University of Warsaw
Faculty of Economic Sciences

Working Papers

No. 9/2012 (75)

Michał Krawczyk

Sex, morals and exam cheating

Warsaw 2012

University of Warsaw
Faculty of Economic Sciences

Sex, morals and exam cheating

Michał Krawczyk
University of Warsaw
Faculty of Economic Sciences
e-mail: mkrawczyk@wne.uw.edu.pl

Abstract

This paper reports a field experiment investigating effectiveness of moral appeal in discouraging exam cheating. Substantial level of cheating was identified using an index of test answers similarity, contrasted with low self-reports. The treatment manipulation made an impact on self-reported but not observed frequency of cheating. Hypothesized gender difference, whereby males took but not gave more illicit information than females was also found.

Keywords:

exam cheating, moral appeal, gender differences, field experiments

JEL:

C93, D82, I23

Acknowledgements:

I would like to thank Aleksandra Gawrysiak-Zabłocka, Gabriela Grotkowska, Andrzej Kondratowicz, Katarzyna Jakubowska-Krawczyk and Katarzyna Kopczewska for allowing me to conduct experiments during their exams and Maciej Wilamowski who provided excellent research assistance. The study was funded by the European Commission under the European Reintegration Grant ``Academic dishonesty -- a non-survey approach"

1 Introduction and literature review

Exam cheating is an important violation of academic integrity, seriously undermining the quality of teaching, reliability of evaluation processes and public trust in education. Regrettably, such violations of proper academic conduct are not infrequent. Actually, there is evidence that they might be on the rise (Davis, Grover, Becker, and McGregor, 1992; McCabe and Bowers, 1994; McCabe, Treviño, and Butterfield, 2001; Enker, 1987). To date, most studies on cheating are based on surveys (of US students or faculty). Starting from Bowers (1964) scholars found cheating to be endemic at American colleges. While studies comparing prevalence of cheating internationally are much more scarce, some evidence exists e.g. that cheating may be even more common in Central and Eastern Europe (Lupton, Chapman, and Weiss, 2000; Lupton and Chaqman, 2002; Grimes, 2004; Magnus, Polterovich, Danilov, and Savvateev, 2002).

A lot is known about factors that correlate with dishonest exam behavior (Whitley, 1998). From the policy viewpoint it may be most important to investigate which *institutional* arrangements lower the prevalence of cheating. Perhaps not surprisingly, probability of detection and severity of punishment have been found to deter such behaviors (Haswell, Jubb, and Wearing, 1999). However, such "physical" forms of enforcement would seem a short-term solution; one would rather prefer measures aimed at internalization of norms of proper academic conduct. As Gallant and Drinan (2006) put it, "the student cheating problem [is] an adaptive challenge (one that requires learning and changes in attitudes, behaviors, or values) rather than a technical problem (one that can be solved in routine ways)".

One such attempt that seems to be effective is the college "honor codes" under which students pledge to obey rules of conduct and take more responsibility for proper behavior of themselves and colleagues, e.g. exams are not invigilated (McCabe and Trevino, 1993; McCabe, Treviño, and Butterfield, 2001). McCabe, Treviño, and Butterfield (1996) find that honor codes may matter even after leaving the college–alumni of code schools were less likely to display dishonest behavior in the workplace. More generally, a number of other recent papers reporting a link between ethical choices at school and in the workplace, point out yet another reason why we should care about college cheating.

Even at no-code schools, numerous studies have succesfully altered ethical attitudes, often measured by Defining Issues Test, by means of (business) ethics courses. In particular, changes in attitudes related to cheating have been reported e.g. in (Glenn, 1992) and summarized in (Bampton and Cowton, 2009), although results tend to be mixed. While substantial gap may

generally exit between attitudes and behavior (Ajzen and Fishbein, 2005), moral beliefs and Simkin and McLeod (2010) point at moral beliefs as the only statistically significant deterrent of cheating.

One would wonder how much time and effort is needed to alter the moral attitude towards cheating. Would "moral appeal" be immediately effective? Actually, the findings of lower level of reported cheating in honor code schools provide relatively weak evidence in favor of general effectiveness of such "soft" measures. For one thing, it may well be that lower prevalence of cheating in schools where moral code is implemented result from selection of (more ethical) students. Even worse, it cannot be excluded that this low level of cheating found in survey studies is an artifact–students in honor-code schools may be more reluctant to admit to cheating. More generally, several authors (Spiller and Crown, 1995) question reliability of self-reports on cheating behavior and find biases, i.a. related to social desirability (Bernardi and Adamaitis, 2006). Schmelkin, Gilbert, Spencer, Pincus, and Silva (2008) also point out that surveys tend to impose researchers' frame upon the participants and agree with Allen, Fuller, and Luckett (1998) that indirect (non-survey measures) may turn out to be more accurate. These concerns will be just as relevant for survey-based ethics course interventions studies.

This calls for a non-survey approach. Unfortunately, up to date field data studies and especially experiments have been scarce in this field. Nowell and Laufer (1997) returned quizzes to students, asking them to grade their own work. Subjects were unaware of the fact that researchers have graded the quizzes themselves in the meantime. The authors found that 23% of students cheated at least once but only some 9% later admitted it, despite use of the Randomized Response Technique. In a natural experiment by West, Ravenscroft, and Shrader (2004) as many as 74% of students used unauthorized material or sought others' help on a take-home assignment.

Ponemon (1993) reports that 10-week ethics course intervention had no impact on ethical judmgment score (Defining Issues Test) or actual ethical choices. Note that the latter were not related to classroom cheating; instead, students were asked to contribute to the costs of course materials.[1] Bloodgood, Turnley, and Mudrack (2008) looked at subjects' tendncy to cheat on self-evaluated word-search tasks taken in a classroom setting. They found no main effect of having completed an ethics course, although the effect seemed to be present when only behavior of highly intelligent or highly religious

[1]The subjects were misled to believe that their contributions were anonymous and that they would be necessary to continue providing handouts for students. It is not clear to what extent low reliability of these bogus assertions may have led to the observed lack of treatment effect.

students was analyzed.[2]

I am not aware of studies aimed at assessing the impact of honor codes or other "moral appeal" approaches on classroom cheating. This study is an example of such a field experiment. Additionally, in order to assess reliability of survey methods, I also distribute questionnaires to my participants and I am able to link them to actual behavior.

Another issue that is worth investigating using non-survey data is the gender difference in cheating behavior (Tibbetts, 1999). To the extent that girls are socialized to hold higher moral standards (Franke, Crown, and Spake, 1997) we expect them to be less willing to cheat in exams. Indeed, a meta study by Whitley, Nelson, and Jones (1999) finds that cheating is slightly more frequent in males. The authors note that this result corresponds to findings of gender difference in other risky and (mildly) socially inappropriate activities, such as petty crimes, gambling or excessive alcohol consumption. However, Whitley and colleagues also find that the difference in *attitude* toward cheating is substantially larger, again, males being more lenient. Given that vast majority of studies analyzed by Whitley et al. are surveys and that *admitting* to cheating is more likely in individuals who find it more acceptable, one may wonder whether there is any difference in actual behavior or we merely observe an artifact of the method.[3]

Substantial literature has also emerged on the gender difference in types of moral reasoning. As proposed by Gilligan and others, females tend to perceive morals in terms of *caring* and males in terms of *justice*. This gives rise to the interesting prediction that different types of violations may be observed in the two genders, women, compared to men, being relatively more eager to let another student copy their answers rather than copy somebody else's answers themselves. However, previous studies (Newstead, Franklyn-Stokes, and Armstead, 1996) failed to find such an effect.

2 Design and procedures

The experiment took place in January of 2010 during an exam at the Department of Law of one of Polish universities. The exam involved forty closed-end (A-D) multiple choice test questions. At the beginning of the exam, students were divided into two conditions: those in the left-hand side of the exam

[2]Interestingly, a recent paper by (Christensen, Rees, and Barnes, 2011) finds exactly the opposite for moral judgments, i.e. only religious students' are positively affected by an ethics course.

[3]Another related difficulty is that response rates tend be different (higher in females), as noted by McCabe and Trevino (1997)[p. 386]

hall were assigned to the "Moral Appeal" (MA) condition and the others to the control group (CG).[4] In bot conditions students would be informed on the opening page of the exam that they were not allowed to communicate or make use of books and notes. Those in the MA condition would additionally be told that college cheating was a form of theft, that it could hurt other students directly (e.g. when grade thresholds are higher because of cheaters' higher scores) and in the long run (by reducing the value of a university diploma). They were solemnly told not to cheat or let others copy their answers. Efforts were made that nobody finds out that two different treatments were used or that any experiment was actually taking place for that matter.

All the answer sheets were secretly coded such that each student's exact location could be determined after the exam. Additionally, after the examination, students were asked to respond to a short questionnaire regarding attitudes and beliefs concerning cheating, involvement in cheating at the university, including the very examination in question, as well as background characteristics such as age, gender and Grade Point Average (GPA). These data could also be linked with a specific location in the hall and thus with the responses to exam questions.

This somewhat tedious procedure had been perfected in two small pilot "sessions" and two full-scale attempts that are not reported here because they involved a very low level of cheating, preventing identification of treatment or gender effects.

3 Results

3.1 Cheating

A total of 108 students seated in a 12 x 9 grid participated in the exam and hence the study. Two versions of the test were used, with 55 and 53 students respectively. Mean age was 23.5 years and 69% were female.

In order to assess prevalence of cheating I used the K index developed by Holland (1996). It determines for each pair of participants the probability that any incorrect answer given by the suspected "source" (S) is also given by the suspected "copier" (C), based on matches of incorrect answers between S and other participants, who have (about) as many incorrect answers as C. This probability is then used in a binomial formula to compute the p value for the null hypothesis of independent answers. I have calculated these p

[4]Students would be individually directed to specific locations selected at random

values for 5726 ordered pairs of participants who had the same version of the test.

The findings may be summarized as follows. First, cheating was fairly common. Using an arbitrary significance threshold of .001 we should expect about 6 cases, for which p-value is lower than that in our nearly 6 thousand pairs. In fact, there 63 such pairs, more than ten times the expected number. If we still assumed that these low p-values arose randomly, we should expect, given the spatial structure of our sample, that about 20% of those would consist of individuals seated not more than two rows and two seats apart. In fact, two-thirds of them (42 out of 63) did (similar results obtain if I used another distance as threshold).

Sixty-two out of 108 participants (57.4%) were involved in at least one of these suspiciously similar pairs, either as a source or copier or both. One can also look at other, simpler statistics: eighteen pairs of subjects had more than 30 identical responses. Of these, 17 were pairs of individuals who sat at most two rows and two seats away from each other, see Table 1.

Table 1: Locations of pairs with more than 30 matched answers

	row c	row s	column c	column s	treat. c	treat. s	matches
1	11	11	5	6	MA	MA	31
2	7	9	3	10	MA	CG	31
3	4	5	7	8	CG	CG	33
4	9	10	10	10	CG	CG	35
5	7	6	2	1	MA	MA	39
6	7	6	4	5	MA	MA	33
7	9	9	4	6	MA	MA	38
8	8	9	4	3	MA	MA	40
9	6	5	10	9	CG	CG	36
10	10	10	1	2	MA	MA	32
11	10	9	7	8	CG	CG	32
12	10	10	7	8	CG	CG	33
13	10	10	7	6	CG	MA	37
14	9	10	8	8	CG	CG	39
15	9	10	8	6	CG	MA	31
16	10	10	8	6	CG	MA	32
17	8	9	5	5	MA	MA	39
18	7	6	8	7	CG	CG	32

Second, there was no impact of the treatment variable on the prevalence

of cheating, see Table 2. Similarly, the Kolmogorov-Smirnov statistic showed no significant difference in the distribution of similarity indices between treatments.

3.2 Self-reported cheating

Table 2 also shows students' response to questions concerning cheating behavior.

Table 2: Treatment comparison

treatment	CG	MA
similarity index: $p < .001$		
copier	32.7%	35.7%
cource	36.5%	35.7%
Self-reports		
cheated today	13.4%	5.7%
ever cheated in college	68.0%	63.6%
guess % cheaters today	27.9%	28.9%
guess % cheaters typical exam	55.4%	42.1%
probability of getting caught (1-7)	3.45	3.27
pr. of punishment when caught (1-7)	2.59	2.33
cheating unethical (1-7)	4.84	4.70

Overall, the self-reported level of cheating was quite low. Interestingly, a marginally significant difference was observed, whereby participants in the MA condition were marginally less likely to admit to cheating, $p = .087$ in the test of proportions (one-sided), although, as mentioned before, there was no significant difference in actual behavior. Expectedly, there was no difference in estimated fraction of cheaters, either in the exam in question or a "typical exam" at the university (the latter, however, correlated highly ($r = .32, p = .001$) with self-reported cheating in the past, a possible indication of false consensus). It is also noteworthy that students generally did not consider it very likely that an attempt to cheat in exam is detected and, especially, punished. However, they generally perceived it as (somewhat) unethical.

Of special importance is the low prevalence of self-reported cheating among the 62 likely cheaters (as measured by the similarity indices). It turns out only six of them admitted, the fraction was not significantly higher than among "non-cheaters".

3.3 Determinants of cheating. Gender differences.

Based on women's higher moral standards but also greater tendency to display a caring attitude, I predicted that males would copy but not share more often than females. This is precisely what we observe. Of 33 males in the sample 17 (51.5%) were identified as copiers. Among 75 females only 20 (26.7%) seemed to have copied from peers ($p = .01$ in a two-sided test of proportions). The difference was much smaller for the variable indicating being an alleged source: it was positive in 45.5% male students and 32% females, ($p = .18$). This gender-specific pattern is particularly striking given that sources were not necessarily sharing voluntarily (and could even be unaware of somebody else peeking at their exam copy), that the similarity indices yield little power to distinguish between the copier and the source in the first place and that these two variables are highly correlated ($r = .391$).

In order to identify other determinants of cheating behavior I have run logistic regressions with the indicator–copier or source–as dependent variable and treatment dummy as well as responses to questionnaire items as explanatory variables. Exam score as well as interaction between gender and treatment were also considered. In a step-wise procedure I would remove variables that were not significant at 20% level. The resulting model for the copier is given in Table 3 and for the source in Table 4.

Table 3: Determinants of copying: logistic regression

	Coef.	Std. Err.	p value
male	1.51	.67	0.023
score	-.15	.07	0.049
work	-1.27	.62	0.040
GPA	-1.21	.75	0.106
cheated today	1.35	1.00	0.176
ever cheated in college	-2.37	.68	0.001
guess % cheaters typical exam	.026	.012	0.029
cons	7.15	3.12	0.022

$n = 87$, Prob $> \chi^2 = 0.0003$, Pseudo $R^2 = 25.5\%$. Note: lower number of obs. due to missing GPA scores. Results are very similar when this variable is skipped.

Regression results confirm findings reported previously: Treatment has no impact and males copy but not share more than females. On top of that we find there is no interaction between treatment and gender. Exam score has a significant and predictable impact–better students copy less but are copied from more. This result seems to confirm that the roles of copiers and

Table 4: Determinants of sharing: logistic regression

	Coef.	Std. Err.	p value
score	-.25	.07	0.000
ever cheated in college	-2.16	.61	0.000
guess % cheaters typical exam	.030	.010	0.004
cheating unethical (1-7)	-.29	.14	0.042
cons	-4.55	1.51	0.003

$n = 101$, Prob $> \chi^2 = 0.0000$, Pseudo $R^2 = 0.2809$

sources are correctly identified. Admitting to cheating during the exam is not significantly related to being a copier or a source. Admitting to cheating ever has an impact but it goes in the unexpected direction: we would expect positive autocorrelation in cheating behavior but we observe that compared to non-cheaters, a smaller fraction of cheaters reported having cheated in the past! One possible interpretation is that nearly all students had cheated in one of previous exams, yet those who have also cheated just minutes before are instinctively afraid to provide what could be considered as incriminating evidence (although they are assured that results would only be used for research purposes). Indeed, the estimated percentage of peers that cheat in a "typical" exam (which may be an unbiased, however indirect, indicator of own behavior (Fisher, 1993)) has the expected positive impact on behavior. Also individuals who consider cheating unethical tend to copy and share less than others. I was not able to find differences described in (Tibbetts, 1999) such as greater impact of grades in females. This null result may be due to limited size of the male sample.

4 Conclusion

The lessons from the study are as follows. Firstly, it seems to suggest that one-time instance of moral appeal is ineffective in deterring cheating. Although baseline level of cheating was rather high, there is no trace of reduction as a result of experimental manipulation. Of course, it may well be that more prolonged exposure to the environment that promotes academic integrity does make a substantial difference; a proper randomized experiment involving such a manipulation would be rather difficult to implement.

Secondly, it shows that relying on self-reports will generally be insufficient as participants are often unwilling to answer truthfully. While most of my subjects endorse the more general notion of having behaved unethically in the

past, they are not willing to admit to cheating during a particular exam, even though they could trust the answers would not be used against them.[5] Self-reports were actually quite misleading in that real cheaters tended to report less cheating in the past. The fact that treatment manipulation only affected reported, not actual cheating, also suggests that un-incentivized self-reports on such a sensitive matter should be treated with great caution. Ultimately, combination of survey-based and experimental techniques, benefiting from relative merits of the two approaches, should prove most useful in the future.

Thirdly, I confirm the standard survey result of males cheating more. On top of that, the data allows positive verification of the previously hypothesized but not empirically observed pattern of gender differentiation in which females are relatively more willing to let their colleagues copy their answers as compared to copying themselves. Such results can help tailor academic integrity training programs, as noted previously by Tibbetts (1999). For example, educators can appeal to male copiers by indicating that cheating is unjust and dishonorable and to female sharers by explaining that refusing illicit help may encourage the requester to study on his own and ultimately benefit him.

[5]In related studies I used the randomized response approach (Warner, 1965) in order to encourage truthful responses. Unlike in (Kerkvliet, 1994) it did not make any difference.

References

Ajzen, I., and M. Fishbein (2005): "The influence of attitudes on behavior," in *The handbook of attitudes*, ed. by D. Albarracin, B. Johnson, and M. Zanna. Lawrence Erlbaum.

Allen, J., D. Fuller, and M. Luckett (1998): "Academic integrity: Behaviors, rates, and attitudes of business students toward cheating," *Journal of Marketing Education*, 20(1), 41–52.

Bampton, R., and C. Cowton (2009): "Taking stock of accounting ethics scholarship: a review of the journal literature," Working Paper of the Financial Ethics and Governance Research Group University of Huddersfield.

Bernardi, R., and K. Adamaitis (2006): "Data contamination by social desirability response bias: an international study of students' cheating beahvior," *Research on Professional Responsibility and Ethics in Accounting*, 11, 149.

Bloodgood, J., W. Turnley, and P. Mudrack (2008): "The influence of ethics instruction, religiosity, and intelligence on cheating behavior," *Journal of business ethics*, 82(3), 557–571.

Bowers, W. (1964): *Student dishonesty and its control in college*. New York: Bureau of Applied Social Research, Columbia University.

Christensen, D., D. Rees, and J. Barnes (2011): "The Effect Of Church And Service On Ethics Education: An Experiment With Accounting Students," *Journal of College Teaching & Learning (TLC)*, 3(5).

Davis, S., C. Grover, A. Becker, and L. McGregor (1992): "Academic dishonesty: Prevalence, determinants, techniques, and punishments," *Teaching of Psychology*, 19(1), 16–20.

Enker, M. (1987): "Attitudinal and normative variables as predictors of cheating behavior," *Journal of Cross-Cultural Psychology*, 18(3), 315.

Fisher, R. (1993): "Social desirability bias and the validity of indirect questioning," *Journal of Consumer Research*, pp. 303–315.

Franke, G., D. Crown, and D. Spake (1997): "Gender differences in ethical perceptions of business practices: A social role theory perspective.," *Journal of Applied Psychology*, 82(6), 920.

Gallant, T., and P. Drinan (2006): "Organizational theory and student cheating: Explanation, responses, and strategies," *Journal of Higher Education*, pp. 839–860.

Glenn, J. (1992): "Can a business and society course affect the ethical judgment of future managers?," *Journal of Business Ethics*, 11(3), 217–223.

Grimes, P. (2004): "Dishonesty in academics and business: A cross-cultural evaluation of student attitudes," *Journal of Business Ethics*, 49(3), 273–290.

Haswell, S., P. Jubb, and B. Wearing (1999): "Accounting students and cheating: A comparative study for Australia, South Africa and the UK," *Teaching Business Ethics*, 3(3), 211–239.

Holland, P. (1996): "Assessing unusual agreement between the incorrect answers of two examinees using the K-index: Statistical theory and empirical support," *Research Report–Educational Testing Service Princeton RR*.

Kerkvliet, J. (1994): "Cheating by economics students: A comparison of survey results," *Journal of Economic Education*, pp. 121–133.

Lupton, R., K. Chapman, and J. Weiss (2000): "International Perspective: A Cross-National Exploration of Business Students' Attitudes, Perceptions, and Tendencies Toward Academic Dishonesty," *The Journal of Education for Business*, 75(4), 231–235.

Lupton, R., and K. Chaqman (2002): "Russian and American college students' attitudes, perceptions and tendencies towards cheating," *Educational Research*, 44(1), 17–27.

Magnus, J., V. Polterovich, D. Danilov, and A. Savvateev (2002): "Tolerance of cheating: An analysis across countries," *The Journal of Economic Education*, 33(2), 125–135.

McCabe, D., and W. Bowers (1994): "Academic dishonesty among males in college: A thirty year perspective.," *Journal of College Student Development*.

McCabe, D., and L. Trevino (1993): "Academic dishonesty: Honor codes and other contextual influences," *Journal of Higher Education*, pp. 522–538.

——— (1997): "Individual and contextual influences on academic dishonesty: A multicampus investigation," *Research in Higher Education*, 38(3), 379–396.

McCabe, D., L. Treviño, and K. Butterfield (1996): "The influence of collegiate and corporate codes of conduct on ethics-related behavior in the workplace," *Business Ethics Quarterly*, pp. 461–476.

——— (2001): "Cheating in academic institutions: A decade of research," *Ethics &Behavior*, 11(3), 219–232.

Newstead, S., A. Franklyn-Stokes, and P. Armstead (1996): "Individual differences in student cheating.," *Journal of Educational Psychology*, 88(2), 229.

Nowell, C., and D. Laufer (1997): "Undergraduate student cheating in the fields of business and economics," *The Journal of Economic Education*, 28(1), 3–12.

Ponemon, L. (1993): "Can ethics be taught in accounting?," *Journal of Accounting Education*, 11(2), 185–209.

Schmelkin, L., K. Gilbert, K. Spencer, H. Pincus, and R. Silva (2008): "A multidimensional scaling of college students' perceptions of academic dishonesty," *The Journal of Higher Education*, pp. 587–607.

Simkin, M., and A. McLeod (2010): "Why Do College Students Cheat?," *Journal of business ethics*, 94(3), 441–453.

Spiller, S., and D. F. Crown (1995): "Changes over time in academic dishonesty at the collegiate level," *Psychological Reports*, 76(3), 763–768.

Tibbetts, S. (1999): "Differences between women and men regarding decisions to commit test cheating," *Research in Higher Education*, 40(3), 323–342.

Warner, S. (1965): "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, pp. 63–69.

West, T., S. Ravenscroft, and C. Shrader (2004): "Cheating and moral judgment in the college classroom: A natural experiment," *Journal of Business Ethics*, 54(2), 173–183.

Whitley, B. (1998): "Factors associated with cheating among college students: A review," *Research in Higher Education*, 39(3), 235–274.

Whitley, B., A. Nelson, and C. Jones (1999): "Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis," *Sex Roles*, 41(9), 657–680.

Faculty of Economic Sciences
University of Warsaw
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl