

Cohn, Alain; Maréchal, Michel André

**Working Paper**

## Laboratory measure of cheating predicts school misconduct

Working Paper, No. 205

**Provided in Cooperation with:**

Department of Economics, University of Zurich

*Suggested Citation:* Cohn, Alain; Maréchal, Michel André (2016) : Laboratory measure of cheating predicts school misconduct, Working Paper, No. 205, University of Zurich, Department of Economics, Zurich,  
<https://doi.org/10.5167/uzh-113256>

This Version is available at:

<https://hdl.handle.net/10419/162414>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



**University of  
Zurich** <sup>UZH</sup>

University of Zurich  
Department of Economics

Working Paper Series  
ISSN 1664-7041 (print)  
ISSN 1664-705X (online)

---

Working Paper No. 205

# **Laboratory Measure of Cheating Predicts School Misconduct**

Alain Cohn and Michel André Maréchal

Revised version, February 2016

---

# Laboratory Measure of Cheating Predicts School Misconduct\*

Alain Cohn and Michel André Maréchal

February 2016

## Abstract

We conducted an experiment with middle and high school students to test the external validity of a common laboratory measure of cheating. Subjects performed several coin tosses and earned money depending on the outcomes they reported. Because the coin tosses were not monitored, subjects faced a financial incentive to misreport their outcomes without having to worry about getting caught. We linked the responses from the lab experiment with three measures of school misconduct (i.e., disruptiveness in class, failure to complete homework, and absenteeism). The findings show that cheating in the lab significantly predicts misbehavior in school, suggesting that the experimental measure of cheating generalizes qualitatively to naturally occurring environments.

**JEL classification:** C93, K42

**Keywords:** Cheating, Honesty, Experiment, External Validity, School Misbehavior.

---

\* Cohn: Booth School of Business, University of Chicago, 5807 S Woodlawn Ave, Chicago, IL 60637, U.S., email: alain.cohn(at)chicagobooth.edu. Maréchal: University of Zurich, Department of Economics, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland, email: michel.marechal(at)econ.uzh.ch. We thank Jonas Addor and Joël Farronato for excellent research assistance and David Tannenbaum and John List for helpful comments. Financial support from the Gottlieb Duttweiler Institute is gratefully acknowledged. Correspondence to: Cohn and Maréchal.

# 1 Introduction

Cheating, misconduct and other forms of rule violating behavior are pervasive problems in many important areas of social and economic life. Examples range from scandals in the business world (e.g., Volkswagen’s recent emission fraud or interest and exchange rate manipulations in the financial industry) to rigged sport competitions (Duggan et al. 2002), rampant corruption in developing countries (Pande and Olken 2012; Banerjee et al. 2013), and student and teacher cheating (Jacob and Levitt 2003; Levitt and Lin 2015).

Given the prevalence and cost of dishonesty to society, a rapidly growing literature on behavioral ethics has emerged with the aim to provide a better understanding of the determinants of lying, cheating, and stealing (see Ariely 2012; Irlenbusch and Villeval 2015, and Shalvi et al. 2015 for recent reviews). Due to its clandestine nature, dishonest behavior is typically difficult to measure reliably using observational field data (Zitzewitz 2012). As a consequence, the majority of empirical findings originates from controlled laboratory environments.<sup>1</sup>

A widely used experimental paradigm to measure cheating is to instruct subjects to perform a simple task of chance (e.g., flipping coins or rolling dice) and asking them to report their outcomes. Because the actual outcomes are not observed by the experimenter and only certain outcomes are rewarded, subjects face the temptation to increase their earnings by misreporting their outcomes without any risk of getting caught (e.g., Bucciol and Piovesan 2011; Shalvi et al. 2011; Fischbacher and Föllmi-Heusi 2013; Cohn et al. 2014; Abeler et al. 2014).<sup>2</sup> Although cheating cannot be detected at the individual level, researchers can measure cheating at the group level as the true distribution of the underlying random process is known. Moreover, because higher earnings are less likely

---

<sup>1</sup>See Pierce and Balasubramanian (2015) for a survey of the literature on dishonest behavior that uses observational data and field experiments.

<sup>2</sup>Another common approach are interactive sender-receiver games where senders can increase their earnings by sending deceptive messages to the receiver (e.g., Gneezy 2005; Sutter 2009).

to be the result of chance, earnings claimed by individual subjects can serve as a proxy for their cheating behavior. While this paradigm has been used extensively to study the determinants of dishonesty, the extent to which the insights gained from the lab can be extrapolated to naturally occurring environments remains unclear. Common objections to the generalizability of lab experiments are that subjects make low-stakes decisions in artificial environments and that they know their behavior is being recorded and analyzed (Levitt and List 2007; Falk and Heckman 2009).

In this paper we investigate whether cheating in the lab predicts rule violating behavior in the field. To this end, we matched a common laboratory measure of cheating with teacher evaluations of students' misbehavior in school. We experimentally measured cheating by asking the students to toss ten coins in private and report their outcomes. Students only received financial rewards when reporting "heads," and thus had a financial incentive to misreport their outcomes for unsuccessful coin flips. Our measures of school misbehavior are based on the US National Education Longitudinal Survey. Specifically, we asked teachers to assess their students along three dimensions: disruptiveness in class, non-completion of homework, and absenteeism. These measures of school misconduct are important as they have been shown to reliably predict future educational achievement and labor market outcomes (Segal 2013).<sup>3</sup>

We found a positive and significant correlation between the laboratory measure of cheating and students' misbehavior in school. The difference in school misbehavior between a student who reported a successful coin toss on every single trial and a student who reported a successful outcome in 50 percent of the cases is the same as between two students with a difference in cognitive ability of 2.7 standard deviations. The relationship between the laboratory measure of cheating and school misbehavior remains strong after adding controls for age, gender, nationality, school level, and parental education.

---

<sup>3</sup>Disruptive and noncompliant behavior in school also seem to matter for students' current academic performance as we found negative and significant correlations between students' self-reported grade point average (GPA) and the three measures of school misbehavior (disruptiveness:  $p = 0.001$ , homework:  $p < 0.001$ , absenteeism:  $p = 0.002$ , Spearman tests).

Moreover, we found that behavior in the coin tossing task explains about one-fifth of the gender gap in school misbehavior. Together, these results suggest that the cheating paradigm from the lab provides an externally valid measure of rule violating behavior in the field.

Our paper contributes to several strands of the literature. First, a growing number of studies combines lab and field data from the *same* subjects to examine the external validity of laboratory measures of behavior.<sup>4</sup> For example, Karlan (2005) found that second-mover behavior in a trust game correlates with the likelihood of loan repayment among participants of a microcredit program in Peru.<sup>5</sup> Using experimental measures of present bias, Sutter et al. (2013) show that more impatient children and adolescents are more likely to buy alcohol and cigarettes, are more likely to be obese, and are less likely to save money.<sup>6</sup> Our findings suggest that cheating in the lab provides a reliable indicator of rule violating behavior in the field. Two other studies analyzed the relationship between rule violation in the lab and the field, but they both used rather unusual participants drawn from the extreme ends of the honesty distribution. Hanna and Wang (2014) examined cheating in a sample of government nurses in India. They found that nurses who cheated more in a dice task also tended to show up at work less often. Cohn et al. (2015*b*) conducted a coin tossing experiment with inmates from a maximum-security prison. They found a positive correlation between claimed earnings from the coin tosses and misconduct in prison (e.g., illegal drug possession or aggression against guards and other inmates). It is reassuring that these two papers arrive at the same conclusion as our study despite using different methods and subject pools.<sup>7</sup>

---

<sup>4</sup>See Camerer (2015) for an overview of experimental studies linking behavior in the lab and field.

<sup>5</sup>Benz and Meier (2008); Carpenter and Myers (2010); Fehr and Leibbrandt (2011); Burks et al. (2015), and Cohn et al. (2015*a*) provide further evidence for positive associations between lab and field measures of prosociality.

<sup>6</sup>Meier and Sprenger (2010) show that experimentally elicited present bias is a reliable predictor of credit card borrowing.

<sup>7</sup>Dai et al. (2015) recently presented the results from a dice-rolling experiment with public transport passengers and showed that the proportion of fully dishonest participants is higher among those who did not hold a valid ticket. List (2009) analyzed a subsample of 17 sellers from open air markets for which he observed lab and field behavior. He found that sellers who breached collusive agreements in

Second, our paper also speaks to a growing literature on the causes and consequences of school misconduct. For example, Segal (2013) shows that students misbehaving in eighth grade are almost three times less likely to finish high school and have almost 10 percent lower earnings as adults relative to non-disruptive students. Bertrand and Pan (2013) found that behavioral problems in school are more prevalent among boys, especially if they grow up in single-mother households. This finding may explain the widening gender gap in academic achievement in the United States and other developed countries (Goldin et al. 2006; Becker et al. 2010; Fortin et al. 2015). While school misconduct might affect labor market success through its negative impact on academic achievement (e.g., due to impaired learning), our findings suggest an additional, complementary channel, namely deficits in social and behavioral skills that are increasingly valued by employers (Deming 2015). Our paper therefore also links to an emerging literature on the role of behavioral or noncognitive skills in explaining educational achievement, labor market success, and other important adult outcomes (e.g., Bowles et al. 2001; Heckman et al. 2006, 2013; Campbell et al. 2014).

The remainder of this paper proceeds as follows. The next section presents the details of the experimental design. Section 3 discusses the results of the experiment and section 4 provides concluding remarks.

## 2 Design

We conducted a paper-and-pencil experiment with 162 students from eight classes in two Swiss public schools—one middle and one high school. Students were between 12 and 20 years old, and 43 percent of them were female. They were informed that their data will be treated confidentially. Although participation was voluntary, all students gave their consent to participate in the study. We ran the experiment simultaneously in all four classes at each school to avoid cross-talk between subjects. The experiment took

---

contextualized lab experiments were also more likely to do so in the field.

place in the classrooms in absence of the teachers. We set up a mobile laboratory and installed partition walls to shield subjects from sight and therefore ensure privacy.

In the first part of the experiment, we asked subjects some basic socio-demographic questions including age, gender, nationality, and parental education. In part two, we measured their cognitive ability using two short tests from Dohmen et al. (2010): the *word fluency test* and the *symbol-digit correspondence test*.<sup>8</sup> Both tests are related to working memory and processing speed, which is often part of the reason children thrive or struggle in school, but they measure distinct concepts of reasoning capability (Carroll 1993).<sup>9</sup> The word fluency test measures “crystallized intelligence” (ability to solve problems using knowledge and experience) by asking subjects to list as many different animals as possible within 90 seconds. Subjects received one point for each correct and unique animal named. The symbol-digit correspondence test measures “fluid intelligence” (innate ability to solve problems) and consists of decoding sets of unfamiliar symbols into single digits as fast as possible within 90 seconds. For each set, subjects had to write down the correct numbers under a grid of nine symbols using a predefined mapping between symbols and digits. Subjects earned one point for each correct symbol-digit pair.

The last part of the experiment comprised the coin tossing task—our laboratory measure of cheating. Subjects first opened an envelope containing ten coins, each worth 0.5 Swiss francs (about US \$0.55). Then, they were instructed to toss each coin in private and report their outcomes on paper. For every coin toss for which subjects reported the outcome “heads” they were allowed to keep the coin; they had to put the coin back into the envelope otherwise. Participants thus faced a financial incentive to cheat by misreporting the outcomes of their coin flips without any risk of getting caught. The stakes were considerable as the maximum possible payoff in this task corresponds roughly to half the amount students of similar age receive in pocket money every week (e.g., see

---

<sup>8</sup>The two cognitive ability tests are based on submodules of the Wechsler Adult Intelligence Scale (WAIS), one of the most frequently used intelligence tests.

<sup>9</sup>Test scores are positively correlated in our sample (Spearman’s  $\rho = 0.423, p < 0.001$ ).



`www.budgetberatung.ch`). After completing the coin tossing task, subjects were asked to put their envelope with the remaining coins into a container.

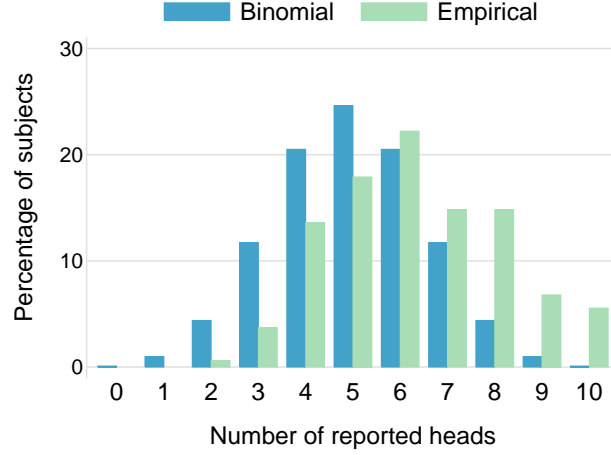
Teachers were asked to assess their students along three dimensions: disruptiveness in class, non-completion of homework, and absenteeism. For each item the teachers evaluated the students on a scale from “never misbehaves” ( $= 0$ ) to “always misbehaves” ( $= 6$ ). These measures of school misbehavior were inspired by the US National Educational Longitudinal Survey—a study that followed a nationally representative sample of more than 20,000 students over several years. We chose these measures of school misbehavior as they have been shown to reliably predict future educational achievement and labor market outcomes (Segal 2013). Because the three items are strongly correlated (Cronbach’s  $\alpha = 0.718$ ) we created an index of school misbehavior using the unweighted average of all three items. Our regression analysis uses the school misbehavior index to reduce the influence of measurement error, but we also report the results using the three measures of misbehavior separately. We matched teachers’ evaluations with the experimental data using identification codes to preserve subjects’ anonymity.

### 3 Results

The results indicate that a significant proportion of the subjects cheated by inflating their number of successful coin tosses. Figure 1 shows that the empirical distribution of reported heads is shifted towards a higher number of heads relative to the honest benchmark provided by the binomial distribution. The outcomes ten, nine, and eight times heads are significantly over-represented ( $p < 0.001$  for all three outcomes, binomial tests), whereas the outcomes two, three, four, and five times heads are significantly underreported ( $p = 0.011$ ,  $p < 0.001$ ,  $p = 0.032$ , and  $p = 0.055$ , binomial tests). On average, the students reported heads for 62.8% of the coin flips (95% confidence interval: 60.0%, 65.7%). Assuming that none of the participants cheated to his or her disadvantage

we estimate that 25.7% of the coin flips were misreported.<sup>10</sup>

Figure 1: Students' behavior in the coin tossing task



The figure indicates that a significant proportion of students cheated in the coin tossing task. The empirical distribution of reported heads (green) is shifted towards higher numbers of heads—i.e., higher earnings—relative to the binomial distribution implied by fully honest reporting (blue).

We also analyzed individual determinants of cheating using multivariate regression analysis. Column (1) of Table 1 indicates that female students behaved more honestly than male students as they reported heads significantly less often ( $p < 0.000$ , t-test).<sup>11</sup> Moreover, we found that high school students cheated significantly less than those from middle school after controlling for age ( $p = 0.011$ , t-test), which could be explained by less deviant students selecting into higher education. However, cheating is not related to measures of cognitive ability, neither to crystallized nor to fluid intelligence ( $p = 0.599$  and  $p = 0.744$ , t-tests).

<sup>10</sup>The calculation of percentage of misreported coin tosses is straightforward if we assume that none of the participants cheated to his or her disadvantage (see Houser et al. 2012). Let  $h$  be the percentage of heads reported and  $m$  be the percentage of misreported coin tosses. For any given coin toss, a participant who cheats reports heads with a probability of 1. By contrast, a participant who is truthful reports heads with a probability of 0.5. Thus, the percentage of heads reported is  $h = m*1 + (1-m)*0.5 = 0.5*(1+m)$ . Solving the equation yields the percentage of misreported coin tosses  $m = 2 * h - 1$ .

<sup>11</sup>Dreber and Johannesson (2008) document a similar gender difference in dishonest behavior.

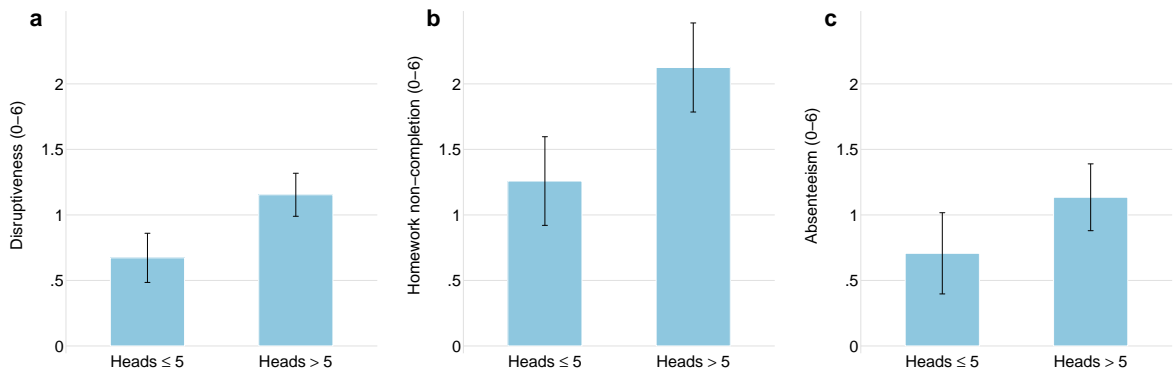
Table 1: Determinants of Cheating and School Misbehavior

Dependent variable	(1) # of heads	(2) School Misbehavior Index	(3) School Misbehavior Index	(4) School Misbehavior Index
# of heads		0.150** (0.023)	0.145** (0.022)	
Age	-0.038 (0.731)	0.489*** (0.006)	0.472*** (0.006)	0.467*** (0.008)
Female	-1.061*** (0.000)	-0.621** (0.013)	-0.663*** (0.004)	-0.817*** (0.004)
Swiss	-0.411 (0.140)	0.143 (0.224)	0.186 (0.192)	0.127 (0.422)
High school	-1.018** (0.050)	-1.360*** (0.005)	-1.033** (0.031)	-1.181** (0.020)
Parental education	-0.120 (0.657)	0.462 (0.168)	0.532 (0.138)	0.514 (0.136)
Crystallized intelligence	-0.080 (0.599)		-0.267** (0.041)	-0.279** (0.037)
Fluid intelligence	-0.041 (0.744)		0.040 (0.788)	0.034 (0.827)
Constant	8.145*** (0.000)	-6.321*** (0.005)	-6.239*** (0.004)	-5.056** (0.014)
Observations	161	161	161	161
$R^2$	0.226	0.310	0.333	0.303

This table reports OLS coefficient estimates. P-values are reported in parenthesis. In column (1), we regress the number of heads reported on a set of individual characteristics and two measures of cognitive ability. Age is measured in years. Female, Swiss, high school, and parental education are dummy variables. Parental education equals to one if at least one parent has a university degree. Our measures of crystallized and fluid intelligence are based on the scores from the word fluency test and the symbol-digit correspondence test, respectively. Both cognitive ability measures are normalized to have a mean of zero and a standard deviation of one. In columns 2 to 4, the dependent variable is the school misbehavior index, which is constructed by averaging the three items of school misconduct, including disruptiveness in class, failure to complete homework, and absenteeism (all measured on a scale from “never misbehaves” (= 0) to “always misbehaves” (= 6)). Because the models in columns 2 to 4 use teacher evaluations, we computed p-values that are robust to clustering at the class level. To account for the low number of clusters we use the wild cluster bootstrap procedure (Cameron et al. 2008) using the 6-point distribution of weights proposed by Webb (2013). The number of observations is 161 instead of 162 because one subject did not provide his age. Significance levels: \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

We next examined whether our experimental measure of cheating is related to misbehavior in school. Panels (a) to (c) in Figure 2 show that behavior in the coin tossing task is positively associated with all three measures of school misbehavior. Subjects who reported more than five times heads score higher by 0.5 points (or 72 percent) on disruptiveness in class, 0.9 points (or 69 percent) on non-completion of homework, and 0.4 points (or 61 percent) on absenteeism relative to the others. In the raw data, the correlations are statistically significant for disruptiveness and homework ( $p = 0.003$  and  $p = 0.020$ ), but absenteeism fails to reach statistical significance ( $p = 0.136$ , Spearman tests).

Figure 2: Cheating and School Misbehavior



The figure shows that, relative to those who reported five times heads or less, students who reported more than five times heads disrupt the class to a larger degree (a), fail to do their homework more often (b), and are more frequently absent from school (c). Error bars indicate the standard error of the mean (adjusted for clustering at the class level).

We additionally estimated regression models to control for factors that might jointly influence cheating and school misbehavior. In the regression analysis, we use the school misbehavior index, which is the average score of all three individual measures of school misbehavior. Our main results are similar if we analyze each measure of school misbehavior separately.<sup>12</sup>

Column (2) of Table 1 confirms that behavior in the coin tossing task is significantly

<sup>12</sup>The results are available from the authors upon request.

related to school misbehavior when controlling for age, gender, nationality, education level, and parental education. A higher number of heads reported is associated with increased behavioral problems in school ( $p = 0.022$ , t-test). Interestingly, in addition to being more honest, female and high school students also misbehave less frequently ( $p = 0.030$  and  $p < 0.001$ , respectively, t-tests). The model reported in column (3) additionally includes our measures of cognitive ability as control variables. Crystallized intelligence is negatively associated with school misbehavior ( $p = 0.033$ , t-test), but fluid intelligence is not ( $p = 0.780$ , t-test). While differences in cognitive ability explain some variation in disruptive and noncompliant behavior, the predictive power of the coin tossing task for school misbehavior remains high after controlling for key background characteristics as well as cognitive ability ( $p = 0.024$ , t-test). For comparison, the difference in school misbehavior between a student who reported 10 times heads and a student who reported 5 times heads is about the same as between two students with a (crystallized) intelligence gap of 2.7 standard deviations. The difference in school misbehavior between presumable cheaters and honest students is also larger than the widely discussed gender gap in misbehavior (e.g., Bertrand and Pan 2013). In column (4) of Table 1 we removed our laboratory measure of cheating from the regression model and found that the gender coefficient increases from -0.663 to -0.817. This suggests that gender differences in experimentally elicited rule violating behavior explains almost one-fifth of the gender gap in school misbehavior.<sup>13</sup>

## 4 Conclusion

In this paper, we examined whether a common laboratory measure of cheating is a reliable indicator of rule violating behavior in the field. We present evidence on the link between rule violating behavior in the lab and field using middle and high school

---

<sup>13</sup>We found very similar results using the pooled Blinder-Oaxaca decomposition method—a technique that was initially developed for studying the gender gap in labor market earnings (Blinder 1973; Oaxaca 1973).

students. We combined experimental data from an incentivized coin tossing task with measures of disruptive and noncompliant behavior at school. Our main result is that students who cheated more in the coin tossing task also misbehave more often at school. The relationship holds when controlling for students' socioeconomic background and cognitive ability.

From a methodological view, our findings contribute to the active debate about the generalizability of laboratory experiments, i.e., whether data obtained in the lab can be extrapolated to naturally occurring environments (Levitt and List 2007; Falk and Heckman 2009). We found a strong relationship between lab and field measures of rule violating behavior despite obvious differences across the two settings, including the context of the choice situation and the degree of scrutiny—factors which have been argued to make inferences from lab to field environments difficult. Our findings are particularly reassuring for the usefulness of laboratory methods given that behavior in the cheating paradigm has been conjectured to be more context-sensitive than other experimental measures of behavior, such as cooperativeness and consumption choices (Abeler et al. 2014).

Given the profound and long-term consequences of misbehavior in school (e.g., Segal 2013), the fact that a simple experimental measure of cheating is able to predict a variety of school misconduct has also clear implications for policymakers. In particular, our findings provide a proof of concept that policymakers can successfully identify at-risk school children early on in development, precisely at the point in childhood when targeted interventions are most likely to be effective (e.g., Heckman 2006). Thus, a deeper investigation of the psychological and neurobiological mechanisms tied to rule violating behavior will likely provide fruitful insights into the development of targeted interventions aimed to reduce delinquent classroom behavior.

## References

- Abeler, Johannes, Anke Becker, and Armin Falk (2014): Representative Evidence on Lying Costs, *Journal of Public Economics*, Vol. 113, pp. 96–104.
- Ariely, Dan (2012): The (honest) truth about dishonesty: how we lie to everyone—especially ourselves, New York: Harpercollins.
- Banerjee, Abhijit, Rema N Hanna, and Sendhil Mullainathan (2013): Corruption, in: Handbook of Organizational Economics, Princeton University Press.
- Becker, Gary S, William HJ Hubbard, Kevin M Murphy, et al. (2010): The market for college graduates and the worldwide boom in higher education of women, *American Economic Review*, Vol. 100, No. 2, pp. 229–233.
- Benz, Matthias and Stephan Meier (2008): Do people behave in experiments as in the field? Evidence from donations, *Experimental Economics*, Vol. 11, No. 3, pp. 268–281.
- Bertrand, Marianne and Jessica Pan (2013): The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior, *American Economic Journal. Applied Economics*, Vol. 5, No. 1, p. 32.
- Blinder, Alan S (1973): Wage discrimination: reduced form and structural estimates, *Journal of Human resources*, pp. 436–455.
- Bowles, Samuel, Herbert Gintis, and Melissa Osborne (2001): The determinants of earnings: A behavioral approach, *Journal of economic literature*, pp. 1137–1176.
- Buccioli, Alessandro and Marco Piovesan (2011): Luck or Cheating? A Field Experiment on Honesty with Children, *Journal of Economic Psychology*, Vol. 32, No. 1, pp. 73–78.
- Burks, Stephen V, Daniele Nosenzo, Jon Anderson, Matthew Bombyk, Derek Ganzhorn, Lorenz Goette, Aldo Rustichini, et al. (2015): Lab Measures of Other-Regarding Preferences Can Predict Some Related on-the-Job Behavior: Evidence from a Large Scale Field Experiment, University of Nottingham Working Paper CeDEx 2015-21.
- Camerer, Colin F (2015): The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Reply to Levitt and List, in: G Fréchet and A Schotter (Eds.), The Methods of Modern Experimental Economics, Oxford University Press Oxford, UK.
- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2008): Bootstrap-based improvements for inference with clustered errors, *The Review of Economics and Statistics*, Vol. 90, No. 3, pp. 414–427.
- Campbell, Frances, Gabriella Conti, James J Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan (2014): Early childhood investments substantially boost adult health, *Science*, Vol. 343, No. 6178, pp. 1478–1485.

- Carpenter, Jeffrey and Caitlin Knowles Myers (2010): Why volunteer? Evidence on the role of altruism, image, and incentives, *Journal of Public Economics*, Vol. 94, No. 11, pp. 911–920.
- Carroll, John B (1993): Human cognitive abilities: A survey of factor-analytic studies, Cambridge University Press.
- Cohn, Alain, Ernst Fehr, and Lorenz Goette (2015a): Fair Wages and Effort Provision: Combining Evidence from a Choice Experiment and a Field Experiment, *Management Science*, Vol. 61, No. 8, pp. 1777–1794, doi: 10.1287/mnsc.2014.1970, URL <http://dx.doi.org/10.1287/mnsc.2014.1970>.
- Cohn, Alain, Ernst Fehr, and Michel A. Maréchal (2014): Business Culture and Dishonesty in the Banking Industry, *Nature*, Vol. 516, pp. 86–89.
- Cohn, Alain, Michel A. Maréchal, and Thomas Noll (2015b): Bad Boys: How Criminal Identity Salience Affects Rule Violation, *Review of Economic Studies*, Vol. 82, No. 4, pp. 1289–1308.
- Dai, Zhixin, Fabio Galeotti, and Marie Claire Villeval (2015): Can Cheating in the Lab Predict Fraud in Real Settings? An Experiment in Public Transportations, in: 6th conference of the French Experimental Economics Association (ASFEE).
- Deming, David J (2015): The growing importance of social skills in the labor market, NBER Working Paper No. 21473.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2010): Are Risk Aversion and Impatience Related to Cognitive Ability?, *American Economic Review*, Vol. 100, No. 3, pp. 1238–60.
- Dreber, Anna and Magnus Johannesson (2008): Gender differences in deception, *Economics Letters*, Vol. 99, No. 1, pp. 197–199.
- Duggan, Mark, Steven D Levitt, et al. (2002): Winning Isn’t Everything: Corruption in Sumo Wrestling, *American Economic Review*, Vol. 92, No. 5, pp. 1594–1605.
- Falk, Armin and James J Heckman (2009): Lab experiments are a major source of knowledge in the social sciences, *Science*, Vol. 326, No. 5952, pp. 535–538.
- Fehr, Ernst and Andreas Leibbrandt (2011): A field study on cooperativeness and impatience in the tragedy of the commons, *Journal of Public Economics*, Vol. 95, No. 9, pp. 1144–1155.
- Fischbacher, Urs and Franziska Föllmi-Heusi (2013): Lies in Disguise - An Experimental Study on Cheating, *Journal of the European Economic Association*, Vol. 11, No. 3, pp. 525–547.
- Fortin, Nicole M, Philip Oreopoulos, and Shelley Phipps (2015): Leaving Boys Behind: Gender Disparities in High Academic Achievement, *Journal of Human Resources*, Vol. 50, No. 3, pp. 549–579.



- Gneezy, Uri (2005): Deception: The role of consequences, *American Economic Review*, Vol. 95, No. 1, pp. 384–394.
- Goldin, Claudia, Lawrence F Katz, and Ilyana Kuziemko (2006): The Homecoming of American College Women: The Reversal of the College Gender Gap, *The Journal of Economic Perspectives*, Vol. 20, No. 4, pp. 133–156.
- Hanna, Rema and Shing-Yi Wang (2014): Dishonesty and Selection into Public Service, NBER Working Paper No. 19649.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev (2013): Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes, *The American Economic Review*, Vol. 103, No. 6, pp. 2052–2086.
- Heckman, James J (2006): Skill formation and the economics of investing in disadvantaged children, *Science*, Vol. 312, No. 5782, pp. 1900–1902.
- Heckman, James J, Jora Stixrud, and Sergio Urzua (2006): The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics*, Vol. 24, No. 3, pp. 411–482.
- Houser, Daniel, Stefan Vetter, and Joachim Winter (2012): Fairness and Cheating, *European Economic Review*, Vol. 56, No. 8, pp. 1645–1655.
- Irlenbusch, Bernd and Marie Claire Villeval (2015): Behavioral ethics: how psychology influenced economics and how economics might inform psychology?, *Current Opinion in Psychology*, Vol. 6, pp. 87–92.
- Jacob, Brian A and Steven D Levitt (2003): Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating, *The Quarterly Journal of Economics*, Vol. 118, No. 3, pp. 843–877.
- Karlan, Dean S. (2005): Using Experimental Economics to Measure Social Capital and Predict Financial Decisions, *American Economic Review*, Vol. 95, No. 5, pp. 1688–1699.
- Levitt, Steven D and Ming-Jen Lin (2015): Catching Cheating Students, NBER Working Paper No. 21628.
- Levitt, Steven D and John A List (2007): What do laboratory experiments measuring social preferences reveal about the real world?, *The journal of economic perspectives*, pp. 153–174.
- List, John A (2009): The economics of open air markets, NBER Working Paper No. 15420.
- Meier, Stephan and Charles Sprenger (2010): Present-biased preferences and credit card borrowing, *American Economic Journal: Applied Economics*, pp. 193–210.

- Oaxaca, Ronald (1973): Male-female wage differentials in urban labor markets, *International economic review*, pp. 693–709.
- Pande, R and B Olken (2012): Corruption in Developing Countries, *Annual Review of Economics*, Vol. 4, No. 1, pp. 479–509.
- Pierce, Lamar and Parasuram Balasubramanian (2015): Behavioral field evidence on psychological and social factors in dishonesty and misconduct, *Current Opinion in Psychology*, Vol. 6, pp. 70–76.
- Segal, Carmit (2013): Misbehavior, education, and labor market outcomes, *Journal of the European Economic Association*, Vol. 11, No. 4, pp. 743–779.
- Shalvi, Shaul, Jason Dana, Michel J.J. Handgraaf, and Carsten DeDreu (2011): Justified Ethicality: Observing Desired Counterfactuals Modifies Ethical Perceptions and Behavior, *Organizational Behavior and Human Decision Processes*, Vol. 115, No. 2, pp. 181–190.
- Shalvi, Shaul, Francesca Gino, Rachel Barkan, and Shahar Ayal (2015): Self-Serving Justifications Doing Wrong and Feeling Moral, *Current Directions in Psychological Science*, Vol. 24, No. 2, pp. 125–130.
- Sutter, Matthias (2009): Deception through telling the truth?! experimental evidence from individuals and teams\*, *The Economic Journal*, Vol. 119, No. 534, pp. 47–60.
- Sutter, Matthias, Martin Georg Kocher, Stefan T Trautmann, et al. (2013): Impatience and Uncertainty: Experimental Decisions Predict Adolescents’ Field Behavior, *American Economic Review*, Vol. 103, No. 1, pp. 510–31.
- Webb, D., Matthew (2013): Reworking Wild Bootstrap Based Inference for Clustered Errors, Queen’s Economics Department Working Paper No. 1315.
- Zitzewitz, Eric (2012): Forensic economics, *Journal of Economic Literature*, pp. 731–769.