

# **Reducing Organizational Rule Breaking through Task Variety: How Task Design Supports Deliberative Thinking**

Rellie Derfler-Rozin  
Robert H. Smith School of Business  
University of Maryland, College Park  
4510 Van Munching Hall  
Tel: +1 301 405 9495  
[rderfler@rhsmith.umd.edu](mailto:rderfler@rhsmith.umd.edu)

Celia Moore  
Department of Management and Technology  
Bocconi University  
Via Roentgen, 1  
Milan 20136 Italy  
Phone: +39 02 58365746  
[celia.moore@unibocconi.it](mailto:celia.moore@unibocconi.it)

Bradley R. Staats  
Kenan-Flagler Business School  
University of North Carolina-Chapel Hill  
Campus Box 3490, McColl Building  
Chapel Hill, NC 27599-3490  
Tel: +1 919 962 7343  
[bstaats@unc.edu](mailto:bstaats@unc.edu)

IN PRESS, *Organization Science*

The authors wish to thank Francesca Gino, Madan Pillutla, Lakshmi Ramarajan, Subra Tangirala, Ann Tenbrunsel, and Kristen Smith-Crowe for helpful comments on earlier drafts of this paper. We also thank the Senior Editor and three reviewers for their constructive comments throughout the revision process, as well as the support of management at our field site. In addition, we wish to thank Zareen Choudhury, Monika Heller, Nasima Kaker, and Ovul Sezer for their assistance with data collection.

# **Reducing Organizational Rule Breaking through Task Variety: How Task Design Supports Deliberative Thinking**

## **Abstract**

In this paper we argue that task design affects rule breaking in the workplace. Specifically, we propose that task variety activates deliberative (Type 2) processes as opposed to automatic/intuitive (Type 1) processes, which, in turn, helps prevent individuals from breaking rules in order to serve their own hedonic self-interest. We use data from the home loan application processing operations of a Japanese bank to establish the phenomenon in the field. We document that increased task variety at a daily level is associated with lower levels of rule breaking in the form of violating corporate break time policies (Study 1). We further explore the relationship between task variety and rule breaking in three lab experiments, using different operationalizations of rule breaking (Studies 2, 3a and 3b) and provide direct evidence for the mediating effect of deliberative thinking in this relationship (Studies 3a and 3b). We discuss implications for rule compliance in organizations, behavioral ethics and work design.

**Keywords:** rule breaking, task variety, deliberative thinking, Type 1 and Type 2 processes, job design

## Introduction

Formal rules and policies that direct employees' behavior in organizations are critical for organizations to function smoothly on a daily basis (Tyler and Blader 2005). Rules help ensure that employees do their work in an appropriate manner, behave suitably with each other, treat customers well, and stay within the bounds of the law. However, ample evidence documents that employee rule breaking is rampant. Employees steal (Greenberg 1997), loaf (Kidwell and Bennett 1993; Roy 1959), harass others (Fitzgerald et al. 1997), and take safety shortcuts (Hofmann and Stetzer 1996), to an alarming degree. Often (though not always), employees violate these rules to serve their own hedonic self-interest to the detriment of others, which is why traditional perspectives on employee deviance consider organizational rule violations to be unethical (Griffin and Lopez 2005; Robinson and Bennett 1995, 1997). Given how widespread rule breaking is, and the tremendous costs it imposes on society (Martin et al. 2013), it is important to examine what tools organizations may have at their disposal to reduce it.

Violations of rules against deviant behavior are traditionally understood as stemming from failures of character (Hogan 1973), motivation (Vardi and Weiner 1996), or incentives (Harris and Bromiley 2007; Tenbrunsel 1998). However, recent work also reveals that this type of behavior can be a consequence of unchecked automatic or intuitive inclinations (Knoch et al. 2006; Moore and Loewenstein 2004; Shalvi et al. 2012). Barring strong internal (Chugh et al. 2014; Reynolds et al. 2010) or external (Desai and Kouchaki 2016; Haley and Fessler 2005) cues to behave otherwise, we will do what is in our own immediate self-interest, to satisfy hedonic needs (Lindenberg and Steg 2007), even if doing so entails violating explicit rules we have been instructed to follow (Babcock et al. 1995; Epley and Caruso 2004; Kunda 1990). Moreover, when making these hedonic choices, we will fail to notice their detrimental implications for others (Chugh et al. 2005; Tenbrunsel and Messick 2004), and experience them as legitimate and fair (Batson et al. 1997; Diekmann et al. 1997).

This natural human tendency is dangerous, as many contexts offer opportunities to break rules in order to serve our immediate self-interest. For example, hospitality workers can earn more tips if they contravene policies to provide everyone high quality service (Salaz 2002), nurses can finish treatments

more quickly by not complying with hand washing regulations (Dai et al. 2015; Grant and Hofmann 2011), workers can have a more relaxing day if they contravene corporate break time policies (Roy 1959), and students can earn higher grades when they violate instructions not to collaborate on take-home exams (Pérez-Peña 2013). If our instinctive response is often to break rules when complying contradicts our immediate (hedonic) self-interest, then organizations need to find ways to turn off this “automatic pilot”.

Our paper addresses the conundrum of how to reduce workplace rule violations by exploring how task design equips us to resist lapsing into self-serving behavior that contravenes formal policies. Linking work on the benefits of deliberative thinking (Kahneman 2003, 2011) with task design literature (Grant et al. 2010; Hackman and Oldham 1976, 1980; Humphrey et al. 2007), we explore how we can design tasks to trigger deliberative processing, in turn reducing tendencies to break rules to satisfy hedonic self-interested goals. We argue that organizing work tasks in a highly varied way activates our deliberative/reflective cognitive capacities (“Type 2” processes), instead of leaving one to rely on more intuitive and automatic responses (“Type 1” processes, see Evans and Stanovich 2013; Stanovich and West 2000). We propose that this deliberative mindset supports rule compliance, especially when rule violations serve individuals’ immediate, hedonic self-interest to the detriment of others. We posit that work that constantly shifts among different subtasks forces individuals to remain cognitively alert, which helps individuals override the impulse to respond to any given task in an automatic, hedonic way (Baumeister et al. 1994; Shallice and Burgess 1993).

Our work contributes to literature on organizational rule compliance (Martin et al. 2013; Scott et al. 2009; Tyler and Blader 2005), behavioral ethics in organizational settings (Moore and Gino 2015; Treviño et al. 2006), and job design (Grant et al. 2010; Humphrey et al. 2007; Ilgen and Hollenbeck 1991). Theoretically, we explore the role of Type 1 and Type 2 processes in ethical behavior (Bazerman and Tenbrunsel 2011), and provide evidence both that task design choices can activate Type 2 processes, and that this more deliberative functioning dampens individuals’ tendencies to break rules in order to meet their immediate hedonic needs. We also respond to Grant and his colleagues’ call (2010) for more work on the ethical implications of job design, by showing that small changes to how tasks are sequenced

(one aspect of job design) trigger positive benefits that have not been explored to date (enhanced deliberative thinking), ultimately decreasing organizational rule breaking. These outcomes can accrue over periods as short as a few minutes (in the case of our experimental results) or over a day (in our field data). Ultimately, our work offers a new tool in the organizational battle against rule breaking, as we provide a workplace intervention—sequential task variety—that may support more desirable and compliant behavior when hedonic self-serving options are an easy and available option.

### **The role of deliberative thinking in decreasing rule breaking**

In a complex world, people cannot attend to all stimuli nor process each piece of information or environmental cue they encounter in depth. To explain how humans function in the face of this limited capacity, Herbert Simon (1982) introduced the idea of bounded rationality. He explained that individuals cannot optimize all decisions in every situation, as traditional economics would predict, because our rational capacity is bounded by the situation and by human computational power. While Simon's original framework centered around basic perceptual-cognitive 'failures', his ideas have since been applied to ethical behavior more directly (Chugh et al. 2005). Chugh and her colleagues coined the term "bounded ethicality" to describe how individuals act unethically in systematic and predictable ways, often without their awareness, driven by the bounded nature of our cognitive capacities.

Two aspects of bounded ethicality are particularly relevant here. The first is the extent to which it is supported by self-serving biases. Individuals interpret situations in ways that position themselves in a beneficial light (Greenwald 1980). This leads to behavioral consequences: given the opportunity, people serve their own immediate self-interest, while thinking of those actions as legitimate and fair (Batson et al. 1997; Diekmann et al. 1997; Messick and Sentis 1979). In fact, Epley and Caruso go so far as to claim that humans have an "automatic egocentric default" (2004, p. 174). The second is the extent to which our bounded ethicality occurs below the level of conscious awareness (Chance et al. 2011; von Hippel et al. 2005). Indeed, "...most of the bias that arises from human occupations and preoccupations cannot be described correctly as rooted in dishonesty—which perhaps makes it more insidious" (Simon 1983, p. 96). If our hedonic self-serving tendencies are so automatic, and so unacknowledged, the challenge of

fixing this default is particularly complex.

### **Intuitive/automatic “Type 1” processes vs. deliberative/reflective “Type 2” processes**

Building on work dating back to William James (1890/1950), Stanovich and West (2000) coined the term System 1 to describe cognitive processes underpinned by automaticity and heuristics, and System 2 to describe cognitive processes that are more effortful and deliberative. In response to an active debate in this area (Keren and Schul 2009; Kruglanski and Gigerenzer 2011; Osman 2004), the theory was amended to acknowledge that these different processes may rely on plural brain systems, but retained the useful distinction between rapid autonomous processes (Type 1) and higher order reasoning processes (Type 2) (Evans and Stanovich 2013). The Type 1/Type 2 distinction helps to build an understanding of how individuals navigate the tension between gratifying automatic preferences, and overriding those tendencies (such as complying with rules that constrain hedonic self-serving choices).

The defining feature of Type 1 processes is that they are automatic and do not require working memory (Evans and Stanovich 2013; Kahneman 2011; Stanovich and West 2000). The theory assumes that, unless Type 2 actively interferes, our natural tendency will be to default to Type 1 responses. Given “automatic egocentric defaults”, this often (though not always<sup>1</sup>) means that Type 1 processing motivates us towards the most viscerally attractive (hedonic) option. In contrast, the defining feature of Type 2 processes is that they require working memory and mental effort (Evans and Stanovich 2013; Kahneman 2011; Stanovich and West 2000). In Stanovich and West’s original theorizing (2000), they claimed that one of the functions of Type 2 processes is to serve as an override for potential dangers inherent in Type 1 processes. The deliberative processing of Type 2 offers a route through which to rein in automatic tendencies to serve one’s hedonic interest, particularly when those automatic tendencies contravene

---

<sup>1</sup> We do not suggest that Type 1 processes will always lead to higher rates of rule breaking, nor that Type 2 processes lead to universally lower ones. Research on intuitive responses to moral dilemmas confirms that many of our more instinctive Type 1 responses can lead to positive ethical outcomes (for an overview, see Haidt and Kesebir 2010; see also Zhong 2011). However, we are suggesting that heightened Type 2 processes will help in circumstances that pit self-interested options that are intuitively appealing against rules intended to constrain that behavior. Since they are more deliberative, Type 2 processes will more likely incorporate understandings of the relevant behavioral norms in the context as well as the consequences of one’s behavior, compared to more hedonic Type 1 processes that might prioritize short-term gain.

explicit behavioral expectations that have been set for us (Gino et al. 2011; Mead et al. 2009). Or, as Daniel Kahneman writes, “System 2 is in charge of self-control” (2011, p. 26).

Support for our claim that Type 2 processes are associated with an increased tendency to comply with the expectations that have been set for our behavior is found in studies that link deliberative processing with ethical behavior. Shalvi and colleagues (2012) asked participants to self-report the money they earned after rolling a dice privately, while experiencing either low or high time pressure. Participants lied less about their earnings in the low time pressure condition, which the authors argue allowed for time to deliberate. Similarly, Gunia and colleagues (2012) showed that when participants contemplate potential courses of action, they are more truthful when communicating to another person in a deception game. These findings suggest that more deliberative cognitive processing helps individuals resist making hedonic choices that cheat another person. While these studies do not test directly whether deliberative thinking leads to lower levels of rule breaking, they suggest that contexts that facilitate deliberation increase ethical behavior and facilitate rule compliance in the face of immediate and hedonic self-serving options. If this is the case, it is important to look for organizational interventions that might activate Type 2 processing.

### **Task Variety and Deliberative Thinking**

Even though there have been calls for research into how to activate Type 2 processes in an effort to improve decision outcomes, there has not been extensive research on this topic (Milkman et al. 2009). Interventions with demonstrated success in activating Type 2 processes include individual mental exercises such as considering an opposite perspective (Mussweiler et al. 2000), or mentally removing oneself from the specific situation (Kahneman and Lovallo 1993). Interventions an organization might use to activate Type 2 processes include training employees in statistical modeling, engaging in group decision-making, or increasing individual accountability for decisions (Larrick 2008; Lerner and Tetlock 1999). In this paper, we investigate whether increased task variety represents an intervention to encourage Type 2 processes, in order to decrease rule breaking. To our knowledge, it is the first research to test this feature of job design to support rule compliance, and the first time that a more traditionally cognitive

mechanism has been explored to explain the positive effects of more varied work.

We focus on a particular type of task design here: the extent to which the tasks at one's work are organized in a varied (rather than repetitive) way. Traditionally, task variety has been conceptualized in terms of "the degree to which a job requires a variety of different activities in carrying out the work" (Hackman and Oldham 1975, p. 161). This type of task variety focuses on the content and number of *skills* required by a job, and a good deal of work has focused on its effects (Hackman and Oldham 1976, 1980; Pierce and Dunham 1976). Indeed, high levels of task variety, operationalized in terms of the different skills a job requires, has positive motivational effects in the workplace (Humphrey et al. 2007). However, it is difficult to enrich the content and skills variety of the tasks that comprise many jobs. Jobs in data processing, assembly line operations, call centers, or diagnostic testing in healthcare all involve a finite set of steps and little flexibility to change them. In jobs like these, altering the content of the work is challenging, and the number of skills they require is limited.

Another way to experience variety at work relates to task sequence. Two jobs, identical in the number of skills required, can be experienced differently in terms of how the tasks that comprise them are ordered and how often individuals switch among them (Pentland 2003, p. 857). It is this type of task variety—variety in terms of how many times one switches from one task type to another—that is our focus. If an individual needs to complete three "Task As", "Task Bs", and "Task Cs", these can be sequenced with low levels of task variety (few switches between task types – e.g., AAABBBCCC – which involves two switches), or with high levels of task variety (constant changes among subtasks – e.g., ABCABCABC – which involves eight switches). We argue that the frequent cognitive switches required by higher levels of task variety mobilize more deliberative processing.

A given situation can elicit either Type 1 or Type 2 processing: when deciding whether it is more dangerous to travel by car or by an airplane, a Type 1 typical response might automatically bring to mind horrible images of plane crashes, concluding it is more dangerous to fly (an example of an availability heuristic, see Tversky and Kahneman 1974). A typical Type 2 response might compare the average miles driven per accident to the average miles flown per accident, concluding that car travel is more dangerous



than flying (Alter et al. 2007). Which process is more likely to operate and when? As we have noted, the Type 1 response is more effortless and automatic, and represents the “default” approach (Kahneman, 2011; Stanovich & West, 2000). We discuss two cues relevant to choices about task design – disfluency and interleaved tasks – that can activate Type 2 processes in the face of Type 1 defaults.

***Disfluency.*** Disfluency describes a subjective experience of difficulty associated with cognitive functioning. It leads to deeper processing of available information and results in decreased dependence on biased reasoning (Diemand-Yauman et al. 2011; Hernandez and Preston 2013). Disfluency also leads to increases in deliberative, Type 2 processes. One way to measure deliberative processing is to use Frederick’s (2005) Cognitive Reflection Task (CRT), a set of questions for which one’s intuitive responses are incorrect: correct answers require individuals to suppress their automatic responses, and engage in more deliberative reasoning processes. Alter and his colleagues (2007) found that participants who responded to the CRT in a font that was difficult to read (creating cognitive strain or disfluency) performed significantly better than those who solved the questions presented in an easy to read font. Therefore, situational constraints that cause individuals to expend more cognitive resources can trigger Type 2 processes, which help to override the more impulsive responses of Type 1 processes.

***Interleaved practice.*** Support for the prediction that task variety will activate Type 2 processes also comes from recent studies in education. Work on interleaved practice (Monsell 2003) focuses on how tasks that are practiced in an interleaved way (forcing switching between types of tasks, ABCABC, i.e., high variety) affect performance differently from tasks practiced in “blocks” (AABBCC, i.e., low variety). These studies find that interleaved practice leads to better long-term performance (Kornell and Bjork 2008; Taylor and Rohrer 2010). These studies show enhanced performance following tasks that require frequent task switching, indicating that this type of task variety triggers more deliberation.

Connecting this work to job design research, we suggest that undertaking sequentially varied tasks creates disfluency, triggering Type 2 processes. High levels of task variety entail frequent subtask switching, requiring an individual to maintain a state of heightened cognitive awareness (Cokely 2007). Though frequently alternating among types of tasks involves “switching costs” (Allport et al. 1994;

Shallice and Burgess 1993), potentially leading to longer reaction times and lower levels of short term performance (Jersild 1927; Staats and Gino 2012), we argue that it also activates Type 2 processes.

Ultimately, we posit that higher levels of task variety are associated with more deliberative thinking, which makes one vigilant about long term consequences over short term hedonic gains (Lindenberg and Steg 2007).

In sum, we argue that rule breaking at the workplace is often a result of employees' reliance on autonomous Type 1 processes that lead to self-serving (hedonic) behavioral defaults at the expense of ethical behavior. Just as deliberative thinking can help overcome cognitive decision biases (Frederick 2005), and reduce the impact of defaults in judgment (Alter et al. 2007), we argue it will also make people less likely to behave in ways that violate rules in favor of short-term self-interested gain. Thus, a task design intervention that activates Type 2 processes is also likely to reduce employees' rule breaking and encourage more ethical workplace conduct. Thus, we propose and test the following hypotheses:

**Hypothesis 1.** Engaging in tasks with high levels of variety (requiring frequent subtask switching) will lead to less rule breaking than engaging in tasks with low levels of variety.

**Hypothesis 2.** Engaging in tasks with high levels of variety (requiring frequent subtask switching) is more likely to lead to more deliberation (Type 2 processes)/less automaticity (Type 1 processes) than engaging in tasks with low levels of variety.

**Hypothesis 3.** More deliberation/less automaticity will mediate the relationship between high levels of task variety and lower levels of rule breaking.

## Overview of the present research

We tested these hypotheses in a field study and three laboratory experiments. Using data from the mortgage application processing department in a bank, Study 1 documents the relationship between task variety and how often employees violate organizational rules for permissible lunch breaks (Hypothesis 1). In Study 2, we manipulate task variety by varying task sequences to replicate the relationships between task variety and rule breaking in the form of cheating. Finally, Studies 3a and 3b test deliberative thinking

as the mechanism, using a different operationalization of rule breaking and both behavioral (Study 3a) and self-report (Study 3b) measures of Type 1 vs. Type 2 processing. We show that a subtle manipulation of task variety triggers higher levels of deliberative thinking (Hypothesis 2), and that higher levels of deliberative thinking/lower levels of automaticity mediate the relationship between sequential task variety and rule breaking (Hypothesis 3). We also use Studies 3a and 3b to rule out alternative explanations for our effects, including intrinsic motivation as the dominant of these.

## **Study 1: The phenomenon in the field**

In Study 1, we analyze data from a home loan mortgage processing operation at a Japanese bank, providing an ecologically valid demonstration of our hypothesized effect. We examine the effect of task variety on a clear rule breaking behavior: contravening corporate break time policies.

### **Method and Setting**

At this bank, there were seventeen distinct stages involved in processing a mortgage application. After submitted applications were scanned into the information technology system (an activity done by a different group of workers) the application processing group began its work. Each of the stages was considered independent by the bank and this was the level at which activities were assigned. Mortgage application processors sat at a desk with a computer and two monitors. One monitor showed the work to be completed (e.g., the form with data to be entered) and the other monitor showed the spaces in which the data were entered. Each of the seventeen stages involved a defined set of activities. For example, at the Custodian stage, operators confirmed that all scans of the materials were legible. At Document Tagging, individuals marked sub-sections of the document to be processed at subsequent stages. Then, in stages such as Application Capture or Preliminary Information, individuals entered data from the application into the system. Average task length was 2.7 minutes with a standard deviation of 3.2 minutes. Table 1 provides a brief description of each of the 17 stages and Figure 1 is a process flow diagram of the home loan mortgage process.

Workers received no specialized training and were eligible to complete any of the stages in the

process. A worker completed all work within a stage, there were no physical handoffs between stages, and workers did not have individual queues. Instead, when a worker finished one task, she was assigned a new one by the system. In setting up the mortgage processing line, management's objective was to eliminate as much of the human element as possible. As such, employees could not request different tasks. As workers completed their tasks, if there were a backup at one stage then the system would automatically assign the next task at that stage to the next available worker. This approach to task assignment meant that over the course of a single day, workers could remain on the same stage, or switch among multiple stages, depending on the system dynamics.

Our sample analyzes data collected between June 1, 2007, when the new mortgage processing line began, until December 30, 2009. We begin with information on all workers who processed loan applications during this period. We drop the records from a total of twenty-nine individuals who completed fewer than 200 transactions each because they were short-term temporary workers, workers separated before the end of their two-week probationary period at the start of their employment, or managers who conducted transactions when workers were away. Since we seek to predict if individuals contravene corporate break policies, and workers take one lunch break per day, we analyze all models at the day-level. Thus, we are left with 17,161 person-day observations across 110 employees. We note that the researchers had no role in setting up the corporate procedures and all procedures were already in use by the organization.

## Measures

*Dependent Variable.* We examine whether individuals take a lunch break that exceeds the length permitted under the company's official regulations. Employees are expected to perform at a high standard throughout the day, ensuring the accuracy of the mortgage applications and treating their co-workers respectfully. According to a bank manager, "supervisors are not expected to monitor the employee lunch break." Workers log out of the system when they begin lunch, and log back in when they return, and we can measure whether their break exceeded the allowable time, and if so, by how long.

Taking a longer than mandated lunch represents clear rule breaking behavior in this organization.

In addition, this behavior has wider negative implications for the organization and the employees' colleagues. Workers not returning on time can disrupt operational functioning and leave others with extra work to complete, possibly requiring them to delay their own scheduled breaks. Since employees on the processing line receive a fixed hourly wage with no incentive compensation (pay raises are determined by an individual's tenure at the organization, not from the individual's performance), complying with break time regulations also represents behavior with important ethical implications.<sup>2</sup> In fact, five of the eleven items in Bennett and Robinson's measure of workplace deviance (2000) tap different ways of wasting the company's or others' time, including one that directly addresses our dependent variable here: "Taken an additional or longer break than is acceptable at your workplace". We measure our outcome of interest in two ways: *Long lunch* is a dichotomous variable set to one if an individual exceeds the regulated lunch length and is zero otherwise. *Length of long lunch* is a continuous measure of the number of minutes exceeding regulation time that the employee returned late from lunch – the variable either takes the value of zero or a number greater than zero.

*Independent variable.* To examine our key research question, we wish to capture the changes in task type that an individual experiences prior to the start of her lunch break. Therefore, we test our

---

<sup>2</sup> We have no data from the bank that would allow us to confirm that employees perceived taking long lunch breaks as a rule violation that is unethical. However, to assess whether this behavior is considered inappropriate in the general population, we conducted a short validation survey using Amazon Mechanical Turk ( $N=399$ , 48% female,  $M_{\text{age}}=35.47$ ,  $SD=10.88$ ). We adapted a set of questions from a dissertation-based field study about ethics (used for but not reported in Martin 2015), to establish in a sample of working adults that coming back late from lunch in a similar interdependent setting (as our scenario was worded) is perceived as 'inappropriate'. We asked participants to rate 10 different scenarios in terms of their appropriateness on a 9-point scale (1=completely inappropriate; 9=completely appropriate). Five were "filler" scenarios, that depict ethically neutral behavior, to serve as a benchmark (e.g. leaving a line when one is tired of waiting), and an additional five that represent inappropriate behavior (e.g. over-reporting one's work hours, tax fraud etc.), including our focal scenario of interest: ("It is 12:29 and your lunch break at work finishes in one minute. However, you don't feel like going back to work yet. You know that no one from management really monitors when exactly you come back from work so you decide to stay for another 15 minutes on break even though you know this will delay your colleagues' scheduled breaks and may also hurt your team performance. You come back from lunch break at 12:45 instead."). The mean rating on the "filler" scenarios ( $M=7.61$ ,  $SD=1.08$ ) differed from the mean rating of the inappropriate ones ( $M=3.72$ ,  $SD=1.69$ ,  $t(398)=34.80$ ,  $p < .001$ ). Note that lower ratings mean more inappropriate behavior. The mean ratings for the late lunch scenario ( $M=3.60$ ,  $SD=2.12$ ) was not significantly different from the mean rating for understating one's taxes ( $M=3.54$ ,  $SD=2.28$ ) or over-reporting one's work hours ( $M=3.55$ ,  $SD=2.40$ ). These results suggest that returning late from a lunch break represents a rule violation that is considered as inappropriate as tax fraud and time theft. A full description of the scenarios is available from the authors upon request.

hypothesis with a variable, *pre-lunch task variety*, which counts the number of times a worker changes between different stages prior to her break. As an example, if a worker completed four tasks from two different Stages (e.g., Application Capture 1 and Credit Check 2), two possible assignments of the tasks could be: 1122 or 1212. In the first example, task variety would equal 1; in the second it would equal 3.

*Control variables.* We include a number of control variables. First, we control for how workers' tasks are allocated across different stages utilizing the Blau measure. This is calculated for each individual on each day, as follows, where  $i$  captures each of the 17 stages in the data:

$$Pre - lunch task type = 1 - \sum_{i=1}^{17} \left( \frac{Pre - lunch cumulative volume_i}{Pre - lunch cumulative volume} \right)^2$$

Task type is thus captured as a continuous variable on a daily basis. The maximum value of the Blau measure is equal to  $(N-1)/N$  where  $N$  is the number of stages, in our case 17. As an example of the calculation, if a worker were to do two tasks each in two different stages in the morning then her pre-lunch task type would equal  $1 - [(\frac{1}{2})^2 + (\frac{1}{2})^2] = \frac{1}{2}$ . Thus, with the inclusion of this control variable we are able to separate out the allocation of tasks to different categories from the number of task changes that a worker experiences. We believe that this makes our test a more conservative test of our hypothesis.<sup>3</sup>

Second, we control for the utilization of a worker by calculating the number of minutes she was working on tasks, prior to lunch, and dividing that value by the total number of minutes she was available to work on tasks (*pre-lunch workload*). When worker utilization is higher, workers may feel more pressure to follow company policy and return within the regulated time. Third, we control for a worker's cumulative morning activity by constructing a variable, *pre-lunch cumulative volume*, which counts the total number of tasks that a worker has completed prior to the start of lunch. Workers may complete a higher volume of tasks in the morning on days where the overall level of activity is higher and may feel pressure to conform to policy and be less likely to overstay their break. Fourth, to control for a worker's

---

<sup>3</sup> Another measure to capture task allocation across stages is the Teachman-Shannon index (Harrison and Klein 2007). We repeated all reported models substituting this index for the Blau measure and generated similar support for Hypothesis 1.

prior experience we include *prior-day cumulative volume*, a count of the number of transactions that each worker has completed on all days prior to the current day. Workers who have completed more prior tasks may feel less constrained by company policy and may be more likely to take a longer lunch, compared to newer, less experienced workers. Fifth, to control for environmental differences across days, such as overall load on the system or external concerns, we include indicators for each day captured in the data. Finally, we control for any time-invariant characteristics of the worker, such as predispositions to follow rules, by including individual indicators. As such, all analyses are conducted “within-worker.” We note Hypothesis 1 remains supported even if all control variables are dropped.

### **Empirical Approach**

As our first dependent variable is dichotomous, we use a logistic regression model to examine it. For our second dependent variable (the continuous measure) we use two approaches. First, we use a Tobit regression model. We use this model since our continuous data are constrained at zero as a minimum and then increases from there. Second, we run a linear regression model just examining those cases where individuals take a longer than expected lunch.

### **Results and Discussion**

Table 2 presents summary statistics and a correlation table for all variables, and Table 3 presents the regression results. Column 1 examines the dichotomous outcome variable. Consistent with Hypothesis 1, we find that greater task variety is related to a decreased likelihood to violate company regulations regarding lunch breaks. A one-standard-deviation increase in pre-lunch task variety decreases the odds of taking a long lunch by 32%. We note that workers who took a longer lunch did not compensate by staying longer that day. This support for our hypothesis remains if we drop day or individual indicators, as well as if we estimate Column 1 with either a conditional fixed effects logistic regression model or a hierarchical linear regression model. In Column 2 we examine results from the Tobit regression model and see that the coefficient on pre-lunch task variety is negative and statistically significant, and in Column 3 we also see that pre-lunch task variety is negative and statistically significant. A one-standard-deviation increase in

pre-lunch task variety decreases the length of lunch by 10-15 minutes, conditional on taking a long lunch.

Our field data illustrate that greater task variety is related to a lower likelihood of violating company break time regulations. The results of this study are important as they show the effect of ‘naturally occurring’ task variety on rule breaking in an actual workplace setting with real employees. However, these results do not show support for our proposed mechanism, nor do they allow us to establish causality in a compelling way. As such, we now turn to the laboratory.

## Laboratory Experiments

### Study 2

In Study 1, we relied on a continuous measure of task variety experienced by workers on a daily basis. However, workers each completed a different set of unique tasks each day. In Study 2, we manipulate task variety experimentally to examine its effect on rule breaking. Compared to the field data, in our experiment all participants completed exactly the same tasks, but we varied their sequence.

*Participants.* We recruited 221 participants (49% female,  $M_{\text{age}} = 35.50$ ,  $SD = 11.47$ , 76% employed) through Amazon Mechanical Turk (Buhrmester et al. 2011; Hauser and Schwarz 2016; Paolacci et al. 2010) for \$2.00, as well as an additional \$0.10 for each correct answer they reported in Round 2 (see below). The study was advertised as an examination of different problem solving techniques. At the end of the study, we asked participants to state what they thought the study was about, and excluded nineteen who wrote that they thought the study related to cheating or ethics. The final analysis includes the remaining 202 participants.

*Procedure and Materials.* Participants were told they would solve a set of 15 mathematical, verbal and spatial problems, with 30 seconds to respond to each problem, after which the correct answer would appear on the screen. In the low variety (‘blocked’) condition ( $N = 101$ ), participants solved five math problems, five verbal ability problems, and five spatial ability problems that were blocked together



(e.g. AAABBBCCC).<sup>4</sup> In the high variety ('interleaved') condition ( $N = 101$ ), participants saw exactly the same problems as in the low variety condition, but ordered in an interleaved way, forcing participants to alternate between different types of problems each time (e.g. ABCABCABC).

Following Round 1, participants received instructions for Round 2, in which we assessed rule breaking. Participants were assigned to either a 'no opportunity to cheat' ( $N = 110$ ) or an 'opportunity to cheat' ( $N = 92$ ) condition (using a paradigm from Chance et al. 2011). In both conditions, participants read that they were about to start the second round of problems and were told that they would win \$0.10 for each correct answer out of the 10 questions in this round. However, in the 'opportunity to cheat' condition they also read:

*"You are now about to start the next set of problems. In this Round we make it possible for you to check after each question whether you got it right or wrong. We want you to mark down the answer you arrived at on your own. **After** you mark your answer you may scroll down and see whether you were correct. Please do so only after you have answered the question."*

All participants then started Round 2. The questions were again presented in a blocked way to participants in the low variety condition, and in an interleaved manner for participants in the high variety condition. In addition, those in the 'opportunity to cheat' condition were able to scroll down and see the correct answer for each question. We chose more challenging questions for the second round to increase the likelihood of cheating. We then debriefed, thanked and paid the participants.

## Results and Discussion

This paradigm allowed us to assess the extent to which our manipulation of task variety differentially encouraged participants to violate the explicit instructions not to look at the answers in advance of reporting their responses. To do this, we first need to show that the paradigm encouraged rule breaking. That is, we need to show that, given the opportunity to look at the answers in advance, individuals would do so, even with explicit instructions not to. The results indicated that those in the 'opportunity to cheat' condition had a significantly higher percentage of correct answers in Round 2

---

<sup>4</sup> An earlier pilot indicated that the order of the blocks did not affect the results, so in the low variety condition participants saw the mathematical 'block' first, followed by the verbal 'block' and then the spatial 'block'.

( $M=.79$ ,  $SD=.19$ , note that all performance measures are reported as proportions from 0 to 1, where 1 represents getting all questions correct), compared to those in the ‘no opportunity to cheat’ condition ( $M=.60$ ,  $SD=.19$ ;  $F(1,200)=49.11$ ,  $p<.001$ ,  $\eta_p^2=.19$ ), suggesting that individuals who had the opportunity to cheat did indeed look at the answers in advance of reporting their own. Thus, consistent with other work using this paradigm (Chance et al. 2011), the experiment did encourage rule breaking (see Figure 2). While this paradigm does not allow us to identify specific individuals who broke the rules, it does allow us to compare how different experimental conditions affect rule breaking in aggregate (e.g., Chance et al. 2011; Shalvi et al. 2012).

Our key expectation was that, provided with the opportunity to break the rules and look at the correct answers in advance, individuals would do so *less often* when they were in the high variety condition compared to the low variety condition. We expected that the variety manipulation would *not* affect Round 2 performance when participants did not have an opportunity to cheat, but that it would affect Round 2 performance when participants did have an opportunity to cheat. When there was an opportunity to cheat, a significant difference in performance as a function of participants’ prior task variety (assuming random allocation of participants to experimental conditions) could be attributed to higher levels of rule breaking behavior. Performance in Round 1 did not differ by task variety condition ( $M_{low\_variety}=.65$ ,  $SD=.15$ ;  $M_{high\_variety}=.63$ ,  $SD=.16$ ,  $F(1,200) = 1.19$ ,  $p=.27$ ). However, age was a significant predictor of performance, so we controlled for age in this and all subsequent analyses. As predicted, in the ‘no opportunity to cheat’ condition, there was no significant difference between the low variety ( $M=.62$ ,  $SD=.17$ ) and the high variety ( $M=.58$ ,  $SD=.21$ ) conditions in terms of their Round 2 performance,  $F(1,107)=1.81$ ,  $p=.18$ . However, when there was an opportunity to cheat, there was a significant difference between the low variety ( $M=.84$ ,  $SD=.18$ ) and high variety ( $M=.76$ ,  $SD=.21$ ) conditions,  $F(1,89)=4.24$ ,  $p=.04$ ,  $\eta_p^2=.05$ , see Figure 2).

While the size of the effect we observe for the role of task variety on rule-breaking is not large given conventional rules of thumb (Cohen 1988), it is important to assess effects within the context which

elicits them (Cortina and Landis 2009; Prentice and Miller 1992). An effect that can be elicited using a subtle manipulation is valuable, because it can be detected in “the least auspicious of circumstances” (Cortina and Landis 2009, p. 296). In light of meta-analytic findings that the presence of a code of conduct (something intended to directly influence ethical behavior in a positive way) has no effect on employees’ ethical behavior (Kish-Gephart et al. 2010), the fact that a subtle change in the way someone’s tasks are ordered (something completely unrelated to ethical behavior) has any effect on rule breaking behavior is important. The magnitude of this effect is consistent with other experimental research that focuses on interventions that can shift morally relevant organizational behaviors, such as conflicts of interest (Sah and Loewenstein 2014), and overbilling (Desai and Kouchaki 2015).

An additional way to think about the effect of task variety on rule breaking is to compare whether the large main effect we observe for rule breaking when there is an opportunity to cheat compared to when there is not ( $\eta_p^2=.19$ , from above) is *larger* within the low variety condition than it is in the high variety condition. In the low variety condition, the difference in Round 2 performance between the ‘no opportunity to cheat’ condition ( $M=.62$ ,  $SD=.17$ ) and the ‘opportunity to cheat’ condition ( $M=.84$ ,  $SD=.18$ ) was significant ( $F(1,98)=39.93$ ,  $p<.001$ ,  $\eta_p^2=.29$ ). This effect also emerged in the high variety condition: there was a significant difference between performance in Round 2 in the ‘no opportunity to cheat’ ( $M=.58$ ,  $SD=.21$ ) and the ‘opportunity to cheat’ condition ( $M=.76$ ,  $SD=.21$ ), though, as expected, with a smaller effect size,  $F(1,98)=20.51$ ,  $p<.001$ ,  $\eta_p^2=.17$ . The Cohen’s  $q$  statistic for the difference between these two effect sizes is 0.164, which indicates that the difference between them is statistically significant (Rosnow and Rosenthal 2003), and confirms that the level of rule-breaking we observe in the low variety condition is larger than the level we observe in the high variety condition.<sup>5</sup> Overall, the

---

<sup>5</sup>Although Rosnow and Rosenthal (2003) caution that evaluating whether two effect sizes are significantly different usually requires a substantially larger sample than what is required to detect a single significant effect, we nevertheless followed their guidelines and calculated Cohen’s  $q$  in order to compare the two effect sizes. We first transformed the eta squared effect sizes obtained from SPSS to  $r$ , and obtained Cohen’s  $q$  based on the subtraction of the Fisher’s  $Z$  transformed correlations.

pattern of results as depicted in Figure 2 is consistent with our prediction.

Finally, since task switching might slow response times, and thus have efficiency implications, we explored whether task variety affected the average number of seconds participants spent answering each question in both rounds. There was a marginally significant difference in the time spent on each question in Round 1 between the variety conditions ( $M_{low\_variety}=19.46$ ,  $SD=3.44$ ;  $M_{high\_variety}=20.32$ ,  $SD=3.76$ ;  $F(1,200)=2.87$ ,  $p=.09$ ,  $\eta_p^2=.01$ ). In Round 2, there was no difference in the time spent on each question between the variety conditions within the ‘no opportunity to cheat’ condition ( $M_{low\_variety}=19.61$ ,  $SD=4.38$ ;  $M_{high\_variety}=19.65$ ,  $SD=5.44$ ;  $F(1,108)=.003$ ,  $p=.96$ ), nor within the ‘opportunity to cheat’ condition ( $M_{low\_variety}=18.28$ ,  $SD=6.58$ ;  $M_{high\_variety}=19.19$ ,  $SD=5.83$ ;  $F(1,90)=.49$ ,  $p=.48$ ).

### Study 3a

While Study 2 replicates the effect we observe in the field in a controlled experimental context, it does not test the proposed mechanism behind this effect. In Study 3a we examined deliberative thinking as our hypothesized mechanism (Hypotheses 2 and 3), as well as three motivational alternative explanations for our effect. The most obvious of alternative explanation for our effect is intrinsic motivation: people who engage in varied tasks may be more intrinsically motivated (Hackman and Oldham 1976, 1980; Pierce and Dunham 1976), and thus will perform without lapsing into rule breaking. A second alternative explanation for our effect is reciprocity. Social exchange theorists (Blau 1964; Emerson 1976) suggest that people are motivated to respond reciprocally to the way they are treated (Gouldner 1960). Individuals engaged in varied tasks may appreciate the more interesting sequence of work, and reciprocate that appreciation by not breaking the experimenter’s rules. Finally, less varied tasks may lead to more negative affect, which may motivate deviant behavior such as rule breaking and represents a third alternative potential explanation for our effect.

*Participants.* Ninety-four participants from a participant pool at a large U.S. university’s behavioral lab (59% female,  $M_{age}=25.51$ ,  $SD=7.14$ ) earned a \$12 show-up fee as well as up to a \$2.50 performance-based bonus. We excluded participants who indicated they were less than completely fluent

in English ( $N=10$ )<sup>6</sup>. In addition, two participants reported technical errors with pressing the buttons on the keyboard during the study and two additional participants had to be excluded because their data were recorded using the same participant number. Thus, the final analysis includes 80 participants.

*Procedure and Materials.* After entering the lab, participants were seated at computers. The procedure up to the end of Round 1 was the same as Study 2, with the computer randomly assigning participants to either a high task variety ('interleaved') condition ( $n=42$ ) or a low task variety ('blocked') condition ( $n=38$ ), and presenting 15 math, spatial and verbal questions for them to answer. Participants then completed a measure of deliberative thinking. The computer screen then instructed them to notify the experimenter, so they could proceed to Round 2. Following a procedure adapted from von Hippel, Lakin, and Shakarchi (2005), when the experimenter came to initiate Round 2, she told them:

*"You will now start the second set of problems. However, we just discovered a 'bug' in the programming of this set of problems, and within two seconds from the appearance of the problem, the correct answer appears on the screen. The only way to prevent this from happening is to press the f6 key when the problem first appears. This will prevent the right answer from appearing on the screen. So please remember to press f6 immediately after you see each problem on the screen".*

Participants then started the final 10 problems, which again contained either low or high levels of variety. Again, participants had 30 seconds to solve each problem, but if they did not press f6 within two seconds of the appearance of the problem, the correct answer appeared on-screen. We then assessed participants' motivation and affect. Finally, we debriefed, paid, and thanked the participants.

## Measures

*Deliberative thinking.* We measured deliberative thinking using the Cognitive Reflection Task (Frederick 2005). As noted above, the CRT consists of three questions for which identifying the correct answers "requires the suppression of an erroneous answer that springs 'impulsively' to mind" (Frederick, 2005: 27). The CRT has been validated as a measure of deliberative, Type 2 thinking (Cokely 2007;

---

<sup>6</sup> Given the nature of the verbal problems we used in the study, we wanted to ensure our participants were fluent in English. This was especially important in Study 3b, which was conducted using a population that includes a substantial proportion of participants for whom English is not a first language. Wanting to ensure consistency, we report our results excluding anyone who did not report being completely fluent in this sample as well. We note, however, that repeating the analyses including all participants who marked themselves above the mid-point on English fluency yields exactly the same results.

Frederick 2005; Kahneman and Frederick 2005; Toplak et al. 2011), with low levels of performance associated with more automatic/Type 1 processing, and high levels of performance associated with more Type 2/deliberative processing. An example question is: “*A bat and a ball cost \$1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? (in cents)*”. An automatic (and incorrect) answer to this question is 10 cents. The correct answer is 5 cents. Participants scored between 0 and 3 on this measure ( $M=1.05$ ,  $SD=1.13$ ).

*Rule breaking.* Our measure of rule breaking was the number of times participants failed to press the f6 key (to avoid seeing the right answer). This “bug” provides an opportunity for participants to let the right answer appear on the screen before answering themselves. This paradigm not only captures the extent to which an individual violates explicit instructions, it also measures: every time the participant sees an answer before answering represents illegitimately earning \$0.10 (see Shu and Gino 2012; Vohs and Schooler 2008). Participants received no feedback about their performance during the rounds, and only learned how many correct answers they provided across both rounds at the end of the experiment.

*Intrinsic motivation.* To measure intrinsic motivation, we used the interest/enjoyment subscale of the Intrinsic Motivation Inventory (IMI; Deci and Ryan 1985; McAuley et al. 1989). Participants were asked to think about the problem-solving questions they solved at the beginning of the experiment, and then responded to 7 items (e.g., “*I enjoyed doing this activity very much*”) on a 5-point scale ( $\alpha=.90$ ).

*Felt obligation.* We adapted 4 items from Eisenberger et al. (2001) on a 5-point scale (1=not at all true; 5= very much true,  $\alpha=.64$ ) to measure feelings of reciprocity towards the experimenter (e.g., “*I owe it to the experimenter to give 100% of my energy to her goals while I am completing surveys*”).

*Affect.* We measured participants’ mood using the 20-item PANAS (Watson et al. 1988) on a 5-point scale (1=not at all or slightly; 5=extremely). Sample items are “*upset*” and “*irritable*” for negative affect (NA,  $\alpha=.81$ ) and “*enthusiastic*” and “*interested*” for positive affect (PA,  $\alpha=.84$ ).

## Results and Discussion

Performance in Round 1 did not differ by condition ( $M_{low\_variety}=.64$ ,  $SD=.14$ ;  $M_{high\_variety}=.68$ ,  $SD=.11$ ,  $F(1,78) = 1.96$ ,  $p=.16$ ). And, as in Study 2, we again observed no difference between the low and

high variety conditions in terms of the average number of seconds it took to answer in Round 1 ( $M_{low\_variety}=17.25$ ,  $SD=20.41$ ;  $M_{high\_variety}=16.97$ ,  $SD=19.52$ ,  $F(1,78)=.38$ ,  $p=.54$ ) or in Round 2 ( $M_{low\_variety}=16.65$ ,  $SD=39.26$ ;  $M_{high\_variety}=15.41$ ,  $SD=29.18$ ,  $F(1,76)=2.59$ ,  $p=.11$ ).<sup>7</sup>

However, as expected, participants in the high variety condition scored better in the CRT ( $M=1.36$ ,  $SD=1.16$ ) compared to those in the low variety condition ( $M=.71$ ,  $SD = 1.01$ ,  $F(1,78)= 6.96$ ,  $p=.01$ ,  $\eta_p^2 = .08$ ). This result supports Hypothesis 2, that task variety fosters more deliberative thinking. To examine if deliberative thinking mediates the relationship between variety and rule breaking we used the SPSS macro PROCESS (Hayes 2013). We ran a model with rule breaking (number of times f6 was not pressed) as the dependent variable, condition (low vs. high task variety) as the independent variable, and performance on the CRT as the mediator. To be consistent with Study 2, we controlled for age. Using 5,000 bootstrap samples, a bias-corrected 95% confidence interval (CI) for the indirect effect did not include zero (point estimate=-.37; 95 % CI =-.91 to -.06), indicating a significant indirect effect. This supports Hypothesis 3: changes in sequential task variety are associated with changes in rule breaking *through* deliberative/Type 2 thinking.<sup>8</sup>

We note that, in this study, the main effect of task variety on rule breaking was not significant ( $M_{low\_variety}=3.08$ ,  $SD=2.83$ ;  $M_{high\_variety}=2.40$ ,  $SD=2.06$ ,  $F(1,78)=1.50$ ,  $p=.22$ ). However, there is growing consensus that a direct relationship between an independent variable and an outcome is not required in order to show mediation (MacKinnon et al. 2002; Shrout and Bolger 2002), and that the focus of studies that test mediation should be the on the significance of the indirect effect rather than on the significance of the main or direct effects (Hayes 2013; Rucker et al. 2011). In addition, research shows that demonstrating a direct effect in a mediating model requires significantly more power or stronger

---

<sup>7</sup> The analysis for time in Round 2 here as well as in Study 3b was performed only for the time spent on problems that participants solved without first seeing the correct answer.

<sup>8</sup> We replicated the mediating role of deliberative thinking in the relationship between task variety and rule breaking in another study in which we manipulated question breadth in addition to sequence ( $n=45$ ). Given the larger sample size and more subtle manipulation of variety in the current study (which provides a more conservative test of our predictions), we report only those results here. The results from this additional study are available from the authors.

manipulations than showing a significant indirect effect alone (see LeBreton et al. 2003). Particularly given the subtlety of the manipulation of task variety we used, it is meaningful that we find both the direct effect of our manipulation on CRT as a mediator, as well as the significance of the indirect effect of task variety on rule breaking via CRT performance.

Finally, we examined whether our focus on the cognitive effects of task variety may have caused us to neglect alternative explanations for our effect. Participants in the low variety condition did report marginally lower levels of intrinsic motivation ( $M=3.19$ ,  $SD=.88$ ) than those in the high variety condition ( $M=3.56$ ,  $SD=.81$ ,  $F(1,78)=3.87$ ,  $p=.053$ ,  $\eta_p^2=.05$ ), as well as lower levels of felt obligation to the experimenter ( $M_{low\_variety}=3.56$ ,  $SD=.83$  vs.  $M_{high\_variety}=3.85$ ,  $SD=.69$ ,  $F(1,78)=2.92$ ,  $p=.09$ ,  $\eta_p^2=.04$ ). Though significant, the effect of task variety on these two potential mediators was smaller than its effect on the CRT. Moreover, the indirect effects of task variety on rule breaking through both these potential mediators yielded confidence intervals that included zero, suggesting that neither of these variables mediated the effect of our variety manipulation on rule breaking. Lastly, there was no difference in positive affect or negative affect between the two conditions (PA:  $M_{high\_variety}=3.08$ ,  $SD=.65$ ;  $M_{low\_variety}=3.03$ ,  $SD=.59$ ;  $F(1,78)=.16$ ,  $p=.69$ ; NA:  $M_{high\_variety}=1.48$ ,  $SD=.44$ ;  $M_{low\_variety}=1.55$ ,  $SD=.49$ ,  $F(1,78)=.49$ ,  $p=.49$ ).

The results of Study 3a showed that task variety encourages more deliberative thinking (measured directly using the CRT), providing support for Hypothesis 2. Most importantly, the results showed that one's score on the CRT mediated the relationship between task variety and the extent to which participants complied with the experimenter's explicit instructions, providing support for Hypothesis 3.

### Study 3b

Though Study 3a aimed to rule out several motivational alternative explanations for our effect, one notable limitation of this study is that the variables we used to rule out these alternatives were measured following the assessment of our dependent variable. Additionally, while our proposed mechanism was measured directly (manifest in performance on the CRT), the motivational measures were



assessed in a self-report manner. These different assessments may not have given motivational explanations a fair chance as possible alternative candidates for the mechanism underlying the relationship between task variety and rule breaking. In Study 3b, we aimed to compare deliberative thinking and intrinsic motivation more fairly as alternative explanations. We focused on intrinsic motivation as the results from Study 3a indicated it was the strongest potential alternative explanation for our effects.

*Participants.* One hundred and forty-eight participants from a participant pool at a UK-based business school behavioral lab (63% female,  $M_{\text{age}}=27.93$ ,  $SD=9.66$ ) earned a £10 show-up fee as well as up to a £2.50 performance-based bonus. We excluded seven participants who indicated suspicion about the glitch in the system, and twenty participants who indicated they were less than completely fluent in English.<sup>9</sup> Thus, the final analysis includes 121 participants.

*Procedure and Materials.* The procedure was almost identical to the procedure of Study 3a. Upon arrival to the lab participants were randomly assigned to either a high task variety ('interleaved') condition ( $n=60$ ) or a low task variety ('blocked') condition ( $n=61$ ). The only difference from Study 3a was that after Round 1 we assessed both potential mechanisms in a self-report manner. We measured our proposed mechanism using two items to tap automaticity (to represent lack of deliberation). Participants indicated on a 5-point scale (1=not at all; 5=to a large extent) their agreement with the statements: (1) "*I feel impatient now*" and (2) "*I feel like I am on automatic mode right now*" ( $r=.51$ ). To assess intrinsic motivation, participants indicated on a 5-point scale their responses to the statements about Round 1: (1) "*I enjoyed doing this activity*" and (2) "*This activity was fun to do*" ( $r=.84$ ). Participants then continued to Round 2, with the same computer glitch as in Study 3a, when we assessed rule breaking. Finally, we debriefed, paid, and thanked the participants.

## Results and Discussion

As in previous studies, performance in Round 1 did not differ by condition ( $M_{\text{low\_variety}}=.59$ ,

---

<sup>9</sup> We note that repeating the analyses with all non-suspicious participants who marked themselves above the mid-point in English fluency level yields the same results, with a 90% CI for the main mediation result.

$SD=.13$ ;  $M_{high\_variety}=.58$ ,  $SD=.14$ ,  $F(1,119) = .11$ ,  $p=.74$ .), nor did we observe any difference between the low and high variety conditions in terms of the average number of seconds it took to answer the problems in Round 1 ( $M_{low\_variety}=19$ ,  $SD=27.66$ ;  $M_{high\_variety}=19$ ,  $SD=21.45$ ,  $F(1,119)=0$ ,  $p=1$ ) or in Round 2 ( $M_{low\_variety}=17.23$ ,  $SD=38.61$ ;  $M_{high\_variety}=16.79$ ,  $SD=31.28$ ,  $F(1,116)=.45$ ,  $p=.50$ ).

Participants in the low variety condition reported higher automaticity ( $M=2.79$ ,  $SD=1.15$ ) compared to participants in the higher variety condition ( $M=2.42$ ,  $SD=1.01$ ;  $F(1,119)=3.34$ ,  $p=.07$ ,  $\eta_p^2=.03$ ). There was no difference between the variety conditions in terms of intrinsic motivation ( $M_{low\_variety}=3.41$ ,  $SD=1.12$ ;  $M_{high\_variety}=3.31$ ,  $SD=1.07$ ,  $F(1,119) = .22$ ,  $p=.64$ ). To examine both deliberative thinking and intrinsic motivation simultaneously as mediators of the relationship between task variety and rule breaking, we used Model 4 of the SPSS macro PROCESS (Hayes 2013), controlling for age as in all other analyses. Using 5,000 bootstrap samples, the results showed that a (lack of) deliberative thinking mediated the relationship between task variety and rule breaking (point estimate=-.15; 95 % CI =-.52 to -.002). Variety was associated with *less* automaticity (coefficient = -.42,  $p=.04$ ), and automaticity was associated in turn with *more* rule breaking (coefficient =.36,  $p=.07$ ). In contrast, the confidence interval for the indirect effect via intrinsic motivation contained zero (point estimate=-.03; 95% CI =-.25 to .04). These results provide evidence that allow us to rule out intrinsic motivation as an alternative explanation for our effect more rigorously.

Similarly to Study 3a, the main effect of task variety on rule breaking was not significant ( $M_{low\_variety}=2.52$ ,  $SD=2.28$ ;  $M_{high\_variety}=2.63$ ,  $SD=2.44$ ,  $F(1,119)=.06$ ,  $p=.80$ ). In Studies 1 and 2 we observe a direct effect of task variety on rule breaking, but do not have a measure of the mechanism. Once we introduce the mechanism in between the manipulation and the dependent variable (in both Studies 3a and b), the direct effect is not significant, but we have substantial support for the prediction that changes in sequential task variety are associated with changes in rule breaking *through* a deliberative mindset. Given both our *a priori* theoretical framework and arguments regarding deliberative thinking, and current thinking on mediation (Hayes 2013; Rucker et al. 2011), we believe that our analyses provide solid support for our model.

## General Discussion

An ongoing challenge for organizations is to find ways to support employees to follow the rules. Unfortunately, due to egocentric defaults (Epley and Caruso 2004), employees often behave in self-serving and unethical ways. Given the organizational and societal costs of rule breaking behavior (Martin et al. 2013) it behooves us to identify organizational interventions that increase compliance. In this paper, we explore one such intervention: task variety. Drawing on literature on deliberative/Type 2 thinking (Evans and Stanovich, 2013; Kahneman, 2011), we predicted that structuring our tasks in a varied way would activate Type 2 processes, which in turn would support our ability to act more in line with societal norms and resist violating rules to serve our own hedonic self-interest.

In order to investigate our predictions, we conducted a field study and three laboratory experiments. We first aimed to identify the phenomenon and establish the main effect in a naturally occurring organizational setting, and then to explore causal effects as well as examine the mechanism behind the hypothesized effect in a controlled and rigorous way. Study 1 provided evidence that daily levels of task variety (the number of times mortgage application processors switch between different types of subtasks) are associated with fewer violations of employee lunch length regulations. Although we were not able to collect data that allowed us to test the mechanism behind this effect in this study, our *within-subject* analysis shows that the same person can be more or less ethical (i.e. take proper breaks or longer than allowed breaks) as a function of the *variety* of tasks she performed that morning. Study 2 replicated this effect in the lab using a subtle manipulation that did no more than vary the sequence in which the same set of math, spatial and verbal questions were presented to participants. Finally, using the same subtle manipulation, Studies 3a and 3b also provided evidence for increased deliberative, Type 2 processing as our mechanism (behaviorally in Study 3a and self-reported in Study 3b). Studies 3a and 3b also provided evidence to rule out intrinsic motivation as a supported alternative explanation.

## Contributions

This research offers several contributions to the literature. First, our research helps us understand

the levers that organizations may use to increase employees' compliance to rules or regulations. Most work on rule compliance has focused on motivational, social influence or incentive-driven strategies (Lehman and Ramanujam 2009; Martin et al. 2013; Tyler and Blader 2005; Wenzel 2005). Here, we offer a new perspective on increasing rule compliance in organizations, one that does not require direct or coercive regulation of employee conduct, and thus, in many ways, may suffer less backlash from employees who may desire autonomy rather than corporate control (Salaz 2002).

Second, our work contributes to theory about behavioral ethics. While many studies on ethical decision making focus on motivational triggers of unethical behavior (Greenberg 1997; Roy 1959), an emerging body of work proposes that our bounded cognitive capacities may also drive unethical behavior (Chugh et al. 2014). Our study offers one way that we can use to mitigate these bounded capacities: subtle changes in the way tasks are designed are likely to bolster deliberative/Type 2 processes, and in turn reduce rule breaking and improve ethical behavior. Our results suggest that the way we organize our tasks at work has the potential to make one less likely to fall prey to automatic self-interested defaults.

Demonstrating that enhanced deliberative thinking mediates the relationship between variety and unethical behavior also provides empirical evidence confirming largely untested intuitions about how Type 2 processes motivate desirable behavior. Interestingly, Kahneman (2011) discusses the nature of the CRT (our measure of deliberative thinking in Study 3a) and posits that it answers the question: "How closely does System 2 monitor the suggestions of System 1?" A person who responded 10 cents on the first question on the CRT was not being vigilant, and the Type 1 processes that would have led to the intuitive answer could have been overridden with a small investment of cognitive effort (Type 2). Importantly, those who give the intuitive answer have also missed an obvious social cue: they should have wondered why anyone would ask a question with such an obvious answer. As Kahneman argues, that we often fail to check our intuitive response to this question is remarkable, because the cost of checking is so low (only a few seconds of mental effort, which is different from intelligence). People who avoid the intuitive answer have more vigilant minds. It is this vigilant mind that we find mediates the mechanism between task variety and rule breaking.

Our studies also have implications for research on job or task design. Though task variety has received considerable attention in the organizational literature, primarily as a tool of improved job performance (Humphrey et al. 2007), most of this literature has focused on its motivational benefits (Hackman and Oldham 1976) through increased employee engagement (Herzberg 1968; Roethlisberger and Dickson 1934). However, this work mainly conceptualizes task variety in term of the range of skills required by one's job rather than in terms of how one's tasks at work are organized. Our studies elevate the importance of a more subtle (and potentially more practical) type of variety, by showing that the cognitive switching that varied task sequencing requires supports less rule breaking and more ethical behavior. Our results show that task variety can have cognitive (as well as motivational) benefits, with organizationally meaningful behavioral implications.

### **Future Research Directions**

Our findings suggest many potentially fruitful research directions.<sup>10</sup> One important potential boundary condition for our effects concerns the nature of the tasks. In both our field and lab contexts, individuals engaged in relatively simple tasks. It is possible that with more complex tasks, the cost of switching associated with high variety may be too depleting (Hamilton et al. 2011), hampering ethical behavior (as well as performance). There may be an 'ideal' amount of variety that best dampens the tendency to make self-interested hedonic choices, but these gains may begin to decline after a certain point. This idea is consistent with the literature on job complexity, which shows that the benefits of an "enhanced" job begin to backfire above a certain level of complexity (Xie and Johns 1995). Task complexity is thus an important moderator for future research to explore.

Related to the idea that task complexity may moderate our effects, it is interesting to note that when there was no opportunity to cheat in our lab experiments, there was no difference in performance between the high and low variety conditions, as the literature on switching costs would predict (Shallice and Burgess, 1993). It might be that switching costs only manifest with switches among more different or

---

<sup>10</sup> We are grateful to our three anonymous reviewers and the senior editor for raising several of these points.

more complex tasks than our experiments required. It could also be that the deliberation activated by the high variety condition made up for these switching costs to some extent. This suggests, though, that more research is needed in order to understand the conditions under which the costs and benefits associated with task switching are mitigated or amplified.

Another interesting question involves the extent to which the variety is predictable or expected. One feature of our sequential task variety manipulation is that participants may have come to expect a certain pattern of tasks (i.e., an ABCABC pattern). And though the workers in our field study could not have known the specific level of task variety they would experience on any given day, these workers also would have known to expect an unpredictable sequence of tasks. Though the consistent pattern of results across our field and lab is encouraging, given the other differences between these settings, we encourage future research to explore how the predictability of subtask switching may affect the relationship between task variety and rule breaking.

A third important question is whether the effects of variety on rule breaking are limited to violations of certain types of rules. We acknowledge that we focused on violations of rules intended to constrain or direct individual behavior so that it meets organizational or ethical standards. While this type of rule breaking is extremely common, and of serious concern to most organizations (e.g., Dai et al. 2015), it is not the only type of rule breaking. Indeed, individuals sometimes break the rules in order to benefit the organization or its stakeholders (Dahling et al. 2012; Morrison 2006), and there are cases in which rule breaking is associated with more creativity and innovation (Gino and Wiltermuth 2014). It would be interesting to examine the role of task variety in these different kinds of rule breaking.

Relatedly, Gino and Wiltermuth (2014) connected creativity to unethical behavior through the mechanism of cognitive flexibility. This raises a question about the overlap (or lack of) between cognitive flexibility and deliberation. Our view is that one can be reflective while still being rule-following. At the same time one can be intuitively flexible (and therefore creative). In fact, Evans and Stanovich (2013) discuss how an associative mindset (related to cognitive flexibility) is a more frequent correlate of Type 1 processes, while rule-based thinking is a more frequently correlate of Type 2 processes. Thus, while we

see deliberation as distinct from cognitive flexibility, future research could still benefit from examining how these constructs are related.

It is important to note that the rule breaking in both our field and experimental studies was relatively invisible (i.e. undetected by management or experimenters), and did not elicit severe consequences if caught. Making the ethical behavior more visible or more consequential (either in terms of formal [punishment] or informal [social ostracism] sanctions) would likely enhance the salience of moral standards. Employees might be more motivated to behave ethically in these cases. Indeed, one likely reason why bank employees do not violate corporate lunch break policy on a daily basis is because doing so would put them at risk of social sanctions from their colleagues for shirking. It would be interesting to examine how belonging to different organizational groups might be another mitigating boundary condition. When the unethical behavior becomes more consequential – either for oneself or for a group that one belongs to – motivational and cognitive accounts are likely to interact. Looking at the visibility, severity, and sociality of unethical behavior as additional potential moderators of these effects (perhaps interacted with motivational explanations) are important avenues for future research.

We note that across our experimental studies we found no difference across conditions in terms of the time it took participants to answer questions. One might have predicted that our variety manipulation (that induces Type 2 processing) would result in slower response times (compared to the low variety condition). It might well be that the nature of our task (relatively simple problems for which the average response time was around 20 seconds, already quite fast), in addition to the upper bound we gave participants (only 30 seconds for each question) resulted in both floor and ceiling effects that did not allow us to detect differences in response times as a function of the variety conditions. However, we can only speculate about this interpretation of this null effect, and we encourage future studies, with possibly more complex tasks, to explore the effect of sequential variety on task completion times, especially since these can have important efficiency implications in organizations.

Finally, while our results provide evidence that deliberative thinking mediates the relationship between task variety and rule breaking, they do not offer evidence as to *why* this link exists. One possible

explanation remains motivational. Although Studies 3a and 3b provided some evidence that intrinsic motivation is not the cause of our effects *per se*, it is possible that the way task variety shifts how individuals view their goals in a given context could be leading to the results we observe. Work on goal framing (Lindenberg and Steg 2007) suggests that individuals have three goal frames, each of which leads to different actions: (1) a hedonic goal frame encourages individuals to seek direct pleasure; (2) a gain goal frame makes people sensitive to changes in their personal resources, and encourages them to focus on accruing gains or preventing losses; and finally (3) a normative goal frame encourages people to focus on the appropriateness, or the “oughts”, of their actions. Hedonic, gain, and normative goal frames have short, intermediate, and long-term time horizons, respectively.

One can think of individuals’ default goal frame as a hedonic one, which is less deliberative, and more associated with meeting one’s immediate self-serving needs. The rule breaking we focus on in our studies involves meeting immediate hedonic needs. The deliberative thinking triggered by high levels of variety likely moves one out of a hedonic frame, to one of the two alternative frames. Variety might move one to a gain frame, which might motivate one to avoid being caught breaking the rules, or a normative frame, which would motivate one to do what they “ought” to and follow the rules. In either of these cases, the deliberation that variety elicits makes people more vigilant about the actions they should take, increasing their motivation to avoid rule breaking.<sup>11</sup> We encourage future research to examine goal frames as a potential micro-mediator of our effects, by connecting deliberative thinking to different goals in general, and to the three goal frames in particular. Especially given research that has critiqued the extent to which Type 2 processes improve ethical behavior (Haidt and Kesebir 2010; Zhong 2011), we need to better understand when and why Type 1 and Type 2 processes will lead to better (or worse) ethical behavior. We hope that our current study contributes to this ongoing and insightful conversation.

## **Practical Implications**

Our work also has important practical implications. First, rule compliance in organizations is

---

<sup>11</sup> We are grateful to the senior editor for directing us to this literature and to a reviewer for further sharpening the contribution.



often lower than socially desirable (Martin et al. 2013). By showing that sequencing work in a more interleaved way drives rule compliance, we identify task variety as a key lever that organizations can deploy to address this challenge. Second, as we noted above, while increasing skill variety in lower level positions is often unfeasible, organizations may have more latitude to vary the sequencing of employees' tasks. Employees often have multiple tasks that may not have a required sequence, but organizations may be able to structure them so that they accrue the benefits identified in this paper.

However, if switching tasks makes the work potentially harder (or is perceived by employees as harder, as our studies did not show hampered performance due to the increased sequential variety), employees with discretion over their task sequencing may choose to order their tasks in an easier way, undermining their deliberative abilities, and, as a result, making them less rule compliant. This elevates the importance of providing training to encourage employees to vary how they organize their work, and incenting them to make this choice. Indeed, ethics training that understands moral behavior as a conscious choice has been criticized as ineffective (Tenbrunsel and Messick 2004). It may be more effective to develop more ethical employees through more indirect routes.

## **Conclusion**

Our studies examine the influence of task design on rule breaking behavior. This lever provides an important avenue for organizations to explore, as it may be that simple changes in employees' exogenous task design can enhance the likelihood that they will comply with rules intended to constrain hedonic self-interested behavior, without detracting from their performance on more focal tasks. By showing that task variety activates deliberative thinking, which, in turn, helps prevent employees from lapsing into ethically compromised behavior, our studies also respond to Grant's calls to examine how task design may influence ethical behavior (Grant 2007; Grant et al. 2010). Our results are encouraging in showing that organizations can positively influence individuals' ethical behavior by making simple changes to their job design. Changing individuals' motivation may well be a much harder endeavor.

### References

- Allport, D.A., E.A. Styles, S. Hsieh. 1994. Shifting intentional set: Exploring the dynamic control of tasks. C. Umiltà, M. Moscovitch, eds. *Attention and performance xv*. MIT Press, Cambridge, MA, 421-452.
- Alter, A.L., D.M. Oppenheimer, N. Epley, R.N. Eyre. 2007. Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General* 136: 569-576.
- Babcock, L., G. Loewenstein, S. Issacharoff, C. Camerer. 1995. Biased judgments of fairness in bargaining. *Am. Econ. Rev.* 85: 1337-1343.
- Batson, C.D., D. Kobrynowicz, J.L. Dinnerstein, H.C. Kampf, A.D. Wilson. 1997. In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology* 72: 1335-1348.
- Baumeister, R.E., T.F. Heatherton, D.M. Tice. 1994. *Losing control: How and why people fail at self-control*. Academic Press, San Diego, CA.
- Bazerman, M.H., A.E. Tenbrunsel. 2011. *Blind spots: Why we fail to do what's right and what to do about it*. Princeton University Press, Princeton, NJ.
- Bennett, R.J., S.L. Robinson. 2000. Development of a measure of workplace deviance. *Journal of Applied Psychology* 85: 349-360.
- Blau, P.M. 1964. *Exchange and power in social life*. J. Wiley, New York.
- Buhrmester, M., T. Kwang, S.D. Gosling. 2011. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* 6: 3-5.
- Chance, Z., M.I. Norton, F. Gino, D. Ariely. 2011. Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences* 108: (Supplement 3) 15655-15959.
- Chugh, D., M.H. Bazerman, M.R. Banaji. 2005. Bounded ethicality as a psychological barrier to recognizing conflicts of interest. D.A. Moore, D.M. Cain, G. Loewenstein, M.H. Bazerman, eds. *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy*. Cambridge University Press, New York, 74-95.

- Chugh, D., M.C. Kern, H.Y. Kim. 2014. When personality matters: Cognitive load moderates the effect of personality on ethics. Working Paper, New York University, Stern School of Business.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*, 2nd ed. Erlbaum, Hillsdale, NJ.
- Cokely, E.T. 2007. Adaptive mechanisms of superior judgment under uncertainty: Rational choices from simple heuristics and elaborative strategies.
- Cortina, J.M., R.S. Landis. 2009. When small effects tell a big story, and when large effect sizes don't. C.E. Lance, R.J. Vandenberg, eds. *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*. Routledge, New York, 287-308.
- Dahling, J.J., S.L. Chau, D.M. Mayer, J.B. Gregory. 2012. Breaking rules for the right reasons? An investigation of pro-social rule breaking. *Journal of Organizational Behavior* 33: 21-42.
- Dai, H., K.L. Milkman, D.A. Hofmann, B.R. Staats. 2015. The impact of time at work and time off from work on rule compliance: The case of hand hygiene in health care. *Journal of Applied Psychology* 100: 846-862.
- Deci, E.L., R.M. Ryan. 1985. *Intrinsic motivation and self-determination in human behavior*. Plenum, New York.
- Desai, S., M. Kouchaki. In press. Moral symbols: A necklace of garlic against unethical requests *Acad. Management J.*
- Desai, S.D., M. Kouchaki. 2015. Work-report formats and overbilling: How unit-reporting vs. Cost-reporting increases accountability and decreases overbilling. *Organizational Behavior and Human Decision Processes* 130: 79-88.
- Diekmann, K.A., S.M. Samuels, L. Ross, M.H. Bazerman. 1997. Self-interest and fairness in problems of resource allocation: Allocators versus recipients. *Journal of Personality and Social Psychology* 72: 1061-1074.
- Diemand-Yauman, C., D.M. Oppenheimer, E.B. Vaughan. 2011. Fortune favors the (): Effects of disfluency on educational outcomes. *Cognition* 118: 111-115.

- Eisenberger, R., S. Armeli, B. Rexwinkel, P.D. Lynch, L. Rhoades. 2001. Reciprocation of perceived organizational support. *Journal of Applied Psychology* 86: 42-51.
- Emerson, R.M. 1976. Social exchange theory. *Annu. Rev. Soc.* 2: 335-362.
- Epley, N., E.M. Caruso. 2004. Egocentric ethics. *Social Justice Research* 17: 171-187.
- Evans, J.S.B.T., K.E. Stanovich. 2013. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science* 8: 223-241.
- Fitzgerald, L.F., F. Drasgow, C.I. Hulin, M.J. Gelfand, V.J. Magley. 1997. Antecedents and consequences of sexual harassment in organizations: A test of an integrated model. *Journal of Applied Psychology* 82: 578-589.
- Frederick, S. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* 19: 25-42.
- Gino, F., M.E. Schweitzer, N.L. Mead, D. Ariely. 2011. Unable to resist temptation: How self-control depletion promotes unethical decision making. *Organizational Behavior and Human Decision Processes* 115: 191-203.
- Gino, F., S.S. Wiltermuth. 2014. Evil genius? How dishonesty can lead to greater creativity. *Psychological Science*.
- Gouldner, A.W. 1960. The norm of reciprocity: A preliminary statement. *American Sociological Review* 25: 161-178.
- Grant, A. 2007. Relational job design and the motivation to make a prosocial difference. *Acad. Management Rev.* 32: 393-417.
- Grant, A.M., Y. Fried, S.K. Parker, M. Frese. 2010. Putting job design in context: Introduction to the special issue. *Journal of Organizational Behavior* 31: 145-157.
- Grant, A.M., D.A. Hofmann. 2011. It's not all about me: Motivating hand hygiene among health care professionals by focusing on patients. *Psychological Science* 22: 1494-1499.
- Greenberg, J. 1997. The steal motive: Managing the social determinants of employee theft. R.A. Giacalone, J. Greenberg, eds. *Antisocial behavior in organizations*. Sage Publications, Thousand Oaks, CA, 85-108.

- Greenwald, A.G. 1980. The totalitarian ego: Fabrication and revision of personal history. *American Psychologist* 35: 603-618.
- Griffin, R.W., Y.P. Lopez. 2005. "Bad behavior" in organizations: A review and typology for future research. *Journal of Management* 31: 988-1005.
- Gunia, B.C., L. Wang, L. Huang, J. Wang, J.K. Murnighan. 2012. Contemplation and conversation: Subtle influences on moral decision making. *Acad. Management J.* 55: 13-33.
- Hackman, J.R., G.R. Oldham. 1975. Development of the job diagnostic survey. *Journal of Applied Psychology* 60: 159-170.
- Hackman, J.R., G.R. Oldham. 1976. Motivation through the design of work: Test of a theory. *Organizational Behavior & Human Performance* 16: 250-279.
- Hackman, J.R., G.R. Oldham. 1980. *Work redesign*. Addison-Wesley, Reading, Mass.
- Haidt, J., S. Kesebir. 2010. Morality. S. Fiske, D. Gilbert, G. Lindzey, eds. *Handbook of social psychology*, 5th ed. Wiley, Hoboken, NJ, 797-832.
- Haley, K.J., D.M.T. Fessler. 2005. Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior* 26: 245-256.
- Hamilton, R., K.D. Vohs, A.-L. Sellier, T. Meyvis. 2011. Being of two minds: Switching mindsets exhausts self-regulatory resources. *Organizational Behavior and Human Decision Processes* 115: 13-24.
- Harris, J., P. Bromiley. 2007. Incentives to cheat: The influence of executive compensation and firm performance on financial misrepresentation. *Organ. Sci.* 18: 350-367.
- Harrison, D.A., K.J. Klein. 2007. What's the difference? Diversity constructs as separation, variety, or diversity in organizations. *Acad. Management Rev.* 32: 1199-1228.
- Hauser, D.J., N. Schwarz. 2016. Attentive turkers: Mturk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods* 48: 400-407.
- Hayes, A.F. 2013. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford, New York.

- Hernandez, I., J.L. Preston. 2013. Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology* 49: 178-182.
- Herzberg, F. 1968. One more time: How do you motivate employees? *Harvard Business Review* 46: 53-62.
- Hofmann, D.A., A. Stetzer. 1996. A cross-level investigation of factors influencing unsafe behaviors and accidents. *Personnel Psychology* 49: 307-339.
- Hogan, R. 1973. Moral conduct and moral character: A psychological perspective. *Psychological Bulletin* 79: 217-232.
- Humphrey, S.E., J.D. Nahrgang, F.P. Morgeson. 2007. Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. *Journal of Applied Psychology* 92: 1332-1356.
- Ilgén, D.R., J.R. Hollenbeck. 1991. The structure of work: Job design and roles. M.D. Dunnette, L.E. Hough, eds. *Handbook of industrial and organizational psychology*, 2nd ed. Consulting Psychologists Press, Palo Alto, CA, 165-207.
- James, W. 1890/1950. *The principles of psychology*. Dover, New York.
- Jersild, A.T. 1927. Mental set and shift. *Archives of psychology* 89: 5-89.
- Kahneman, D. 2003. Maps of bounded rationality: Psychology for behavioral economics. *Am. Econ. Rev.* 93: 1449-1475.
- Kahneman, D. 2011. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York.
- Kahneman, D., S. Frederick. 2005. A model of heuristic judgment. *The Cambridge handbook of thinking and reasoning* 267-293.
- Kahneman, D., D. Lovallo. 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Sci.* 39: 17-31.
- Keren, G., Y. Schul. 2009. Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science* 4: 533-550.
- Kidwell, R.E., N. Bennett. 1993. Employee propensity to withhold effort: A conceptual model to intersect three avenues of research. *Acad. Management Rev.* 18: 429-456.

- Kish-Gephart, J.J., D.A. Harrison, L.K. Treviño. 2010. Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology* 95: 1-31.
- Knoch, D., A. Pascual-Leone, K. Meyer, V. Treyer, E. Fehr. 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314: 829-832.
- Kornell, N., R.A. Bjork. 2008. Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science* 19: 585-592.
- Kruglanski, A.W., G. Gigerenzer. 2011. Intuitive and deliberate judgments are based on common principles. *Psychological Review* 118: 97-109.
- Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108: 480-498.
- Larrick, R.P. 2008. Debiasing. D.J. Koehler, N. Harvey, eds. *Blackwell handbook of judgment and decision making*. Blackwell, Oxford, UK, 316-338.
- LeBreton, J.M., J.R.D. Burgess, R.B. Kaiser, E.K. Atchley, L.R. James. 2003. The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods* 6: 80-128.
- Lehman, D.W., R. Ramanujam. 2009. Selectivity in organizational rule violations. *Acad. Management Rev.* 34: 643-657.
- Lerner, J.S., P.E. Tetlock. 1999. Accounting for the effects of accountability. *Psychological Bulletin* 125: 255-275.
- Lindenberg, S., L. Steg. 2007. Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of Social Issues* 63: 117-137.
- MacKinnon, D.P., C.M. Lockwood, J.M. Hoffman, S.G. West, V. Sheets. 2002. A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods* 7: 83-104.
- Martin, A.W., S.H. Lopez, V.J. Roscigno, R. Hodson. 2013. Against the rules: Synthesizing types and processes of bureaucratic rule-breaking. *Acad. Management Rev.* 38: 550-574.

- Martin, S.R. In press. Stories about values and valuable stories: A field experiment of the power of narratives to shape newcomers' actions. *Acad. Management J.*
- McAuley, E., T. Duncan, V. Tammen. 1989. Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport* 60.
- Mead, N.L., R.F. Baumeister, F. Gino, M.E. Schweitzer, D. Ariely. 2009. Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology* 45: 594-597.
- Messick, D.M., K.P. Sentis. 1979. Fairness and preference. *Journal of Experimental Social Psychology* 15: 418-434.
- Milkman, K.L., D. Chugh, M.H. Bazerman. 2009. How can decision making be improved? *Perspectives on Psychological Science* 4: 379-383.
- Monsell, S. 2003. Task switching. *Trends in Cognitive Science* 7: 134-140.
- Moore, C., F. Gino. 2015. Approach, ability, aftermath: A psychological process framework of unethical behavior at work. *Acad. Management Annals* 9: 235-289.
- Moore, D.A., G. Loewenstein. 2004. Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research* 17: 189-202.
- Morrison, E.W. 2006. Doing the job well: An investigation of pro-social rule breaking. *Journal of Management* 32: 5-28.
- Mussweiler, T., F. Strack, T. Pfeiffer. 2000. Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Pers. Soc. Psychol. Bull.* 26: 1142-1150.
- Osman, M. 2004. An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review* 11: 988-1010.
- Paolacci, G., J. Chandler, P. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5: 411-419.



- Pentland, B.T. 2003. Conceptualizing and measuring variety in the execution of organizational work processes. *Management Sci.* 49: 857-870.
- Pérez-Peña, R. 2013, February 13. Students disciplined in Harvard scandal. *New York Times*.
- Pierce, J.L., R.B. Dunham. 1976. Task design: A literature review. *Acad. Management Rev.* 1: 83-97.
- Prentice, D.A., D.T. Miller. 1992. When small effects are impressive. *Psychological Bulletin* 112: 160-162.
- Reynolds, S.J., K. Leavitt, K.A. DeCelles. 2010. Automatic ethics: The effects of implicit assumptions and contextual cues on moral behavior. *Journal of Applied Psychology* 95: 752-760.
- Robinson, S.L., R.J. Bennett. 1995. A typology of deviant workplace behaviors: A multidimensional scaling study. *Acad. Management J.* 38: 555-572.
- Robinson, S.L., R.J. Bennett. 1997. Workplace deviance: Its definition, its manifestations, and its causes. R.J. Lewicki, R.J. Bies, B. Sheppard, eds. *Research on negotiation in organizations*. Elsevier Science/JAI Press, Greenwich, CT, 3-27.
- Roethlisberger, F.J., W.J. Dickson. 1934. *Management and the worker*. Harvard University Press, Cambridge, MA.
- Rosnow, R.L., R. Rosenthal. 2003. Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 57: 221-237.
- Roy, D. 1959. Banana time: Job satisfaction and informal interaction. *Human Organization* 18: 158-168.
- Rucker, D.D., K.J. Preacher, Z.L. Tormala, R.E. Petty. 2011. Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass* 5: 359-371.
- Sah, S., G. Loewenstein. 2014. Nothing to declare: Mandatory and voluntary disclosure leads advisors to avoid conflicts of interest. *Psychological Science* 25: 575-584.
- Salaz, J.J. 2002. The house rules: Autonomy and interests among service workers in the contemporary casino industry. *Work and Occupations* 29: 394-427.

- Scott, B.A., J.A. Colquitt, E.L. Paddock. 2009. An actor-focused model of justice rule adherence and violation: The role of managerial motives and discretion. *Journal of Applied Psychology* 94: 756-769.
- Shallice, T., P. Burgess. 1993. Supervisory control of action and thought selection. A. Baddeley, L. Weiskrantz, eds. *Attention: Selection, awareness, and control*. Oxford University Press, Oxford, England, 171-187.
- Shalvi, S., O. Eldar, Y. Bereby-Meyer. 2012. Honesty requires time (and lack of justifications). *Psychological Science* 23: 1264-1270.
- Shrout, P.E., N. Bolger. 2002. Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods* 7: 422-445.
- Shu, L.L., F. Gino. 2012. Sweeping dishonesty under the rug: How unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology* 102: 1164-1177.
- Simon, H.A. 1982. *Models of bounded rationality*. MIT Press, Cambridge, MA.
- Simon, H.A. 1983. *Reasons in human affairs*. Stanford University Press, Stanford.
- Staats, B.R., F. Gino. 2012. Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Sci.* 58: 1141-1159.
- Stanovich, K.E., R.F. West. 2000. Individual differences in reasoning: Implications for the rationality debate? *Behav Brain Sci* 23: 645-665.
- Taylor, K., D. Rohrer. 2010. The effects of interleaved practice. *Applied Cognitive Psychology* 24: 837-848.
- Tenbrunsel, A.E. 1998. Misrepresentation and expectations of misrepresentation in an ethical dilemma: The role of incentives and temptation. *Acad. Management J.* 41: 330-339.
- Tenbrunsel, A.E., D.M. Messick. 2004. Ethical fading: The role of self-deception in unethical behaviour. *Social Justice Research* 17: 223-236.
- Toplak, M., R. West, K. Stanovich. 2011. The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Mem Cogn* 39: 1275-1289.

- Treviño, L.K., G.R. Weaver, S.J. Reynolds. 2006. Behavioral ethics in organizations: A review. *Journal of Management* 32: 951-990.
- Tversky, A., D. Kahneman. 1974. Judgement under uncertainty: Heuristics and biases. *Science* 185: 1124-1131.
- Tyler, T.R., S.L. Blader. 2005. Can businesses effectively regulate employee conduct? The antecedents of rule following in work settings. *Acad. Management J.* 48: 1143-1158.
- Vardi, Y., Y. Weiner. 1996. Misbehavior in organizations: A motivational framework. *Organ. Sci.* 7: 151-165.
- Vohs, K.D., J.W. Schooler. 2008. The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science* 19: 49-54.
- von Hippel, W., J.L. Lakin, R.J. Shakarchi. 2005. Individual differences in motivated social cognition: The case of self-serving information processing. *Pers. Soc. Psychol. Bull.* 31: 1347-1357.
- Watson, D., L.A. Clark, A. Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of Personality and Social Psychology* 54: 1063-1070.
- Wenzel, M. 2005. Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology* 26: 862-883.
- Xie, J.L., G. Johns. 1995. Job scope and stress: Can job scope be too high? *Acad. Management J.* 38: 1288-1309.
- Zhong, C.-B. 2011. The ethical dangers of deliberative decision making. *Admin. Sci. Quart.* 56: 1-25.

## Tables & Figures

**Table 1.** *Description of Steps in the Process (Study 1)*

Name	Separate Stages	Description
Custodian	2	Comparison of scanned document to actual document to confirm scan is of sufficient quality (done for both initial and additional data).
Doc tagging	2	Marking of images on document scans for subsequent steps (done for both initial and additional data).
Application capture	4	Data entry using applications (two separate steps each for initial and additional application capture).
Preliminary information	2	Additional data entry from application forms (two separate steps inputting data from separate pages).
Credit check	2	A two step process – in step 1 a credit check is requested and in step 2 the data is entered from the check.
Income tax	2	A two step process – in step 1 tax verification is requested and in step 2 the data is entered from the verification.
Real estate	2	A two step process – in step 1 a real estate appraisal is requested and in step 2 the data is entered from the appraisal.
Credit approval	1	An individual either accepts, rejects, or routes the application to a credit expert, depending on the underwriting criteria.

**Table 2.** *Descriptive Statistics and Correlations (Study 1)*

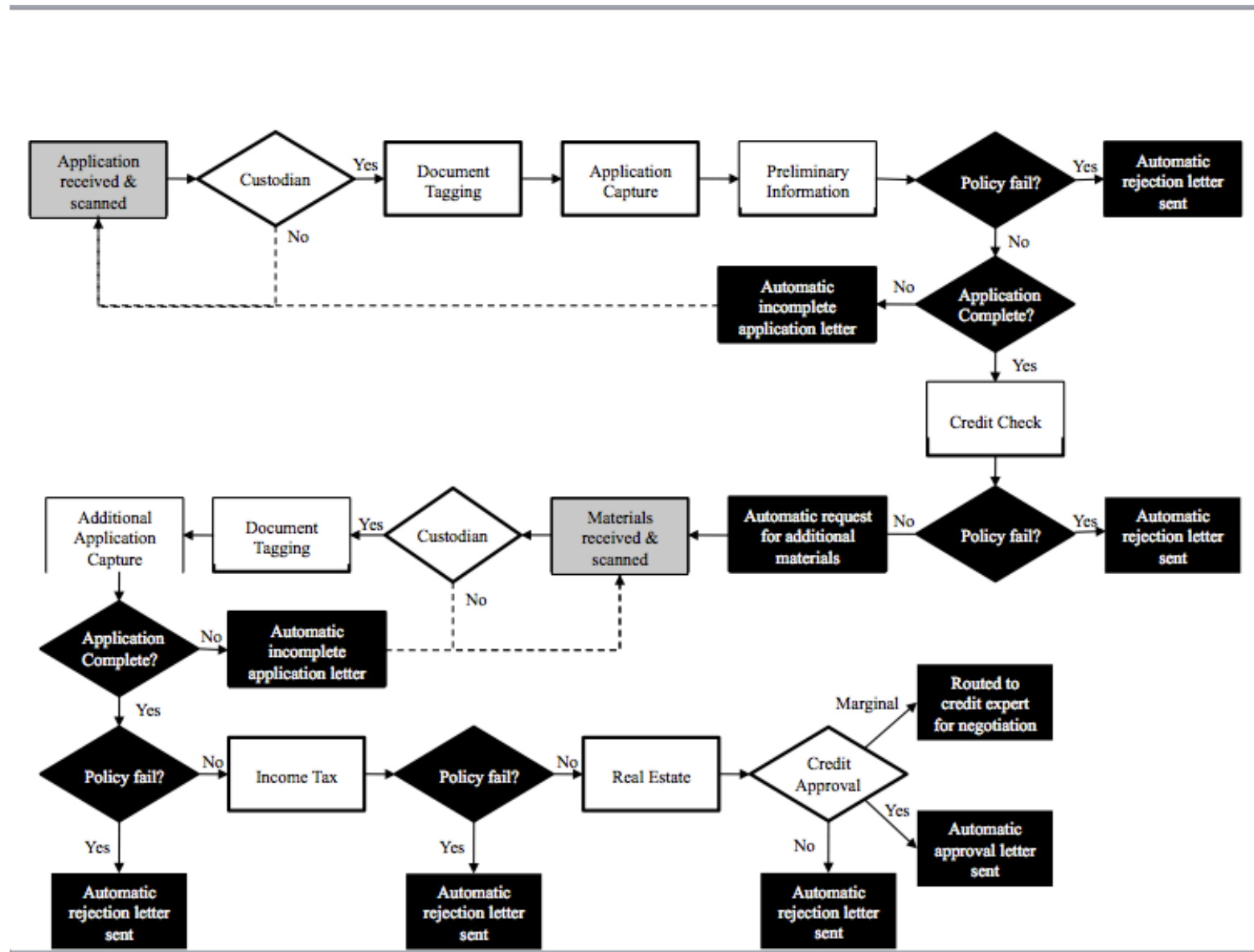
Variable	Mean	$\sigma$	1	2	3	4	5	6
1. Long lunch	0.16	0.37						
2. Length of long lunch	7.53	23.31	0.74					
3. Pre-lunch task variety	0.72	1.51	-0.12	-0.10				
4. Pre-lunch task type	0.14	0.20	-0.10	-0.08	0.65			
5. Pre-lunch cumulative volume	62.84	78.21	-0.14	-0.13	0.40	0.24		
6. Prior day cumulative volume	4,501.4	4,231.8	-0.04	-0.02	0.24	0.31	0.19	
7. Pre-lunch workload	0.48	0.31	-0.10	-0.12	-0.09	-0.14	-0.09	-0.16

Note:  $N = 16,804$ . All correlations are significant at less than 5%.

**Table 3.** *Results of Logistic Regression Predicting Taking a Lunch Break that Exceeds Regulations (Study 1)*

	<b>Logistic Regression on Long Lunch</b>	<b>Tobit Regression on Length of Long Lunch</b>	<b>OLS on Length of Long Lunch &gt; 0</b>
	(1)	(2)	(3)
Pre-lunch sequential variety	-0.2569*** (0.06952)	-10.223*** (1.6436)	-6.8705* (2.7316)
Pre-lunch total variety	0.06819 (0.2403)	-4.3208 (7.9815)	6.2120 (9.5054)
Pre-lunch cumulative volume	-0.003354*** (8.385e-04)	-0.1737*** (0.02112)	-0.09525*** (0.01677)
Prior day cumulative volume	5.062e-05* (2.460e-05)	7.976e-04 (5.569e-04)	-1.261e-04 (7.149e-04)
Pre-lunch workload	-1.3401*** (0.2174)	-56.611*** (3.5216)	-21.862*** (4.2656)
Day indicators	Included	Not Included	Included
Individual Controls	Fixed Effects	Random Effects	Fixed Effects
Constant	1.5308* (0.6602)	-14.573 (20.981)	25.010** (9.4641)
Observations	16,084	17,161	2,577
Number of Individuals	102	110	102

Notes. \*, \*\* and \*\*\* denote significance at the 5%, 1% and 0.1% levels, respectively. Column 1 is a logistic regression model with heteroskedasticity robust standard errors clustered by worker. Column 2 is a Tobit regression model and Column 3 is a fixed effects regression model with standard errors clustered by individual.

**Figure 1.** *Process Flow Diagram for Home Loan Mortgage Process (Study 1)*

**Figure 2.** *Percentage of questions solved correctly ( $\pm 1SE$ ) as a function of task variety, within the opportunity to cheat and the no opportunity to cheat conditions (Study 2)*

