ETHICAL DECISION-MAKING
A Perspective from the Behavioral Sciences

Silvia Saccardo
Department of Social and Decision Sciences, CMU
Why should we care about ethics?
Why should we care about ethics?

Individuals sometimes engage in unethical behavior.
Why should we care about ethics?

Individuals sometimes engage in unethical behavior
Why should we care about ethics?

Individuals sometimes engage in unethical behavior

Madoff’s Ponzi Scheme
Why should we care about ethics?

Individuals sometimes engage in unethical behavior

Madoff’s Ponzi Scheme

Volkswagen emission scandal
Why should we care about ethics?

These decisions can have devastating consequences for individuals and the society as a whole.

- Madoff’s Ponzi Scheme
- Volkswagen emission scandal
Understanding the drivers of unethical behavior can help to prevent these and other types of ethical failure.
Ethics - Moral Philosophy approach

- Prescriptive or normative approach that used insights from philosophy to describe how moral people should behave

- Evaluation of actions from a moral point of view.

“To me, however, it seems certain that every lie is a sin…”
Ethics - Moral Philosophy approach

**Utilitarian view** - examines moral actions from their outcomes: “the greatest good for the greatest number of people”

**Deontological view** - morality of an action depends on the action adherence to rules or principles
Ethics - Moral Philosophy approach

**Utilitarian view** - examines moral actions from their outcomes: "the greatest good for the greatest number of people"

**Deontological view** - morality of an action depends on the action adherence to rules or principles

Trolley Dilemma
Ethics - Moral Philosophy approach

**Utilitarian view** - examines moral actions from their outcomes: “the greatest good for the greatest number of people”

**Deontological view** - morality of an action depends on the action adherence to rules or principles

Trolley Dilemma

Footbridge Dilemma
Ethics - Homo Economics

• Individuals are selfish, profit-maximizing agents who are unconcerned about the well-being of others

• They should engage in unethical behavior (e.g., commit a crime, lie, cheat, break trust) whenever their own benefits of doing so outweigh the costs, regardless of the effect on others
Ethics - Homo Economics

• Individuals are selfish, profit-maximizing agents who are unconcerned about the well-being of others

• They should engage in unethical behavior (e.g., commit a crime, lie, cheat, break trust) whenever their own benefits of doing so outweigh the costs, regardless of the effect on others

What do you think?
Born Selfish?

https://www.youtube.com/watch?v=S0qjK3TWZE8
Behavioral Approach to Ethics

• Using empirical approach, behavioral ethics incorporates insights from economics and psychology and to study the **determinants of ethical and unethical behavior**.

• Does not take a normative approach but rather studies **how decisions are made by people** and **why good people sometimes do bad things**

• Role of context/environment in which decisions are made, situational factors, psychological processes.
Unethical Behavior: Empirical Facts

• Empirical findings on unethical behavior show that individuals are not always acting morally
Unethical Behavior: Empirical Facts

• But they also do not behave as predicted by homo economics

• Indeed, most people cheat, but only do so a little bit (De Paulo, 1996; Mazar et al, 2008; Fischbacher and Heusi, 2013)

• Can you come up with some examples?
The Cheating Experiment  (Fishbacher and Heusi, 2013)

Imagine I give you a dice and ask you to roll it in private
The Cheating Experiment (Fishbacher and Heusi, 2013)

Imagine I give you a dice and ask you to roll it in private

• If you roll a 1, you get $1
• If you roll a 2, you get $2
• If you roll a 3, you get $3
• If you roll a 4, you get $4
• If you roll a 5, you get $5
• If you roll a 6, you get $0
The Cheating Experiment (Fishbacher and Heusi, 2013)

Imagine I give you a dice and ask you to roll it in private

• If you roll a 1, you get $1
• If you roll a 2, you get $2
• If you roll a 3, you get $3
• If you roll a 4, you get $4
• If you roll a 5, you get $5
• If you roll a 6, you get $0

Which number would you report?
A substantial fraction of participants engages in partial cheating.
Behavioral approach to ethics attempts to understand the situations in which individuals are more likely to violate some moral norms.

- Why do people only cheat partially?
- When are individuals more likely to commit moral violations?
- What are the factors that lead “good” people to do wrong?
- How can we design institutions and organizations in order to minimize ethical failures?
Examples from my own research

• The role of self-deception in (un)ethical behavior
• Corruption: the role of verbal persuasion
• Moral dilemmas and discrimination
SELF-DECEPTION AND UNETHICAL BEHAVIOR

Can well-meaning individuals behave unethically without realizing it?

with U. Gneezy, M. Serra-Garcia and R. van Veldhuizen
Some people care about their identity and want to perceive themselves as moral people (see e.g., Bénabou and Tirole 2011; Mazar et al., 2008; Ariely et al., 2009).

At the same time, these people would like to enjoy material gains that could be obtained via unethical behavior.

Dishonesty and Self-Image
To protect such belief, some people may engage in **self-deception** and convince **themselves** that advice that maximizes their material gains is also ethical.
Examples

Medical Advice
Overtreatment estimated to cost $210 billion (IOM, 2012)

Financial Advice
Incentives may lead to biased recommendations
Can we find evidence that the ability to self-deceive enables dishonest behavior?
Individuals informed they will be asked to make a subjective judgment

See the items they will judge

Make a recommendation
Individuals informed they will be asked to make a subjective judgment

See the items they will judge

Make a recommendation

Information about the incentive BEFORE
Individuals informed they will be asked to make a subjective judgment

See the items they will judge

Make a recommendation

Information about the incentive AFTER
Provide recommendation to an uniformed client who chooses between A or B

Paid $1 to recommend a Product, A or B
Paid $1 to recommend a Product, A or B

Product A
50-50 lottery between $2 and $4

Provide recommendation to an uniformed client who chooses between A or B
Paid $1 to recommend a Product, A or B

<table>
<thead>
<tr>
<th>Product A</th>
<th>Product B</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-50 lottery between $2 and $4</td>
<td>50-50 lottery between $1 and $7</td>
</tr>
</tbody>
</table>

Provide recommendation to an uniformed client who chooses between A or B
Treatments

**Control.** No incentive (fixed fee of $1 only)

**Before.** $1 incentive (of top of fixed fee) for recommending A
Information about the incentives provided *before* learning about A and B

**After.** $1 incentive (of top of fixed fee) for recommending A
Information about the incentives provided *after* learning about A and B
**Hypotheses**

<table>
<thead>
<tr>
<th></th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-Interest</strong></td>
<td>Before = After &gt; Control</td>
</tr>
<tr>
<td><strong>Self-Image costs</strong></td>
<td></td>
</tr>
<tr>
<td>• <strong>Without Self-Deception</strong></td>
<td>Before = After = Control</td>
</tr>
<tr>
<td>• <strong>With Self-Deception</strong></td>
<td>Before &gt; After = Control</td>
</tr>
</tbody>
</table>
Results

N=310
Results

![Bar chart showing the fraction of A recommendations for control, before, and after periods with N=310]
Results

(z=3.90, p=.001, test of proportion)
The persistence of self-deception
What happens when we reduce ambiguity over the best option?

<table>
<thead>
<tr>
<th>Multiple Excuses</th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Expected payoff of A, B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$2 and $4</td>
<td>$1 and $7</td>
<td>$3, $4</td>
</tr>
</tbody>
</table>
The persistence of self-deception
What happens when we reduce ambiguity over the best option?

<table>
<thead>
<tr>
<th></th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Expected payoff of A,B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Excuses</td>
<td>$2 and $4</td>
<td>$1 and $7</td>
<td>$3, $4</td>
</tr>
<tr>
<td>Little Excuse</td>
<td>$2 and $4</td>
<td>$2 and $6</td>
<td>$3, $4</td>
</tr>
</tbody>
</table>
The persistence of self-deception
What happens when we reduce ambiguity over the best option?

<table>
<thead>
<tr>
<th></th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Expected payoff of A,B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Excuses</td>
<td>$2 and $4</td>
<td>$1 and $7</td>
<td>$3, $4</td>
</tr>
<tr>
<td>Little Excuse</td>
<td>$2 and $4</td>
<td>$2 and $6</td>
<td>$3, $4</td>
</tr>
<tr>
<td>No Excuses</td>
<td>$2 and $4</td>
<td>$5 and $7</td>
<td>$3, $6</td>
</tr>
</tbody>
</table>
Little Excuse - nothing changes

Fraction of A Recommendations

N=295
No Excuses - dishonesty goes down

Fraction of A recommendations

N=320
The ability to self-deceive can dramatically increase the prevalence of dishonest behavior, even in well-meaning individuals!

Interventions aimed at limiting the scope for self deception can help prevent dishonest advice by those who care to see themselves as honest.

Huge implications in many domains. Vagueness & ambiguity may allow individuals to justify unethical actions. Think about academia!!
MORAL DILEMMAS & DISCRIMINATION

Is discrimination more likely to emerge when people can plausibly justify it?

with A. Danilov
Yet, empirical work shows that discrimination persists in many economics domains (e.g., Bertrand and Mullainathan, 2004; Doleac and Stein, 2013; Fong and Luttmer, 2011; Kaas and Manger, 2012; Milkman et al., 2012; Bartos et al., 2016).

Over the past several decades, shifts in legislations and social norms have led to unprecedented positive outcomes for minorities in Western societies.

Today’s Discrimination
Are individuals more likely to discriminate when able to plausibly attribute their actions to virtuous motives?

• Predictions

• When discrimination is hard to disguise, people who want to appear as “good” will not discriminate
• When discrimination is easy to disguise, discrimination will emerge
An Experimental Test

➢ **German sounding names**
  Bernd, Dirk, Ingo, Johannes, Andreas, Florian, Tobias

➢ **Turkish sounding names**
  Baris, Emrah, Ismail, Mustafa, Ali, Huseyin, Murat
An Experimental Test

- Decision-Makers make decisions that determine the payoffs for themselves and their counterpart

<table>
<thead>
<tr>
<th>Payoff</th>
<th>DM</th>
<th>Receiver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option 1</td>
<td>€10</td>
<td>€10</td>
</tr>
<tr>
<td>Option 2</td>
<td>€9</td>
<td>€15</td>
</tr>
</tbody>
</table>

- We study the fraction of **prosocial choices** as a function of the ethnicity of the counterpart
2 versions

• Prosocial Choice Game
  ➢ Decision-Makers simply have to make a simple choice between Option 1 and Option 2 (the prosocial choice), and are paid accordingly.

  Hard to be “mean” to the minority due to clear norms of being “nice”. We expect to detect little or no discrimination.
2 versions

• **Prosocial Choice**
  ➢ Decision-Makers simply have to make a simple *choice* between Option 1 and Option 2 (the prosocial choice), and are paid accordingly.

  Hard to be “mean” to the minority due to clear norms of being “nice”. We expect to detect little or no discrimination.

• **Prosocial Lie**
  ➢ In order to reach Payoff Option 2, individuals have to lie (i.e., violate a moral norm).

  Being prosocial entails telling a lie. Not clear what is the right thing to do. In this setting discrimination may emerge.
Prosocial Choices in the Choice/Lie Games

Panel A: Prosocial Choice

Error bars denote standard errors of the mean
Prosocial Choices in the Choice/Lie Games

Panel A: Prosocial Choice

Panel B: Prosocial Lies

Error bars denote standard errors of the mean

N=166
Implications

- Discrimination only emerges when people can plausibly justify it.
- Situations with conflicting norms can lead to bias (think about hiring!).
- Again, vagueness & ambiguity may allow individuals to justify unethical actions.
CORRUPTION: BRIBES VS. VERBAL PERSUASION

What are the factors that are more effective in persuading others to violate moral norms?

with U. Gneezy, and R. van Veldhuizen
Discretionary Power and Unethical Behavior

• Many decisions are made by individuals or groups with discretionary power.
• And some subjectivity about what is ‘right’.

• Leaves scope for influencing decisions, leading people to “bend the rules”/ make corrupt decisions.
• Judges, Doctors, Policemen, Editors, etc.

• Can be done in different ways…
Bribes or Gifts

$1.1 trillion (World Bank)

CPI, Transparency International
Verbal Persuasion

Please like my paper...

It does have nice pictures...

But I put so much effort on this

That’s true, maybe I can bend the rules this one time
What distorts decisions?

• Bribes distort judgment (Armantier and Boly, 2013; Malmadier and Schmidt, 2017; Gneezy, Saccardo, van Veldhuizen, 2017)

• Verbal persuasion may make decision-makers more likely to bend the rules/lie/be corrupt

• It may increase the effectiveness of bribes/gifts?
EXPERIMENTAL DESIGN

UCSD, N=416
EXPERIMENTAL DESIGN

• **Worker’s task**: 5 minutes to construct words from a set of letters “aabells”

• Worker gets points for each correct word.

• **Judge** counts the points and indicates whether the worker has **200 points or more**.

• If he indicates that the worker did more than 200 points the worker gets a monetary bonus, if less he/she gets nothing.
We use this experiment to study whether judges are willing to do something unethical (LIE) in order to make the worker PASS.

What persuades judges to be corrupt?

UCSD, N=416
EXPERIMENTAL DESIGN

TREATMENTS

**Control:** Worker submits the task to the Judge

**Bribe:** Worker submits the task to the Judge and add a monetary bribe to it

**Communication:** Worker has 5 minutes to chat with the judge (via an online chat) to persuade him to pass him

**Bribe+Communication:** Worker can send money to the judge and also communicate with him

UCSD, N=416
ARE BRIBES AND VERBAL PERSUASION INFLUENCING THE JUDGE?

- Communication effective \( (p=0.016, \text{ test of proportions}) \).
- Bribery ineffective \( (p=0.809, \text{ test of proportions}) \).
What are people saying to persuade the Judge?

Work in progress — Stay tuned!
Many opportunities for collaborations!

• Behavioral Science research could benefit a lot from interaction/collaboration with computer scientists

• Many opportunities to generate data via randomized controlled experiments in which people interact (e.g., via chat), use NLP tools to analyze interactions, make new predictions based on the results and run new experiments to test them

• Computer scientists interested in social science questions could benefit from the theories and findings from behavioral science.
THANK YOU

Feel free to contact me with questions!

Silvia Saccardo
ssaccard@andrew.cmu.edu