to predict the overall likelihood of a worker being other-regarding versus self-regarding. The variable Self-regarding is a dummy variable with value 1 if the overall assessment predicts the subject as self-regarding; otherwise, the variable has a value of 0.

Also included in the overall assessment were questions intended to gauge an applicant's technical ability. Applicants were asked early in the assessment to self-assess their computer proficiency and they were then tested on several key computer skills. We compared their self-assessment to their actual computer proficiency in order to develop a measure of applicant self-confidence. The variable Confidence is constructed by extracting the residual from a regression of actual skills (i.e., measured skills) on promised skills (i.e., given by the worker). This variable is a measure of how much the actual skills exceed or fall short of the promised skills.[8]

We acknowledge that this variable could also be a measure of honesty. Though, 11% of workers actually *under-promise* their skills level, which would make this an unlikely measure of dishonesty in such an incentivized setting in which a worker who under-promises would be reporting against her best interest. Furthermore, if it becomes apparent after hire that a worker has lower-than-promised skills, there is a real chance that she will be terminated. Thus, it seems this variable is more a measure of confidence in one's own abilities than a measure of honesty.

Some questions on the assessment asked applicants about their propensity to follow rules. We were also able to obtain these questions and the applicant responses in order to understand whether there was a relationship between the applicant's response and her likelihood to engage in toxic behavior. Applicants were asked to choose one option from each of the following sets of statements:

1. I believe that rules are made to be followed. OR

2. Sometimes it's necessary to break the rules to accomplish something

and

1. I like to see new places and experience new things. OR

[8] We also calculate Confidence as simply the difference between stated and actual skills without using regression analysis, and the results are similar. In absolute terms, we find that roughly 11% of workers promise lower skill than they deliver, 55% deliver as they promise, and 34% overpromise.

2. I complete activities according to the rules.

For each of the rule-following variables constructed, a 1 means that the worker chose the statement that rules should be followed (i.e., the first statement in the first set and the second statement for second set). Thus, receiving a 1 on these dummy variables means that a subject is stating that he feels rules should be followed.

The Density of Toxic Workers is a ratio that measures the degree of a worker's exposure to other toxic workers. We calculated this measure using the number of other workers on a worker's team who are ultimately terminated for being toxic, as described below, divided by the current number of workers on the worker's team. Thus, this measure changes over time.

For a subset of the dataset, we also have quantitative performance data. We have a measure of worker output speed and we have the length of time needed to complete one unit of output. The variable Performance Quantity Time FE is an individual worker fixed effect calculated while regressing the time-per-unit of a worker on a cubic function of time-on-the-job experience and controls for job position and the sub-firm where the worker is employed, while achieving a given performance result. We generally have multiple observations of a worker's performance over time; we refer to each observation of performance measurement as a performance result. In addition, we have a measure of worker output quality. This variable Performance Quality is obtained analogously to the variable Performance Quantity Time FE.

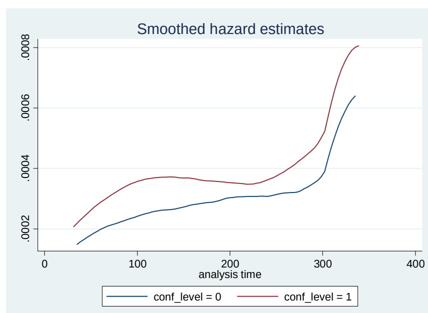
Finally, our dependent variable is an indicator variable based on whether the worker is terminated for toxic behavior. Toxic behavior is defined as involuntary termination due to an egregious violation of company policy. Examples include sexual harassment, workplace violence, falsifying documents, fraud, and general workplace misconduct. The mean of this variable is approximately 1% across all observations. However, in terms of per worker, the mean is 4.5% of all observations. In other words, roughly 1 in 20 workers is ultimately terminated as a toxic worker.

4 Antecedents of Toxicity

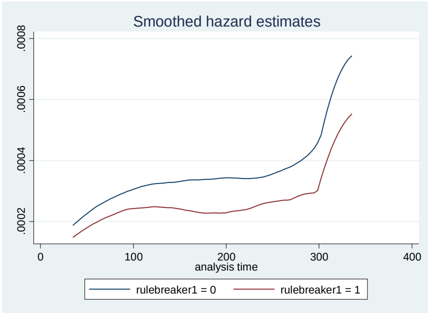
4.1 Hazard Functions

In this section, we show some graphical examples of the hazard rate as a function of time. In the next section, we will conduct a full analysis with controls. To provide sufficient observations we report the hazards for the first 365 days, as over 90% of a worker's tenure is under one year.

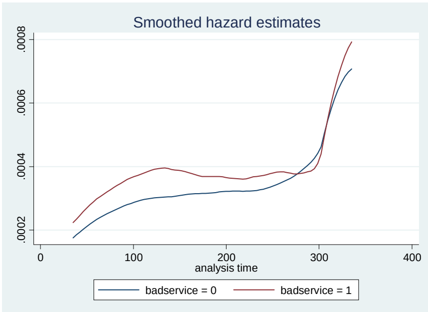
The first chart compares the difference in hazard rates of workers with an above-average (i.e., conf_level=1) and a below-average (i.e., conf_level=0) Confidence. Those who appear overconfident by overreporting their skill level before they start the job are more likely to be terminated for toxic behavior across all time.



Next we estimate the hazard rates of workers that state rules should never be broken (i.e., rulebreaker1=0) and those that suggest sometimes breaking rules is necessary (i.e., rulebreaker1=1). Interestingly, those that claim the rules should never be broken are more likely to be terminated for breaking the rules as a toxic worker across, at all times. This supports Hypothesis 3B, which suggests self-proclamation of rule following that is highly incentivized captures more Machiavellian than conscientious types. This possibility also raises a broader question of how to interpret survey responses when responses are highly incentivized versus when they are not; what might predict one type when not incentivized can possibly predict a very a different type when highly incentivized.



Finally, we compare the hazard rates of those that are Self-regarding to those that are not. As can be seen, having a Self-regarding orientation makes one more likely to be terminated for toxicity. If a Self-regarding worker is not terminated for toxicity in the first year, thereafter their chance of termination for toxicity is more similar to the average worker's chance. It could be that Self-regarding workers are those that also engage in toxicity are largely eliminated from the worker pool by this time.



Although these charts are suggestive, these estimated hazard rates need to be interpreted with care; they do not include potentially important controls (e.g., job position or supervisor). Further, we need to consider different factors simultaneously to determine if they are different predictors of toxic workers or if they are a measure of the same underlying force. For this analysis, we turn to our proportional-hazards regression model.

4.2 Regression Analysis of Toxic Terminations

4.2.1 Baseline

Table 2 reports the results of our baseline regression model.[9] For these regressions, we have a large enough sample to stratify by each sub-firm. This means that each sub-firm is allowed to have a unique baseline hazard function $h(t)$. That is, we estimate each regression while allowing for a different base hazard function $h(t)_j$ for each sub-firm j. As can be seen, greater Confidence results in a greater chance of being terminated for being a toxic worker. In particular, a one standard deviation in Confidence results in an approximate 15%[10] increase in the hazard. That is, conditional on a worker not yet having been terminated as a toxic worker, a one standard deviation increase in Confidence means that there is some 15% greater hazard of termination due to toxic behavior. Similarly, those that are Self-regarding, have more than a 22% increased hazard of toxic termination. If a worker reports that she believes rules are always made to be followed (as opposed to stating that it is sometimes necessary to break the rules to accomplish something), she has about 25% greater hazard of being terminated for actually breaking the rules. Finally, a worker that has a one standard deviation increase in exposure to toxic workers himself experiences a 46% increased hazard in being terminated for engaging in toxic behavior.

If we categorize the first four columns as measures of the person and the last two columns (including particular job type) as measures of the situation, we can state what fraction of a toxic worker's origin is attributable to the person versus the situation. In particular, using McFadden's pseudo R^2, we calculate that approximately 70% of the explanatory power of the model beyond a model that only contains an intercept comes from the person, and the balance (i.e., 30%) from the situation. If each variable in each column mattered equally, we would expect the person to explain 2/3 of outcome (i.e., 4 out of 6). In the next section, when we only analyze a worker's first placement, we find that using McFadden's pseudo R^2 that person explains almost 88% of the outcome. In short, at least in our setting, there is important explanatory power in simply knowing

[9] For all of our hazard models, we test the proportionality assumption (i.e., that the composite hazard rate is of the form $h(t)\,f(t)$) on the basis of the Schoenfeld residuals after fitting a given model (see Grambsch & Therneau (1994)). In all cases, our model is consistent.

[10] Recall that to convert estimates into a hazard ratio, simply raise e to the coefficient value. For example, a coefficient value of .5 results in $e^{.5} \simeq 1.65$. This means a one unit change in the regressor amounts to a 65% increase in the hazard ratio. Alternatively, a one standard deviation increase, when such standard deviation is .225, results in a roughly 14.6% increase in the hazard ratio.

the person, though the situation certainly matters too, and it seems to matter more over time.

Since we can only measure those cases of toxicity that are discovered and eliminated through termination, we could be only partially measuring outcomes. For the kinds of toxic behavior that we're studying (e.g., extreme levels such as sexual harassment and workplace violence) it seems likely that when such behavior is exhibited, discovery and termination will usually occur. However, in principle, it could be that cleverer people are better at somehow hiding their behaviors. To explore this possibility, we obtained the results of two cognitive tests that applicants take. These tests represents two questions, each quantitatively based with an objective correct answer. Table 3 reports the results of adding these two measures to our previous analysis. In doing so, we found that our previous estimates are very similar and that the cognitive test results do not explain the likelihood of toxic terminations. In fact, one test has a positive coefficient point estimate and the other test has a negative one, though neither is significant.

Ideally, we would like to conduct our analysis after randomly allocating all workers to workgroups and then observing their experiences and performance over time. Doing so would average out possible confounds, for which it is difficult to control. For example, perhaps a particular workgroup is better at detecting and eliminating toxic workers and that same group systematically hires more confident workers. However, based on discussions with company executives, conditional on a given sub-firm, a worker's first placement tends to be essentially random. Exactly where an employee is initially placed depends on a variety of factors outside the control of the worker and the workgroup in which she is placed. For example, the work flow of a particular operation, demand and supply shocks, and exactly when a worker turns up looking for a job are all factors determining to which group a new hire will be assigned. Further, a workgroup supervisor does not generally choose her group's new worker, so the supervisor does not observe the new worker's predicted job fit and other individual characteristic covariates that we use in our analysis. Instead, hiring is conducted through a centralized human resource center that is matching worker supply and demand. Thus, it is as if we have almost 60,000 workers randomly allocated across 11 firms on the dimensions of the personal characteristics; this allows us to view links between characteristics and toxicity and performance as causal rather than merely correlational, which is often the case. Nonetheless, although a worker's first placement is essentially random, a worker's second placement may not be random. Thus, for a robustness test, we now redo our above analysis, but only for a worker's first placement.

4.2.2 First Placement Only

Table 4 reports an analysis based only on an employee's first placement. The results are generally similar to the case in which all worker placements are included. Upon closer inspection, we see that the magnitude of the coefficient of the toxic worker density is about 20% smaller. One possible explanation for this is that the exposure to toxic workers has a cumulative effect: the same exposure sustained over an extended period of time has a greater adverse effect on a worker.

In principle, we can test statistically whether a placement is different from random. One common method includes comparing covariates across treatments, where treatments normally total two. However, in our setting, a "treatment" is the initial placement in each workgroup, which amounts to $2,882$ treatments, making a comparison cumbersome. Further, one can only consider relationships pair-by-pair. However, another common method that also allows the covariates to be interdependent is using a logit or probit model to predict treatment. Of course, this method only works when there are two different treatments; again, we have $2,882$ treatments. However, we can analyze a multinomial equivalent where each outcome is considered an unordered outcome of being placed in a given workgroup. We need sufficient observations in order to estimate how each covariate contributes to the likelihood of being placed in a particular workgroup. In the end, we can estimate how covariates predict 985 workgroup placements.

The following table reports the results of these regressions.

	Self-regarding	Confidence	Rules: Sometimes Break	Rules: Prefer Adventure
Number of Workgroups Significant at 5%	47	113	50	101
Fraction of Estimated Workgroups (985)	4.77%	11.47%	5.08%	10.25%
Fraction of All Workgroups (2,882)	1.63%	3.92%	1.73%	3.50%

We find that in 47 of the 985 cases (almost 5% of workgroups), Self-regarding predicts the workgroup in which a worker is placed. Confidence is significant over 11% of the time, whereas Rules covariates are significant at 5% and 10% of the time, respectively. If all placements were generated at random, we would expect each covariate to be significant at the 5% level, 5% of the time, on average. When we consider the full dataset we are using to estimate effects in our main analysis, covariates are only significant less than 3% of the time, on average. The reason we cannot estimate covariate effects on the entire dataset of workgroups is that generally there are too few observations for a particular workgroup, which also means we do not expect such workgroups to create statistical aberrations on their own.

We ran an additional robustness test to explicitly control for a employee's workgroup during her first placement. Specifically, we ran a linear panel model with workgroup fixed effects for a worker's first placement.[11] Here, we collapse the exposure to toxic workers as an average exposure over the placement, whereas before this was the current-period exposure. Results are reported in Table 5. The findings with this linear model are very similar in terms of significance and magnitude when compared with our hazard models, with the exception of the effect of Toxic exposure . In terms of magnitudes, a worker that is Self-regarding has an additional .9% chance of becoming a terminated toxic worker, which is an increase of 20% from the baseline toxic worker rate of 4.5%. A one standard deviation in Confidence results in a roughly 11% chance of becoming a terminated toxic worker. Those who state that rules should never be broken are 20% more likely to be terminated for toxic behavior. A one standard deviation increase in exposure to other toxic workers induces a roughly 98% increased chance of a worker becoming a toxic worker. Finally, In short, these effects are consistent with those found with our previous models.

5 Consequences of Toxic Workers

Now that we have identified some antecedents of toxicity, we will explore some of the consequences. As already seen in the past section, Toxic workers seem to induce others to be toxic. In this section, we further explore the consequences of a toxic worker in terms of direct performance effects. For a subset of the data, we have performance data on the workers. For this group, as discussed in section 3.2, we have a measure of each employee's time to produce one unit of quantity and a measure of their quality of work. We then use this data to calculate a worker-specific fixed effect of each of these measures, which we refer to as Performance Quantity Time FE and Performance Quality FE, respectively.

Looking simply at mean FE values, we find that the average Performance Quantity Time FE

[11] Note that we do not control for position type in these specifications. A particular workgroup typically consists of the same set of position types, and thus the variance matrix naturally becomes unusable when we do attempt to control for position type simultaneously with workgroup.

is less for those ultimately fired for toxicity than those that are not toxic (t-test with unequal variance yields a p-value= .0376). That is, toxic workers are more productive than those that are not ultimately terminated for toxicity. When we consider Performance Quality FE, we find that toxic workers produce lesser quality work than non-toxic workers; however, results do not quite reach conventional levels of statistical significance (p-value= .1233). Of course these are simply means and we should consider analysis with controls. In particular, we should relate the productivity of workers and whether or not that worker is terminated for toxic behavior while controlling for the previous factors that are important for identifying toxicity.

Table 6 reports the results of introducing these additional measures to our original analysis reported in Table 2. The other variables of interest previously studied are qualitatively the same as in table 2, although the levels of significance are diminished for this considerably smaller sample size.

Similar to our findings of comparing the means, those who are terminated for engaging in toxic behavior are more productive than non-toxic employees; equivalently, those who are slower (i.e., large values of Performance Quantity Time FE) are *less* likely to be toxic. In terms of magnitude, a one standard deviation in time per unit of production results in a 56% reduction in the hazard of becoming a toxic worker. However, those workers with poorer *quality* performance are more likely to be toxic. Here, a one standard deviation increase in the quality of production results in a 27% decrease in the hazard.

It might seem that toxic workers are simply those that trade work quality for speed and those workers that produce higher quality must also be slower workers. However, this is not the case. Clearly, there is a natural tradeoff between speed and quality of work. Yet, we found almost 50% more workers that produce high quality work quickly (32.4% of workers) than those that produce low quality work quickly (23% of workers). Thus, although toxic workers are quicker than the average worker, they are not necessarily more productive in a quality-adjusted sense. In the long run, these kinds of workers are not likely to improve overall organizational performance.

5.1 Finding Superstars vs. Avoiding Toxic Workers

A next logical step is comparing the benefits and costs of finding a superstar worker versus avoiding a toxic worker. With performance data we can also compare the strategy of finding a "superstar" worker versus avoiding a toxic worker. As discussed in the introduction, many firms, as well as the extant literature are focused on finding and keeping the next star performer, whereas much less attention is devoted to avoiding toxic workers. Given a firm with limited resources, which strategy is more fruitful? Although we certainly cannot answer this question for all possible settings, we can assess this trade-off for our own setting.

To generate a straightforward comparison of the value of each of these focuses, we quantify the value of a star performer by identifying the cost savings from her increased output level. That is, a superstar is a worker that adds so much value that without her a firm would have to hire additional workers (or pay for additional hours from existing workers) to achieve the same level of output as when they have that single superstar worker. In the table below, the column "Hire a Superstar" reports the cost saving based on the top 1%, 5%, 10%, and 25% performers. We calculate the percent in increased performance for each of these performance levels and multiply it by the average worker salary, based on company records. This is an upper bound of the cost-savings from hiring a superstar since increased performance is often accompanied by increased wages. Note that productivity spillovers of a superstar worker are not found to be material in our setting. Superstars can have both positive (e.g., Azoulay et al. (2010)) and negative (e.g., Brown (2011)) spillovers; thus, these forces seem to approximately net out in our setting.

For comparison, we report in the "Avoid a Toxic Worker" column the induced turnover cost of a toxic worker, based on company figures. Induced turnover cost captures the expense of replacing additional workers lost in response to the presence of a toxic worker on a team. The total estimated cost is $12,489 and does not include other potential costs, such as litigation, regulatory penalty, and reduced employee morale. Also not included are the secondary costs of turnover that come from a new worker's learning curve: a time of lower productivity precedes a return to higher productivity. Thus, this estimate is likely a lower bound on the average cost of a toxic worker, at least for this empirical setting.

	Cost-savings	
Superstar Rank	Hire a Superstar	Avoid a Toxic Worker
top 25%	$ 1,951	$ 12,489
top 10%	$ 3,251	$ 12,489
top 5%	$ 3,875	$ 12,489
top 1%	$ 5,303	$ 12,489

In comparing the two costs, even if a firm could replace an average worker with one who performs in the top 1%, it would still be better off by replacing a toxic worker with an average worker by more than two-to-one.

That is, avoiding a toxic worker (or converting them to an average worker) provides more benefit than finding and retaining a superstar. Assuming that it is no more costly to avoid a toxic worker (or replace them with an average worker) than it is to find, hire, and retain a superstar, it is also more profitable to do the former over the latter. Of course, this differential between the superstar and toxic worker might not be as drastic in other settings. Nonetheless, finding a top 1% worker can be both difficult and costly. Further, sometimes "stars" that hiring managers discover via another firm are not able to transfer their same elevated level of productivity to their next employer (Groysberg (2012)).

5.2 Multi-dimensional Hiring: Performance and Toxicity

Of course, hiring decisions do not simply occur in the "tails" of the distribution. Next we consider hiring for productivity and toxicity simultaneously, across the entire distribution. That is, we assume that multi-dimensional hiring should not only include hiring a worker based on multiple worker characteristics but it should also include hiring a worker based on multiple expected outcomes. Specifically, when hiring a worker based on a collection of traits and experiences that lead to higher productivity, we also consider how that same set of characteristics might lead to toxic behavior. To explore this idea, we analyze our productivity data as a function of the same individual characteristics used in Section (4.2) to predict the hazard of toxicity.

Table 7 reports the results of regressing performance outcomes on these individual characteristics only for a worker's first placement, and controlling for supervisors. As previously discussed, since workers are exogenously placed in workgroups, we can identify causal consequences of characteristics on performance. Here, we find that though Self-regarding workers are no different in terms of productivity (i.e., the speed of their work), they are more likely to produce lower quality work: coefficient estimates are negative at the 10% level. Since these workers are also more likely to be terminated for toxic behavior, there is no apparent tradeoff when choosing no to hire Self-regarding workers.

Similarly, consider hiring workers that claim sometimes the rules need to be broken. As found earlier in our analysis, such workers are less likely to be terminated for toxicity. Although these workers are no different in terms of productivity, they tend to produce higher quality work, as evidenced by the positive estimate coefficient in column (6), which is significant at the 10% level. In short, there is no hiring tradeoff for this characteristic either.

In contrast to the two previous characteristics, there is an tradeoff in hiring more confident workers. From before, we know that more confident workers are more likely to be terminated for toxicity. However, as shown in Table 7, these workers are also more likely to be productive. In fact, the firms in the sample to tend to hire more confident workers more often. That is, they are hiring unidimensionally on this factor of confidence, without considering that such characteristic also predicts the likelihood of a toxic worker. In this case, following the approach in the previous section, we can estimate the net effects of profit from choosing to continue to hire on the characteristic of confidence. In particular, a one standard deviation increase in Confidence results in approximately \$122 of expected saved wages due to greater productivity of that person. However, the same one standard deviation increase in Confidence yields an expected $\$1,327$ increased cost from induced turnover from increased likelihood of toxicity. Thus, on net, a company is still over $\$1,000$ better off in terms of expected profit per worker if they refrain from hiring Self-regarding workers. This is the danger of making hiring decisions for unidimensional reasons, such as productivity.

6 Discussion: Strategic Human Resource Management

Based on our analysis, we have a variety of takeaways for managers. From our study, it seems clear that toxic workers originate both as a function of preexisting characteristics and of the environment in which they work. In particular, we found consistent evidence that those who seem overconfident in their abilities, who are self-regarding, and who claim rules should be followed, are more likely to become toxic workers and break company and legal rules. Thus, one strategy for managers is to screen potential workers for these traits to reduce the chance of hiring toxic workers. However, we also found that toxic workers are more productive, at least in terms of the quantity of output. This could explain why toxic workers are selected and are able to remain in an organization for as long as they do. For example, an investment bank with a rogue trader who is making the firm millions in profits might be tempted to look the other way when the trader is found to be overstepping the legal boundaries. In fact, Pierce and Snyder (2013) find that unethical workers enjoy longer tenures.

This performance finding suggests that toxic workers are similar to what Jack Welch described as "Type 4" workers—those who deliver on the numbers but do not have the right values. Welch claimed that while difficult to do, it was critical to remove such workers: "People are removed for having the wrong values...we don't even talk about the numbers" (Bartlett and Wozny (2005)).[12] We find evidence that such a policy—one that removes the "big shots" and "tyrants" seems to be one that would lead to more productive organizations in general, despite terminating such a productive worker.[13] Similarly, Delong and Vijayaraghavan (2003) argue that the top performers are not always the best workers to pursue over even an average worker, as the former can also create organizational issues, including reckless behavior. In recognizing this tradeoff of productive workers that might be toxic, we were able to directly explore some of the characteristics that lead to better performance and toxicity. In particular, we identified confidence as predicting workers that are both highly productive and toxic. However, when considering these outcomes in tandem, the net consequence in terms of profit is still net negative when hiring such workers. Thus, when considering simultaneously the dimensions of productivity *and* toxicity, certain hires no longer make sense, even setting aside ethical concerns and instead relying solely on profit maximization. Thus, an important take away is that managers should consider toxic and productivity outcomes together rather than relying on productivity alone as the criterion of a good hire. As we found, doing so could allow a manager to avoid a worker who would have caused net profit losses, a worker she would have otherwise hired if she considered productivity outcomes alone. An even more general take away is that managers should hire multi-dimensionally in terms of *outcomes.*

Although we do find certain preexisting traits that predict toxic workers, this does not mean that those traits were always present in the worker. Though it is beyond the scope of this paper, it would be interesting to learn to what extent work-life experiences breed the preexisting traits that we have found to lead to toxic workers. It would be very valuable to discover what firms can currently do to limit the chances of converting a "normal" worker to a future toxic worker.

We did find that a worker's environment also substantially influenced her propensity to become a toxic worker. We documented that holding a particular type of position, as well as exposure to other toxic workers, negatively influenced the likelihood of one becoming toxic. Hence, this suggests

[12] Havard Business School Publishing case # 9-399-150.
[13] We thank Tarun Khanna for this Jack Welch example.

that managing toxic workers is not simply a matter of screening them out of the firm, but also of minding the work environment.

7 Conclusion

A good or bad hiring decision is multidimensional (Lazear & Oyer (2007) and Hermalin (2013)). We have identified several individual and situational factors that lead to a worker engaging in objective toxic behavior. Knowledge of these factors can be used to avoid and better manage for toxic workers. However, we also found the need to hire based on multiple dimensions of expected outcomes: We found that adding the dimension of toxicity can help improve performance by means of avoiding the wrong kind of highly productive workers that would have been thought a preferred hire had we not considered toxicity.

We have also discovered some important effects of toxic workers. However, there are surely additional traits that could be used to identify toxic workers. Similarly, it would be helpful to know which other environmental factors nudge an otherwise normal worker towards becoming a toxic worker and possibly creating the preexisting workplace conditions that lead to toxic behavior. Future research can shed light on these questions. This latter focus seems particularly important, because to the extent that we can reduce a worker's likelihood of becoming toxic, we are helping not only the firm, but the worker himself, those around him, and the potential firms where that employee may work in the future. Since we found some evidence that a toxic worker can have more impact on performance than a "superstar," it may be that spending more time limiting negative impacts on an organization might improve everyone's outcome to a greater extent than only focusing on increasing positive impacts. We have taken a step in exploring this notion and hope that we witness future progress in this area.

References

[1] Andreoni, J. and J. Miller (2002) "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70 (2), 737-753.

[2] Azoulay, P., Graff Zivin, J., & Wang, J. (2010). Superstar Extinction. The Quarterly Journal of Economics, 125(2), 549-589.

[3] Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. Review of General Psychology, 5(4), 323.

[4] Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. American Economic Review, 67-72.

[5] Brown, J. (2011). Quitters never win: The (adverse) incentive effects of competing with superstars. Journal of Political Economy, 119(5), 982-1013.

[6] Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge university press.

[7] Campbell, W., Goodie, A. S., & Foster, J. D. (2004). Narcissism, confidence, and risk attitude. Journal of Behavioral Decision Making, 17(4), 297-311.

[8] De Cremer, D., & Van Lange, P. A. (2001). Why prosocials exhibit greater cooperation than proselfs: The roles of social responsibility and reciprocity.European Journal of Personality, 15, 5-18.

[9] Delong, T. J., & Vijayaraghavan, V. (2003). Let's hear it for B players. Harvard Business Review, 81(6), 96-102.

[10] Derfler-Rozin, R., Moore, C., & Staats BR. (2015). Enhancing ethical behavior through task variety. London Business School working paper.

[11] Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. Current Directions in Psychological Science, 12(3), 83-87.

[12] Ermongkonchai, P. (2010). Understanding Reasons for Employee Unethical Conduct in Thai Organizations: A Qualitative Inquiry, Contemporary Management Research, 6(2), 125-140.

[13] Estes, Z., & Adelman, J. (2008). Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006). Emotion, 8(4), 441-444.

[14] Fisman, R., S. Kariv, and D. Markovits (2007) "Individual Preferences for Giving," *American Economic Review*, 97 (5), 1858–1876.

[15] Folmer, C. P. R., & De Cremer, D. (2012). Bad for me or bad for us? Interpersonal orientations and the impact of losses on unethical behavior. Personality and Social Psychology Bulletin, 38(6), 760-771.

[16] Ford, R. C., & Richardson, W. D. (1994). Ethical decision making: A review of the empirical literature. Journal of Business Ethics, 13(3), 205-221.

[17] Frank, D. H., & Obloj, T. (2014). Firm-specific human capital, organizational incentives, and agency costs: Evidence from retail banking. Strategic Management Journal, 35(9), 1279-1301.

[18] Gibbons, R., & Roberts, J. (Eds.). (2013). The Handbook of Organizational Economics. Princeton University Press.

[19] Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. Psychological review, 98(4), 506.

[20] Gino, F., & Ariely, D. (2012). The dark side of creativity: original thinkers can be more dishonest. Journal of Personality and Social Psychology, 102(3), 445.

[21] Gino, F., & Bazerman, M. H. (2009). When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. Journal of Experimental Social Psychology, 45(4), 708-719.

[22] Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. Biometrika, 81(3), 515-526.

[23] Greenberg, J. (2002). Who stole the money, and when? Individual and situational determinants of employee theft. Organizational Behavior and Human Decision Processes, 89(1), 985-1003.

[24] Greve, H. R., Palmer, D., & Pozner, J. E. (2010). Organizations gone wild: The causes, processes, and consequences of organizational misconduct. The Academy of Management Annals, 4(1), 53-107.

[25] Groysberg, B. (2012). Chasing stars: The myth of talent and the portability of performance. Princeton University Press.

[26] Hegarty, W. H., & Sims, H. P. (1978). Some determinants of unethical decision behavior: An experiment. journal of Applied Psychology, 63(4), 451

[27] Hermalin, B. (2013). Leadership and corporate culture. Handbook of Organizational Economics, 432-478.

[28] Larkin, I. (2014). The cost of high-powered incentives: Employee gaming in enterprise software sales. Journal of Labor Economics, 32(2), 199-227.

[29] Lazear, E. P., & Oyer, P. (2007). Personnel economics (No. w13480). National Bureau of Economic Research.

[30] Lee, K., Ashton, M. C., & Shin, K. H. (2005). Personality Correlates of Workplace Anti-Social Behavior. Applied Psychology, 54(1), 81-98.

[31] Litzky, B. E., Eddleston, K. A., & Kidder, D. L. (2006). The good, the bad, and the misguided: How managers inadvertently encourage deviant behaviors. The Academy of Management Perspectives, 20(1), 91-103.

[32] Loe, T. W., Ferrell, L., & Mansfield, P. (2000). A review of empirical studies assessing ethical decision making in business. Journal of Business Ethics, 25(3), 185-204.

[33] MacLean, T. L., & Behnam, M. (2010). The dangers of decoupling: The relationship between compliance programs, legitimacy perceptions, and institutionalized misconduct. Academy of Management Journal, 53(6), 1499-1520.

[34] Mayer, D. M., Kuenzi, M., Greenbaum, R., Bardes, M., & Salvador, R. B. (2009). How low does ethical leadership flow? Test of a trickle-down model. Organizational Behavior and Human Decision Processes, 108(1), 1-13.

[35] Moore, C., Detert, J. R., Klebe Treviño, L., Baker, V. L., & Mayer, D. M. (2012). Why employees do bad things: Moral disengagement and unethical organizational behavior. Personnel Psychology, 65(1), 1-48.

[36] Moore, C., & Gino, F. (2013). Ethically adrift: How others pull our moral compass from true North, and how we can fix it. Research in Organizational Behavior, 33, 53-77

[37] Moore, C., & Gino, F. (2015). Approach, Ability, Aftermath: A Psychological Process Framework of Unethical Behavior at Work. The Academy of Management Annals, 9(1), 235-289.

[38] Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. Psychological review, 115(2), 502.

[39] Mount, M., Ilies, R., & Johnson, E. (2006). Relationship of personality traits and counterproductive work behaviors: The mediating effects of job satisfaction. Personnel psychology, 59(3), 591-622.

[40] Oberholzer-Gee, F. & Wulf, J. (2012). Incentives-based misconduct: Earnings Management from the Bottom Up: An Analysis of Managerial Incentives Below the CEO. HBS Working Paper 12-056.

[41] O'Boyle Jr, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. Journal of Applied Psychology, 97(3), 557.

[42] Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. Management Science, 58(6), 1122-1140.

[43] Pierce, L., & Balasubramanian, P. (2015). Behavioral field evidence on psychological and social factors in dishonesty and misconduct. Current Opinion in Psychology, 6, 70-76.

[44] Pierce, L., & Snyder, J. (2008). Ethical spillovers in firms: Evidence from vehicle emissions testing. Management Science, 54(11), 1891-1903.

[45] Pierce, L., & Snyder, J. A. (2013). Unethical demand and employee turnover. Journal of Business Ethics, 1-17.

[46] Roberts, J., & G. Saloner. (2013). Strategy and Organization. Handbook of Organizational Economics, 799-854.

[47] Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. Academy of management journal, 38(2), 555-572.

[48] Robinson, S. L., & O'Leary-Kelly, A. M. (1998). Monkey see, monkey do: The influence of work groups on the antisocial behavior of employees. Academy of Management Journal, 41(6), 658-672.

[49] Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. International Journal of Selection and Assessment, 10, 117-125.

[50] Sauermann, H., & Cohen, W. M. (2010). What makes them tick? Employee motives and firm innovation. Management Science, 56(12), 2134-2153.

[51] Tenbrunsel, A. E., & Smith-Crowe, K. (2008). 13 Ethical Decision Making: Where We've Been and Where We're Going. The Academy of Management Annals, 2(1), 545-607.

[52] Trevino, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. Academy of management Review, 11(3), 601-617.

[53] Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty J Risk Uncertainty, 5(4), 297-323

[54] Van Lange, P. A., & Kuhlman, D. M. (1994). Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect. Journal of Personality and Social Psychology, 67(1), 126.

[55] Vardi, Y. (2001). The effects of organizational and ethical climates on misconduct at work. Journal of Business Ethics, 29(4), 325-337.

[56] Vaughan, D. (1999). The dark side of organizations: Mistake, misconduct, and disaster. Annual review of sociology, 271-305.

[57] Weaver, G. R., & Trevino, L. K. (1999). Compliance and Values Oriented Ethics Programs. Business Ethics Quarterly, 9(2), 315-335.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Self-regarding	248370	0.15	0.35	0.00	1.00
Confidence	248370	-0.01	0.23	-0.24	0.92
Rules: Sometimes Break Them	248370	0.14	0.34	0.00	1.00
Rules: Prefer Adventure	248370	0.44	0.50	0.00	1.00
Density of Toxic Workers	248370	0.04	0.04	0.00	0.80
Performance Quantity Time FE	62618	-32.77	213.11	-462.94	1488.31
Performance Quality FE	20089	-0.05	0.13	-0.91	0.23
Terminated for Toxic Behavior	248370	0.01	0.10	0	1

Table 2: Terminations as a Function of Worker Type and Environment

(All Placements)

Outcome: Terminated Toxic Worker

Worker and Environment	(1)	(2)	(3)	(4)	(5)	(6)
Self-regarding	0.2808***	0.2090***	0.2077***	0.2047***	0.2030***	0.2023***
	(5.30)	(3.81)	(3.78)	(3.71)	(3.69)	(3.68)
Confidence		0.5215***	0.5206***	0.5177***	0.5033***	0.5034***
		(6.42)	(6.41)	(6.36)	(6.18)	(6.17)
Rules: Sometimes Break Them			-0.2373***	-0.2284***	-0.2247***	-0.2272***
			(-3.75)	(-3.55)	(-3.49)	(-3.53)
Rules: Prefer Adventure				-0.0274	-0.0193	-0.0184
				(-0.67)	(-0.47)	(-0.45)
Density of Toxic Workers					2.5581***	2.5182***
					(11.74)	(11.49)
Position Controls	No	No	No	No	No	Yes
Log Likelihood	-18446.4880	-18427.3077	-18419.9216	-18419.6990	-18372.5281	-18368.2087
N	246599	246599	246599	246599	246599	246599

Cox proportional hazard model used for estimation

Non parametric hazard functions estimated at the sub-firm level

Z scores reported in parentheses based on standard errors clustered at the worker level

* p<0.10, ** p<0.05, *** p<.01

Table 3: Cognitive Scores and Terminations

(All Placements)

Outcome: Terminated Toxic Worker

Worker and Environment	(1)	(2)	(3)	(4)	(5)	(6)
Cognitive Test I Correct	0.0051	0.0164	0.0200	0.0201	0.0304	0.0258
	(0.10)	(0.32)	(0.39)	(0.39)	(0.59)	(0.50)
Cognitive Test II Correct	-0.1092*	-0.0777	-0.0708	-0.0693	-0.0581	-0.0599
	(-1.76)	(-1.25)	(-1.14)	(-1.11)	(-0.93)	(-0.96)
Self-regarding	0.2730***	0.2051***	0.2043***	0.2015***	0.2004***	0.1996***
	(5.12)	(3.73)	(3.71)	(3.65)	(3.64)	(3.63)
Confidence		0.5126***	0.5129***	0.5104***	0.4982***	0.4978***
		(6.27)	(6.28)	(6.23)	(6.08)	(6.06)
Rules: Sometimes Break Them			-0.2355***	-0.2273***	-0.2241***	-0.2264***
			(-3.72)	(-3.53)	(-3.48)	(-3.52)
Rules: Prefer Adventure				-0.0257	-0.0180	-0.0169
				(-0.63)	(-0.44)	(-0.41)
Density of Toxic Workers					2.5560***	2.5151***
					(11.71)	(11.45)
Position Controls	No	No	No	No	No	Yes
Log Likelihood	-18444.8609	-18426.5004	-18419.2375	-18419.0419	-18371.9870	-18367.6758
N	246599	246599	246599	246599	246599	246599

Cox proportional hazard model used for estimation

Non parametric hazard functions estimated at the sub-firm level

Z scores reported in parentheses based on standard errors clustered at the worker level

* $p<0.10$, ** $p<0.05$, *** $p<.01$

Table 4: Terminations as a Function of Worker Type and Environment

(First Placements Only)

Outcome: Terminated Toxic Worker

Worker and Environment	(1)	(2)	(3)	(4)	(5)	(6)
Self-regarding	0.3279***	0.2617***	0.2608***	0.2638***	0.2595***	0.2597***
	(5.66)	(4.35)	(4.33)	(4.37)	(4.30)	(4.30)
Confidence		0.4574***	0.4573***	0.4603***	0.4545***	0.4513***
		(4.96)	(4.96)	(4.98)	(4.91)	(4.86)
Rules: Sometimes Break Them			-0.2058***	-0.2146***	-0.2111***	-0.2097***
			(-2.95)	(-3.02)	(-2.96)	(-2.95)
Rules: Prefer Adventure				0.0272	0.0311	0.0300
				(0.59)	(0.67)	(0.65)
Density of Toxic Workers					2.1341***	2.1409***
					(7.47)	(7.48)
Position Controls	No	No	No	No	No	Yes
Log Likelihood	-13661.0461	-13649.4261	-13644.9028	-13644.7302	-13627.5793	-13626.9476
N	190178	190178	190178	190178	190178	190178

Cox proportional hazard model used for estimation

Non parametric hazard functions estimated at the sub-firm level

Z scores reported in parentheses based on standard errors clustered at the worker level

* p<0.10, ** p<0.05, *** p<.01

Table 5: Linear Model of Terminations with Workgroup Fixed Effects

(First Placements Only)

Outcome: Terminated Toxic Worker

Worker and Environment	(1)	(2)	(3)	(4)	(5)
Self-regarding	0.0128***	0.0093***	0.0093***	0.0097***	0.0091***
	(4.04)	(2.86)	(2.87)	(2.96)	(2.82)
Confidence		0.0243***	0.0243***	0.0246***	0.0208***
		(5.07)	(5.06)	(5.11)	(4.39)
Rules: Sometimes Break Them			-0.0074***	-0.0083***	-0.0076***
			(-2.74)	(-3.01)	(-2.79)
Rules: Prefer Adventure				0.0028	0.0025
				(1.33)	(1.23)
Avg Density of Toxic Workers					1.1078***
					(20.70)
R Squared	0.044	0.044	0.044	0.045	0.070
Adjusted R Squared	0.022	0.023	0.023	0.023	0.049
N	44710	44710	44710	44710	44710

t statistics reported in parentheses based on standard errors clustered at the workgroup level

* $p<0.10$, ** $p<0.05$, *** $p<.01$

Table 6: Terminations with Worker Performance

(All Placements)

Outcome: Terminated Toxic Worker

Worker and Environment	(1)	(2)	(3)
Performance Quantity Time FE	-0.0036***		-0.0039***
	(-6.79)		(-5.33)
Performance Quality FE		-2.1925***	-2.4419***
		(-4.90)	(-5.28)
Self-regarding	0.0383	-0.0620	-0.0853
	(0.41)	(-0.41)	(-0.55)
Confidence	0.4069***	0.4597**	0.4686**
	(2.93)	(2.50)	(2.54)
Rules: Sometimes Break Them	-0.0215	-0.0226	0.0192
	(-0.20)	(-0.14)	(0.12)
Rules: Prefer Adventure	-0.0823	-0.2627***	-0.2449**
	(-1.17)	(-2.64)	(-2.44)
Density of Toxic Workers	1.5359***	1.7591***	1.5971***
	(5.21)	(5.83)	(5.15)
Position Controls	Yes	Yes	yes
Log Likelihood	-5859.5624	-3233.1108	-3165.6394
N	62419	19983	19751

Cox proportional hazard model used for estimation

Non parametric hazard functions estimated at the sub-firm level

Z scores reported in parentheses are based

on standard errors clustered at the worker level

* p<0.10, ** p<0.05, *** p<.01

Table 7: Performance as a Function of Personal Characteristics

(First Placements Only)

Outcome: Performance

	Performance Metric					
	Speed	Quality	Speed	Quality	Speed	Quality
Individual Characteristics	(1)	(2)	(3)	(4)	(5)	(6)
Self-regarding	-2.6291	-0.0181**	-1.5154	-0.0148*	-1.7836	-0.0139*
	(-0.65)	(-2.25)	(-0.37)	(-1.82)	(-0.43)	(-1.71)
Confidence			-12.3441**	-0.0393***	-12.5026**	-0.0391***
			(-2.29)	(-3.47)	(-2.32)	(-3.45)
Rules: Sometimes Break Them					1.2505	0.0143*
					(0.32)	(1.85)
Rules: Prefer Adventure					-2.2509	0.0038
					(-0.90)	(0.82)
R Squared	0.903	0.126	0.903	0.128	0.903	0.129
Adjusted R Squared	0.899	0.088	0.899	0.090	0.899	0.090
N	6226	5485	6226	5485	6226	5485

t statistics reported in parentheses based on standard errors clustered at the workgroup level

* p<0.10, ** p<0.05, *** p<.01