

**From the holdings of the Department of Special Collections , Princeton University Library,  
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

---

Doctoral Dissertations, 2011-2026

Psychology

---

2020

# Public Acts of Deviance: Determinants and Consequences

Gomila, Robin

Princeton, NJ : Princeton University

---

<http://arks.princeton.edu/ark:/88435/dsp01cv43p074k>

*U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>*

PUBLIC ACTS OF DEVIANCE:  
DETERMINANTS AND CONSEQUENCES

ROBIN GOMILA

A DISSERTATION  
PRESENTED TO THE FACULTY  
OF PRINCETON UNIVERSITY  
IN CANDIDACY FOR THE DEGREE  
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE  
BY THE DEPARTMENT OF  
PSYCHOLOGY  
ADVISER: ELIZABETH LEVY PALUCK

JUNE 2020

© Copyright by Robin P. Gomila, 2020.

All rights reserved.

# Abstract

Social norms are powerful drivers of human behavior, so much so that people conform to norms even when they would prefer to deviate. Researchers have theorized that in these cases, people conform to avoid the costs of deviance. For one, a large body of work suggests that individuals endure psychological costs such as guilt, shame, self-deprecation, or decreased self-esteem for deviating from norms. For two, norm violators are often perceived and labeled negatively, marginalized, stigmatized, or punished by other people. Notwithstanding the possible psychological or social costs, people do decide to violate norms. In the present dissertation, I ask: Who are the deviants, and why? Once individuals have deviated from a norm, how does this experience shape them? These questions are important to address because deviants play critical roles in their communities. Specifically, public acts of deviance can signal disagreement with current practices, and catalyze group and broader social change. In this dissertation, I present evidence from qualitative and quantitative studies in the lab and the field suggesting that the experience of violating a norm increases individuals' inclination to deviance. The results also indicate a possible psychological mechanism explaining this effect: deviating causes people to depreciate their perception of the costs of deviance. After reviewing the relevant literature in Chapter 1, I present findings from in-depth interviews and two rounds of surveys of deviants and conformists (Chapter 2) suggesting that deviance is a pattern: compared to conformists, deviants report a history of deviance and of feeling different from the typical member of their groups. Chapters 3 and 4 focus on the question of how deviance becomes a pattern. Together, the results of online experiments using behavioral games (Chapter 3) and a field experiment in the Hasidic community (Chapter 4) challenge traditional views of deviance as personality or a stable individual trait that develops early in the life span. Instead, these findings suggest that patterns of deviance can emerge from an experience of deviance at any point in one's life.

# Acknowledgements

I feel grateful for spending the past years following my dream of becoming a social psychologist. This journey started about a decade ago when I discovered social psychology as an undergraduate student in Paris. I have been lucky to spend these past years learning from brilliant, stimulating and generous scholars in an incredibly supportive environment. I am filled with gratitude for all of the people who made this happen.

I would like to thank my advisor and mentor, Betsy Levy Paluck, who gave me the chance to become a scholar. Betsy, thank you for believing in me. Throughout the years, you made sure to provide me with structure and intellectual freedom, and you allowed me to become the academic I wanted to be. Thank you for guiding me every step of the way. Thank you for challenging my ideas, making me think big, and teaching me scientific rigor. There are no words to express how grateful I am for receiving your phenomenal mentorship.

I have learned tremendously from my supporting advisors and collaborators. In particular, I would like to thank Eldar Shafir, Alex Todorov, Debbie Prentice, Tom Griffiths, Cecilia Rouse, and Alin Coman for broadening my research horizons and teaching me how to be a scientist. Working with you and receiving your advice has made me a better scholar.

I would also like to thank a number of colleagues. First and foremost, this dissertation would have not been complete without Nechumi Yaffe, who made our study of deviance in the ultra-Orthodox Jewish community possible. Nechumi, thank you for making me think deeper about social norms, religion, culture, and poverty. Working with you has been invaluable. I have also been fortunate to work with a breadth of incredible scholars: Ruth Ditlmann, Ana Gantman, Hana Shepherd, Graeme Blair, Stefan Uddenberg, Bill Thompson, and Sarah Mehta. Thank you for being stimulating and enthusiastic collaborators. I learned so much from each of you.

I am grateful for the generosity of those who have provided feedback on different aspects of my research. I would like to thank Donald Green, Eduardo Morales, Alex Coppock, Brandon Stewart, Pietro Ortoleva and Andi Zhou for encouraging me to think deeply about theory and statistical methods. Your feedback has been invaluable. I would also like to thank my teachers and colleagues from the Joint Degree Program in Social Policy (JDP) for their comments on my dissertation ideas and early drafts of Chapter 2. In particular, I would like to thank Susan Fiske, Sara McLanahan, Ian Lundberg, Daniela Urbina Julio, Elsa Voytas, Herrissa Lamothe, Yo-Yo Chen, Ryan Parsons, Hannah Korevaar, and Jose Maria Rodriguez Valadez. Your interest in social policy and inequality has shaped my research trajectory in many ways. I also received thoughtful comments on my dissertation work from the Graduate Prize Fellows at the University Center for Human Values. Thanks to Olaoluwatoni Alimi, Alisabeth Ayars, Chiara Benetollo, Benjamin Bernard, Stephanie Fan, Emily Silkaitis, Aarthi Vaidyanathan, Andrew Walker-Cornetta, and Niharika Yadav. Finally, I would like to thank Liz Harman for incredibly helpful career advice. Liz, thank you for your guidance, support and generosity.

I would also like to thank the Princeton University institutions that provided me with financial support throughout my PhD. I am grateful for being a Hamid Biglari \*87 fellow at the Woodrow Wilson School of Public and International Affairs and a Laurance S. Rockefeller Graduate Prize fellow at the University Center for Human Values. I have been fortunate to receive a grant from University Center for Human Values, which supported our research in the ultra-Orthodox Jewish community.

I am filled with gratitude for my officemates Rebecca Littmann, Sherry Wu, Roni Porat, and Chelsey Clark. Thank you for being so inspiring. Thank you for challenging me, discussing ideas, motivating me, and giving me so much feedback over the years.

To all of those who passed by the BLP lab, thank you for your energy, generosity, and feedback on my research. In particular, I would like to thank Hannah Alarian, Carolyn Barnett, Nate Cheek, Jason Chin, Jasper Cooper, Ajua Duker, Ana Gantman, Heather Kugelmass, Roh Majumdar, Leon Mait, Joel Martinez, Nathan Matias, Janet Pauketat, John-Henry Pezzuto, Salar Rad, Leslie Jennings Rowley, Hana Shepherd, Jordan Starck, Margaret Tankard, Janet Xu, and Yang-Yang Zhou.

I am tremendously grateful for learning how to teach from Eduardo Morales and Diana Tamir. Your generosity and passion for teaching is contagious. Thank you for shaping my vision of teaching in the best possible way.

I would like to conclude with a few words to my family. To Maria Alessia Rossi, thank you for becoming my wife. I am incredibly lucky because you are truly extraordinary. You have filled my days with love and laughter (and a lot of activities). Your devotion to academic pursuits has been a constant source of inspiration. Your—rather high—eating standards might have made me become a chef (?). And your growing interest in plants and vegetables has gotten me into gardening. I can't wait to develop a lifetime of new hobbies with you. Mom and dad, thank you for giving me roots and wings. You are the reason why I have never stopped following my dreams. You have spent the past decades loving me unconditionally, giving credit to my crazy ideas, and supporting my deviant decisions. You have provided me with the foundations for life. You have given me strength to believe in myself even when I was losing confidence. You have given me the ability to feel loved when I was failing. Thank you for always being there for me.

## Relevant Presentations and Publications by the Author

1. The main findings and method presented in Chapter 4 (in preparation for publication co-authored with Nechumi Yaffe and Betsy Levy Paluck) were presented publicly at the following scholarly conference:

Society for Personality and Social Psychology (2020), New Orleans, LA.

*Preconference: Advances in Cultural Psychology.*

2. The main findings and method of Chapters 2, 3, and 4 were presented publicly in the following workshop:

Kahneman-Treisman Center for Behavioral Science & Public Policy,  
Princeton, NJ. (2020)

*Campus Behavioral Science Initiative Workshop*

3. The main findings and method presented in Chapter 2 were recently published in a peer-reviewed journal:

Gomila, R., & Paluck, E. L. (2020). The social and psychological characteristics of norm Deviants: A field study in a small cohesive university campus. *Journal of Social and Political Psychology*, 8(1), 220–245.

<https://doi.org/10.5964/jspp.v8i1.1134>

4. The main findings and method presented in Chapter 3 were recently published as a preprint:

Gomila, R., & Paluck, E. L. (2020). The emergence of deviance: Experiments testing the personal effects of violating a social norm. *PsyArXiv*.

<https://doi.org/10.31234/osf.io/xk3zw>

To my parents.

# Contents

|                                                                        |           |
|------------------------------------------------------------------------|-----------|
| Abstract . . . . .                                                     | iii       |
| Acknowledgements . . . . .                                             | iv        |
| List of Tables . . . . .                                               | xiii      |
| List of Figures . . . . .                                              | xiv       |
| <b>1 Introduction</b>                                                  | <b>1</b>  |
| 1.1 Social and Psychological Determinants of Deviance . . . . .        | 3         |
| 1.1.1 Contextual factors that may drive deviance . . . . .             | 3         |
| 1.1.2 Individual factors that may drive deviance . . . . .             | 5         |
| 1.2 Consequences of Deviance on Psychological and Behavioral Outcomes  | 7         |
| 1.2.1 Violating a norm makes an individual more inclined to conformity | 8         |
| 1.2.2 Violating a norm makes an individual more inclined to deviance   | 9         |
| 1.3 Outline . . . . .                                                  | 11        |
| 1.3.1 Determinants of deviance . . . . .                               | 11        |
| 1.3.2 Consequences of deviance . . . . .                               | 12        |
| <b>2 Social and Psychological Determinants of Deviance</b>             | <b>14</b> |
| 2.1 Study Context: An Established Social Norm at Princeton University  | 15        |
| 2.1.1 Overview of the studies . . . . .                                | 18        |
| 2.2 Study 1: Qualitative and Quantitative Data Collections on Students |           |
| Who Decided to Deviate vs. Conform . . . . .                           | 19        |

|          |                                                                                                                |           |
|----------|----------------------------------------------------------------------------------------------------------------|-----------|
| 2.2.1    | Study 1a: Interviews of deviant and conformist students . . .                                                  | 19        |
| 2.2.2    | Study 1b: Students deciding to deviate or conform . . . . .                                                    | 27        |
| 2.3      | Study 2: Main Survey of Deviant and Conformist Students . . . . .                                              | 29        |
| 2.4      | General Discussion . . . . .                                                                                   | 40        |
| <b>3</b> | <b>Experiments Testing the Consequences of Deviance in Behavioral Games</b>                                    | <b>46</b> |
| 3.1      | Studying the Consequences of Deviance in Causal Terms: Challenges and Approach . . . . .                       | 46        |
| 3.2      | Method . . . . .                                                                                               | 48        |
| 3.2.1    | Overview of the design . . . . .                                                                               | 48        |
| 3.2.2    | Pre-treatment setup . . . . .                                                                                  | 50        |
| 3.2.3    | Experimental round of the Public Goods game . . . . .                                                          | 51        |
| 3.2.4    | Outcome measures . . . . .                                                                                     | 51        |
| 3.2.5    | Analytic strategy . . . . .                                                                                    | 54        |
| 3.3      | Results . . . . .                                                                                              | 55        |
| 3.3.1    | Description of the sample . . . . .                                                                            | 55        |
| 3.3.2    | Pre-treatment contributions . . . . .                                                                          | 56        |
| 3.3.3    | Noncompliance to treatment . . . . .                                                                           | 57        |
| 3.3.4    | Effect of deviance incentive on violations of subsequent norms in different settings . . . . .                 | 57        |
| 3.3.5    | Effect of treatment incentive to deviate on future violations of norms of fairness and reciprocation . . . . . | 60        |
| 3.4      | General discussion . . . . .                                                                                   | 64        |
| <b>4</b> | <b>Public Acts of Deviance in Settled Cultures: A Field Experiment in the Hasidic Community</b>                | <b>67</b> |
| 4.1      | The Satmar Hasidic Community of Brooklyn . . . . .                                                             | 69        |

|          |                                                                                                   |           |
|----------|---------------------------------------------------------------------------------------------------|-----------|
| 4.1.1    | Study context: A settled culture . . . . .                                                        | 69        |
| 4.1.2    | Ethics of studying the high-end clothing norm in the Hasidic<br>community . . . . .               | 70        |
| 4.1.3    | Hypotheses . . . . .                                                                              | 72        |
| 4.2      | Method . . . . .                                                                                  | 73        |
| 4.2.1    | Experimental design . . . . .                                                                     | 73        |
| 4.2.2    | Data collection . . . . .                                                                         | 75        |
| 4.2.3    | Analytic strategy . . . . .                                                                       | 79        |
| 4.3      | Results . . . . .                                                                                 | 81        |
| 4.3.1    | Description of the sample . . . . .                                                               | 81        |
| 4.3.2    | Measure of the psychological and social costs of the one-day<br>activity . . . . .                | 81        |
| 4.3.3    | Adherence with the community's views . . . . .                                                    | 83        |
| 4.3.4    | Inclination to deviate from social norms . . . . .                                                | 84        |
| 4.3.5    | Psychological outcomes associated with shifts in individuals'<br>inclination to deviate . . . . . | 87        |
| 4.4      | General Discussion . . . . .                                                                      | 90        |
| <b>5</b> | <b>Conclusion</b>                                                                                 | <b>94</b> |
| 5.1      | Emergence of Patterns of Deviance in the Life Span . . . . .                                      | 94        |
| 5.1.1    | What have we learned about theories of deviance, conformity,<br>and the self? . . . . .           | 94        |
| 5.1.2    | Who are the "robust conformers"? . . . . .                                                        | 96        |
| 5.2      | The Possible Consequences of Punishment . . . . .                                                 | 97        |
| 5.2.1    | Misperceived costs of deviance. . . . .                                                           | 97        |
| 5.2.2    | Punishment and the self . . . . .                                                                 | 97        |
| 5.2.3    | Avoiding punishment: A learning process? . . . . .                                                | 98        |

|       |                                                                                                |            |
|-------|------------------------------------------------------------------------------------------------|------------|
| 5.3   | Deviance in Context: Perceived Meaning of Deviance and Social Psychological Outcomes . . . . . | 98         |
| 5.3.1 | Intrinsically vs. extrinsically motivated deviance . . . . .                                   | 98         |
| 5.3.2 | Degree of naturalism . . . . .                                                                 | 99         |
| 5.3.3 | Deviating with others . . . . .                                                                | 100        |
| 5.3.4 | Long-term effects of deviance . . . . .                                                        | 100        |
|       | <b>References</b>                                                                              | <b>101</b> |
|       | <b>A Supplementary materials for Chapter 2</b>                                                 | <b>112</b> |
| A.1   | Summary statistics for the pilot survey . . . . .                                              | 112        |
| A.2   | Factor analysis . . . . .                                                                      | 113        |
| A.3   | Identification with Princeton . . . . .                                                        | 114        |
| A.4   | Summary of the main survey analyses . . . . .                                                  | 115        |
|       | <b>B Supplementary materials for Chapter 3</b>                                                 | <b>116</b> |
| B.1   | Description of Study 1 and Study 2 samples . . . . .                                           | 116        |
| B.2   | Results of the Public Goods Game by Round . . . . .                                            | 118        |
| B.3   | Study 2 Results . . . . .                                                                      | 120        |
| B.4   | Effects Are Not Due to Learning Processes . . . . .                                            | 122        |
|       | <b>C Supplementary materials for Chapter 4</b>                                                 | <b>124</b> |
| C.1   | Instrumental Variable analyses . . . . .                                                       | 124        |

# List of Tables

|     |                                                                                                        |     |
|-----|--------------------------------------------------------------------------------------------------------|-----|
| 3.1 | Description of the sample . . . . .                                                                    | 56  |
| 4.1 | Psychological and social discomfort with the activity. . . . .                                         | 83  |
| 4.2 | Effect of treatment on inclination to deviate . . . . .                                                | 86  |
| 4.3 | Effect of treatment on possible psychological mechanisms . . . . .                                     | 89  |
| A.1 | Description of the sample . . . . .                                                                    | 112 |
| A.2 | Factor analysis . . . . .                                                                              | 113 |
| A.3 | Relative contributions to each component of the PCA . . . . .                                          | 113 |
| A.4 | Table of coefficients for Study 2 . . . . .                                                            | 115 |
| B.1 | Description of Study 1 sample . . . . .                                                                | 116 |
| B.2 | Description of Study 2 sample . . . . .                                                                | 117 |
| B.3 | Intent to Treat Analysis . . . . .                                                                     | 118 |
| B.4 | Intent to Treat Analysis . . . . .                                                                     | 119 |
| B.5 | Intent-to-Treat Analyses (Study 2) . . . . .                                                           | 121 |
| B.6 | Instrumental Variable Analyses (Study 2) . . . . .                                                     | 122 |
| C.1 | Effects wearing the plain dress on discomfort with the activity and<br>inclination to deviate. . . . . | 125 |
| C.2 | Effects wearing the plain dress on possible psychological mechanisms. . . . .                          | 126 |

# List of Figures

|     |                                                                                                     |     |
|-----|-----------------------------------------------------------------------------------------------------|-----|
| 2.1 | Perceived percentage of eating club members . . . . .                                               | 17  |
| 2.2 | Feeling different in the past and present predicts deviance . . . . .                               | 36  |
| 2.3 | Reasons for feeling different . . . . .                                                             | 37  |
| 2.4 | Deviating in the past and present predicts deviance . . . . .                                       | 38  |
| 2.5 | Belonging predicts deviance . . . . .                                                               | 39  |
| 2.6 | Self-monitoring predicts deviance . . . . .                                                         | 40  |
| 3.1 | Description of the studies . . . . .                                                                | 49  |
| 3.2 | Average contributions in the Public Goods game . . . . .                                            | 58  |
| 3.3 | Illustration of the treatment on the treated effect on Public Goods<br>game contributions . . . . . | 60  |
| 3.4 | Causal effect of the incentive to deviate in different games . . . . .                              | 63  |
| A.1 | Scree plot indicating variance captured by each component . . . . .                                 | 113 |
| A.2 | Identification predicts deviance . . . . .                                                          | 114 |
| B.1 | Public Goods Game by round (ITT) . . . . .                                                          | 120 |
| B.2 | Dictator Game (ITT) . . . . .                                                                       | 120 |

# 1

## Introduction

*“Departure from established, accustomed or expected ways of behaving is a prominent feature of human historical change.”*

– Muzafer and Carolyn Sherif

Public acts of deviance can be consequential for individuals and their communities (Durkheim, 1938; Erikson, 2005; Swann and Jetten, 2017). For individuals, deviating from social norms often comes with psychological and social costs. Social norms are defined as a person’s impression of the desirable attitudes and behavior of a group (Miller and Prentice, 1996), and past research suggests that people feel guilty or self-deprecate for violating norms (Chang et al., 2011; Schwartz, 1973; van Kleef et al., 2015). A large body of work also suggests that deviants tend to be stigmatized, marginalized, or relegated to the lowest social status of their group (Crocker et al., 1998; Jetten and Hornsey, 2010; Major and O’Brien, 2005; Monin and O’Connor, 2011). For communities, public acts of deviance often signal disapproval of ongoing practices and can generate social change. Paluck and Shepherd (2012) illustrated this phenomenon with an intervention fostering students’ public displays against harassment in their school. This intervention, which successfully changed students’ percep-

tions of social norms regarding peer-conflict and harassment, also changed patterns of collective behaviors.

Although deviance can play important roles for individuals and groups, modern social psychology emerged and developed with the study of the forces that lead people to conform, not deviate (Monin and O'Connor, 2011; Moscovici, 1976; Swann and Jetten, 2017; Turner, 2006). Intellectual momentum toward studying conformity was started by seminal studies exploring public conformity to the erroneous judgments of peers (Asch, 1955), conformity to negative social roles (Haney et al., 1972), and conformity to the unethical instructions of an authority figure (Milgram, 1963). Since that time, a major focus in social psychological research has been conformity to social norms, with studies highlighting that people most often conform to social norms and that perceptions of social norms often explain behavioral conformity within a group or community (Prentice, 2012). For instance, research has found that perceptions of social norms may influence the average age of marriage, women's level of education, rates of domestic violence, propensity to recycle, or peer harassment (Prentice, 2012; Tankard and Paluck, 2016).

This focus on conformity has left important questions about the determinants and consequences of deviance unanswered. While we have some idea of how a community may treat a deviant, and also the negative emotions deviants can feel after violating a social norm (Chang et al., 2011; Schwartz, 1973; van Kleef et al., 2015), scant research has asked: What are the social and psychological determinants of deviance? How does the experience of violating social norms shape human psychology and behavior?

In the present dissertation, I revisit questions about the determinants and consequences of deviance. In the remaining of this introduction, I review the relevant literature and make theoretical predictions about the factors that lead individuals to violate social norms, and the personal consequences of deviating from social norms. Second, using a field study in a university context with a cohesive student body and

an established social norm, I ask: who deviates from social norms, and why? Finally, I present the results of experiments in the lab and the field exploring the emergence of patterns of deviance in the life span. Specifically, these studies ask: does the experience of deviating from a social norm makes an individual more inclined to deviance?

## **1.1 Social and Psychological Determinants of Deviance**

Who becomes a deviant is driven by both contextual and individual factors. I hypothesize about these drivers by drawing on a wide literature from social psychology and sociology. Contextual factors that drive deviance may be very broad, as in the time period and culture (e.g., Swidler, 1986), or more local, as in the types of identity groups that surround an individual (e.g., Monin and O'Connor, 2011). Individual difference factors may include a person's level of identification with their group (Packer and Chasteen, 2010) personality traits (e.g., Snyder, 1979), or moral beliefs (Monin and O'Connor, 2011).

### **1.1.1 Contextual factors that may drive deviance**

Predicting which members of a given group will deviate from an established social norm may depend on a set of contextual factors. Social norms interact with the institutions, ideologies, symbols, and material products that together form a group's culture (Sewell, 2004), broadly defined as "a system of meaning linked to a set of available behavioral practices" (Paluck and Green, 2009, p. 623). Few perspectives in the psychology of deviance take seriously the contextual factor of culture, although a rich literature exists in sociology. For example, Swidler (1986) differentiates between settled and unsettled cultural periods when studying the power of social norms over behavior.

In settled times, norms are strong, injunctive, and privately endorsed by most ingroup members; few or no resources are available for those who may consider deviating. Swidler (1986) predicted and found through case study evidence that deviation from social norms may be nonexistent or extremely rare during settled times. Ingroup members may conform to social norms against their will, and underestimate the probability that others hold similar attitudes against the norm. This phenomenon, studied in psychology as pluralistic ignorance, constitutes a major obstacle to social change in groups (Miller and Prentice, 1996; Prentice and Miller, 1996). Swidler's (1986) research suggests that, because contextual factors are strongly in favor of status quo norms, deviance in settled times is more extraordinary.

By contrast, in unsettled times there is a greater opening for deviance, due to the emergence of structural opportunities for action against norms. In times of unsettledness or "periods of social transformation" (Swidler, 1986, p.278), resources emerge for those who hold attitudes against the norm. This perspective suggests that unsettled cultural periods may shake perceptions of pluralistic ignorance and facilitate deviance. Deviance may be more frequent during unsettled times, and may not need to involve individuals with extreme levels of moral resolve, or defiant stances toward their social groups.

Research on who deviates in settled times is a crucial component of our understanding of deviance. However, in order to better understand deviance as a broad and more ordinary phenomenon, it is also crucial to study deviance in less settled periods, when change is afoot or resources are available for change. Less research has examined deviance in these cultural periods.

Some psychological researchers have theorized about contextual drivers of deviance, including van Kleef et al. (2015) whose research demonstrates that high power predicts deviance, or Monin and O'Connor (2011) in their work on defiance who categorize deviance according to the intention of the actor. In this framework, when

one deviates unintentionally, deviance constitutes a status. For example, people may unintentionally deviate because they are forced by contextual factors like job loss or living in a foreign country. This research helpfully points out that deviance is not always chosen by an individual, in contrast to deviance as a choice, which is the subject of interest for this paper.

### **1.1.2 Individual factors that may drive deviance**

Perhaps because deviance is often exceptional and thus salient (Ross, 1977; Gilbert and Malone, 1995), or because those who deviate retrospectively report patterns of deviance in their life history, deviance has traditionally been considered a feature of human’s personality or a stable individual motivation or trait (Snyder, 1979; Smith et al., 1983; Bennett and Robinson, 2000; Briggs et al., 1980; Diefendorff and Mehta, 2007; O’Neill and Hastings, 2011; DeBono et al., 2011; Bernberg, 1954; Di Palma and McClosky, 1970; Elms and Milgram, 1966; Burger, 1987). This view of deviance has led scholars to study the stable characteristics of *chronic* norm violators, such as “rebels” (Gino, 2018), “rule-breakers” (Gelfand, 2018), “non-conformists”, (Grant, 2016), “outsiders” (Becker, 1963), or “trendsetters” (Bicchieri, 2017). This framework generally considers deviance as a stable trait that develops early in the life span. For instance, Gelfand (2018) argues that by growing up in a “loose” culture, individuals become “less attentive to social norms, more willing to take risks, more impulsive, and more comfortable with disorder and ambiguity” (Gelfand, 2018, p. 166). Other arguments about the early origins of deviance include the idea that being last in birth order makes one more rebellious and unconventional compared with other siblings because of differences in child rearing across sibling order (Grant, 2016, p. 154-155). In all cases, this framework suggests that certain individuals are more likely to violate norms than others, and that individuals who have often deviated in the past will keep deviating in future.

Psychologists have developed and tested specific theories of the personality traits that predict deviance. The theory of self-monitoring (Snyder, 1979), on the one hand, specifically considers the existence of enduring individual differences in the extent to which individuals regulate their behavior based on the expectations of their social context, with high self-monitors regulating more than low self-monitors (Lennox and Wolfe, 1984; Snyder, 1979). On the other hand, some studies have found associations between deviance and some of the Big Five personality traits. For instance, a study by Bodankin and Tziner (2009) suggests a relationship between deviance in the workplace, and traits such as agreeableness, neuroticism and conscientiousness. In their study, Bodankin and Tziner (2009) distinguish between constructive and destructive deviance, and suggest that the direction of the relationship between these personality traits and deviance depends on the norm at play.

Past research has also explored the motivations that lead some individuals to deviate from social norms. When people have the means and ability to conform but make the intentional decision to deviate, deviance can be considered a stance (Monin and O'Connor, 2011). Social psychologists have theorized that individuals may consciously decide to deviate from a norm because they consider it wrong, because they feel above it, or because they want to be at odds with it (Monin and O'Connor, 2011; Morrison and Miller, 2012).

Psychologists studying dissent, which may or may not include dissent from social norms, have suggested that other individual differences may drive deviance—in particular, levels of identification with an the group. On average, individuals who identify weakly with their group are more likely to voice out minority opinions (Hornsey et al., 2003). However, those who strongly identify with their group may publicly criticize a social norm, as long as they believe that the norm can be harmful for the group. This contrasts with low identifiers, who tend to remain silent or disengage when they notice that a norm may be harmful for the group (Packer, 2012).

Much research remains to be done in order to understand what leads individuals to choose to deviate from social norms. Accumulated research suggests that some individual differences matter, such as a person's group identification or personality traits (Kirsch et al., 2015; Packer, 2012). Most of the research on individual differences has been conducted in the laboratory. It will be interesting to note whether these relationships hold in a real-world context regarding a consequential decision about deviation. Further, it is impossible to know what types of predictors we may be missing as a result of not studying these types of decisions in the world. Monroe's (2008) work on heroes who broke with norms to save lives during genocide is a notable exception to this rule, but also may not inform theories of more ordinary, regular occurrences of deviation in the world.

Finally, while I am interested in the characteristics of individuals who choose to deviate, theory proposes that many individuals are marked as deviants even though they did not choose to deviate in any meaningful behavioral manner (Monin and O'Connor, 2011). These individuals may perceive themselves as different from others because of the words or actions of other people, rather than a choice they initiated. The relationship between feeling different from other people and deviating by choice is of interest. For example, individuals who choose to deviate may end up feeling different from others and as a result, be more likely to deviate again. Or more simply, deviance may become a pattern composed of a history of feeling different and acts of deviance by choice.

## **1.2 Consequences of Deviance on Psychological and Behavioral Outcomes**

How does the experience of publicly deviating from a social norm impact human psychology and behavior? Different theoretical frameworks support alternative hy-

potheses. On the one hand, classic theories of conformity support the idea that violating a norm make individuals more inclined to conformity. On the other hand, different lines of reasoning about the costs of deviance and theories of the self suggest the opposite: that violating a social norm makes an individual more inclined to deviance. This would suggest that deviance is more of a socially constructed pattern than a personality trait.

### **1.2.1 Violating a norm makes an individual more inclined to conformity**

Traditional theories of conformity suggest that people conform to social norms to achieve “fundamental goals to rewarding human functioning” (Cialdini and Goldstein, 2004, p.591). In this framework, human behavior is at the service of two main basic needs: maintaining affiliation to the community and maintaining a positive evaluation of the self (Cialdini and Goldstein, 2004; Cialdini and Trost, 1998; Leary, 1995; Steele, 1988). These fundamental goals are generally thought to be enhanced by conformity to social norms and thwarted by deviance (Cialdini and Goldstein, 2004).

**Conformity to maintain affiliation.** Because people most often disapprove of acts of deviance, violating norms generally limits an individual’s ability to build or maintain rewarding relationships with their community (Jetten and Hornsey, 2010; Monin and O’Connor, 2011). Past research has demonstrated, for instance, that violating social norms comes with negative social outcomes such as being punished, marginalized, or excluded (Festinger, 1950; Jetten and Hornsey, 2010; Link and Phelan, 2001; Major and O’Brien, 2005). These outcomes typically give rise to an increased need for belonging and social approval, which can be achieved by conforming to norms (Cialdini and Goldstein, 2004; Williams et al., 2000). For these reasons, theories that tie conformity to humans’ need for affiliation support the idea that after

violating a social norm, an individual would be more inclined to conforming to norms as a way to reinstate their affiliation.

**Conformity to maintain a positive self-concept.** Theories of norms also tie conformity to the self. Schwartz (1973) theorized that, particularly when a social norm becomes personal, “anticipation or actual violation of the norm result in guilt, self-deprecation, loss of self-esteem; conformity or its anticipation result in pride, enhanced self-esteem, security.” (Schwartz, 1973, p.353). More broadly, social psychologists have proposed that a central motivation to conforming to social norms is the anticipation of the pain of perceiving oneself as a deviant (Chaiken et al., 1996; Cialdini and Trost, 1998; Prentice and Miller, 1993). The idea that social norm violations have negative impacts on the self is also supported by recent social neuroscience research, which has for instance led researchers to formalize negative emotions such as guilt as deviation from a perceived social norm (Chang et al., 2011). Taken together, these bodies of work provide additional support to the hypothesis that violating social norms leads individuals to be more inclined to conforming to social norms, as a way to restore a positive self-concept.

### **1.2.2 Violating a norm makes an individual more inclined to deviance**

Different theoretical framework support the idea that the experience of deviating from a norm may make individuals generally feel like a deviant, adopt deviant views and opinions, or be more likely to engage in deviant behaviors. On the one hand, theories of the self directly anticipate these outcomes (Markus, 1977; Swann and Bosson, 2010, e.g.). On the other hand, reasoning about how people form expectations about the costs of deviating from norms also support the hypothesis that, in many cases, deviance may cause further acts of deviance, in private or in public.

**Reappraising the self in the community.** Theories highlighting the coherence and consistency of the self (Bruner, 1990; Markus, 1977; Swann and Bosson, 2010) suggest that individuals observe and rationalize their behaviors to make them part of a coherent narrative. After violating a social norm, individuals could rationalize their acts of deviance by either shifting their views of the community or integrating deviance as part of their identity or self-schemata. For instance, individuals may develop deviant private views of their community or start perceiving themselves as “the deviant type” after violating a norm, and this can make them more prone to deviance.

**Reappraising the costs of violating norms for the self-concept.** The experience of violating a social norm may make individuals realize that deviating did not feel as bad as they had anticipated. Past research suggests that individuals anticipate that public act of deviance will negatively impact their relationship with others or have negative consequences for the self, such as feeling of guilt, lower self-esteem, or increased instances of self-deprecation (Chang et al., 2011; Cialdini and Goldstein, 2004; Cialdini and Trost, 1998; Prentice and Miller, 1993; Schwartz, 1973). However, if milder than expected, the experience of violating a social norm may lead individuals to reconsider these costs for the self. In particular, lowering one’s expectations of how painful breaking a norm generally feels would, in theory, make the individual more inclined to deviance.

**Reappraising the costs of violating social norms for affiliation.** Similarly, if the experience of violating a social norm provoked milder reactions from other people than the norm violator initially anticipated, deviating could lead the individual to reappraise their perception of the costs of deviance for their relationship with their community. This mechanism may operate in two different but related ways. Individuals who notice that others did not respond as negatively as expected to their act of deviance may depreciate their perception of the strength of the specific norm

that they violated, in particular, its level of prescriptiveness. This could make them more inclined to violate similar norms in related contexts (Tankard and Paluck, 2016). After deviating, individuals may also develop different expectations of the probability that other people punish norm violations in general. I reason that if an individual noticed that others did not react as negatively as expected to their public act of deviance, this individual may infer that people do generally not react negatively to norm violations. These reappraisals of the perception of the strength of the norm or of others' reactions to deviance could lead individuals to be more inclined to be a deviant after deviating from a norm.

## 1.3 Outline

### 1.3.1 Determinants of deviance

In **Chapter 2**, I present the results of a field study conducted on the Princeton University campus. For this study, I conducted in-depth interviews of Princeton University upperclassmen who deviated from a historical and widely known Princeton norm: joining an “eating club,” a social group that undergraduates join at the end of their sophomore year. We explored the themes of these interviews with two rounds of surveys during the semester when students decide whether to join an eating club (pilot survey,  $N = 408$ ; and a random subsample of the pilot survey with 90% takeup,  $N = 212$ ). The surveys asked: what are the social and psychological determinants of deviance from norms? The data suggest that deviance is a pattern: compared to those who conform, students who deviate by not joining clubs report a history of deviance and of feeling different from the typical member of their social group. They also feel less social belonging and identification with Princeton and its social environment. Students who deviate are lower in self-monitoring, but otherwise are comparable to students who conform in terms of personality traits measured by the Big Five, and

of their perception of the self as socially awkward, independent, or rebellious. While some of these findings replicate past research, worth further exploration is the role of previous experience with deviance and its meaning for individuals as they decide whether to deviate.

### 1.3.2 Consequences of deviance

In **Chapter 3**, I present the results of two randomized experiments ( $N = 3,499$ ) exploring the behavioral and psychological consequences of deviating from norms. Participants played social dilemma games governed by norms of cooperation, without a mechanism for players to punish one another. In these studies, participants assigned to the treatment condition, unlike those in the control condition, were offered a monetary incentive to deviate from a strong norm of cooperation in one round of the game. The results suggest that violating this norm significantly increased participants' propensity to violate other social norms in new settings involving new groups and new games (with no incentive to deviate). In addition, post-treatment survey data suggest that violating norms causes people to depreciate their expectation of the costs associated with violation, but does not lead them to update their views of the self. Together, these results challenge traditional views of deviance as personality or a stable individual motivation or trait, and suggest that deviance can emerge from an experience of deviance at any point in one's life.

In **Chapter 4**, we take these findings to the field and explore the consequences of violating a norm in a "settled" culture (Swidler, 1986), in front of other community members. We deployed a field experiment in the Hasidic community in New York, in which non-religious norms are typically strong and prescriptive, and norm violations are rare. We recruited Hasidic women and randomly assigned them to complete one of two possible activities. Participants assigned to the deviance activity (treatment) were invited to publicly deviate from a strong social norm of clothing that is not tied

to Judaism. In contrast, participants assigned to the neutral activity (control) were invited to carry a religious book for a day. To identify the causal effects of public deviance, we asked participants ( $N = 155$ ) to complete a survey within the next day of their activity. The experience of violating a norm did not make participants more inclined to publicly deviate in future, but it made them more critical of their community and more supportive of deviance. Although the experience of deviating was psychologically challenging for the treatment group (e.g., feeling judged, negative affect), it led participants to perceive the norm that they violated as weaker and the costs of violating social norms in general as milder. Taken together, these results suggest that publicly deviating from a social norm in a settled culture may constitute a first step towards developing deviant private views and behaviors.

Finally, I conclude the present dissertation in **Chapter 5** with a discussion of the implications of the present findings for future research on deviance.

## 2

# Social and Psychological Determinants of Deviance

In this chapter, I revisit questions about the social and psychological determinants of deviance from peer influence and social norms. Using a field study in a university context with a cohesive student body and an established social norm, I ask: who deviates, and why? We designed a series of studies that are inductive and qualitative examinations of real-world deviance from a century-old, widely-followed social norm at Princeton University –joining an eating club<sup>1</sup>. We followed qualitative interviews with repeated surveys using representative sampling to assess the robustness of the findings from the qualitative interviews. This approach is in line with a pragmatist philosophy of science, which “urges scientists to observe what behaviors emerge in the complexity of real life” (Gantman et al., 2018, p.18), to describe the phenomenon at play, and to formulate hypotheses about the regularities surrounding the phenomenon (James, 1907; Rozin, 2001). This work is observational, and intended to be gener-

---

<sup>1</sup>We describe the eating club norm and its prescriptive component later on. To provide some context to the reader, eating clubs are student-led private institutions that upperclassmen may join. This is where most Princeton upperclassmen have meals and party. In 2016, 73% of Princeton upperclassmen were members of an eating club, which by definition makes it a descriptive norm.

ative of future hypothesis-testing and theory building about the causes of everyday behavioral deviance.

## **2.1 Study Context: An Established Social Norm at Princeton University**

For the past century, the great majority of Princeton undergraduate students have joined Princeton’s eating clubs, which are independent and private institutions where students have meals and parties. To quote an undergraduate student, becoming a member of one of the eleven eating clubs has been considered a “rite of passage.” In short, belonging to an eating club at Princeton is a social norm.

Technically, any Princeton student who wants to conform to this social norm by joining an eating club has the means and ability to become a member. Although the clubs are private institutions, Princeton increases students’ financial aid packages to assist students to join, in recognition of the eating club’s importance to undergraduate life. Furthermore, while some clubs are selective (“bicker clubs”), others (“sign-in clubs”) will accept any undergraduates who decide to join. In the spring of their sophomore year, all Princeton students make a choice to conform to the eating club norm by becoming a member, or to deviate by becoming an independent.

Historically the eating club norm could be characterized as part of a settled culture at Princeton (Swidler, 1986), because almost all students belonged and projected the idea that one should belong, as a Princeton student (characterizing a strong descriptive and injunctive norm, respectively). Also, few or no resources were available for those who would consider deviating, such as alternative cafeterias or kitchens in dorm rooms.

Today, however, the eating club norm might be characterized as part of a more unsettled culture at Princeton. Fewer students are joining eating clubs, in part because

the University is providing more alternatives for eating and socializing on the campus including co-ops, alternative eating groups where students cook and eat together. As a consequence, the norm is perceived “on the decline” by Princeton undergraduates, which is reflected in recent news articles about Princeton eating clubs, stating that eating clubs Presidents “are combatting declining membership” (Shashkini, 2018). This is also illustrated in the survey data that we collected in Spring 2017 from 408 Princeton students: the eating club norm is perceived by the students, on average, as weaker than it actually is. We asked the participants to estimate the percentage of eating club members on campus, and although the reality is that 73% of the upperclassmen were members in 2016 (L. Schmucki, personal communication, September 7, 2019), participants’ median and the mean response were approximately 10 percentage points lower (median = 65%, mean = 63%). Three quarters of all participants estimated that less than 70% of all students were eating club members (Figure 2.1).

Additionally, the norm of joining clubs is explicitly and publicly debated, as demonstrated by a recent referendum students held on the topic of whether students should have to “bicker” to gain entry to the club. The referendum did not pass, but it was notable as the first time in the history of Princeton that the tradition was up for debate. Statistics released by the eating clubs show that the proportion of sophomores who participated in the admission process between 2013 and 2016 nonetheless remained stable: between 79% and 81% of sophomores took part in the process each year.

Thus, the current climate at Princeton provides an excellent opportunity to study deviants, because they are growing in number even though the norm remains relatively strong, clear, and tied to the Princeton student identity. However, while the choice to become a deviant has grown more common, the experience could not be characterized as an easy one. While some independent students join an alternative eating group,

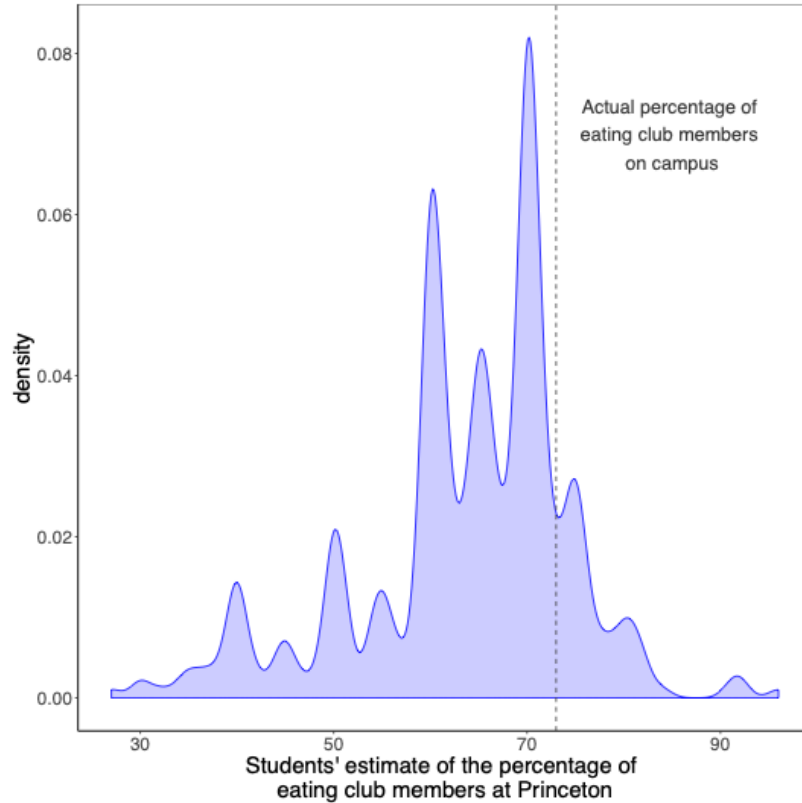


Figure 2.1: Perceived percentage of eating club members. This figure displays the density of the students' estimate of the proportion of eating club members on campus at the time of the study. The vertical dashed line represents the actual proportion of eating club members on campus.

for the most part students who deviate are not automatically part of an alternative social group for eating or socializing.

To conclude, Princeton has recently become an unsettled environment in relation to Eating Club, which provides us with the unique opportunity to study a larger sample of deviants at a time where the norm is clear, but debated. In this context, we take the opportunity to study: who deviates? What guides their choice to not join an eating club? Although some of the reasons why students decide not to join eating clubs may be specific to eating clubs themselves, this context constitutes a unique opportunity to study a reasonable sample of deviants (over 30% of the sophomore class), and provides an opportunity to derive more general lessons about the types of

contexts that give rise to a decision to stray from the norm and the types of people who make those decisions.

### **2.1.1 Overview of the studies**

In this chapter, we integrate qualitative and quantitative research to understand who deviates from a well-established social norm in a real-world setting. We ask: who deviates? How do they perceive their choice to deviate, how do they identify with their group, and how do they describe themselves and their history of deviation or conformity?

We first conducted in-depth interviews with all Princeton upperclassmen who responded to an advertisement to talk about their decision to join or not join an eating club at Princeton. The first author used a semi-structured interview approach, in which general questions guided the conversation but did not attempt to test hypotheses. The purpose of the interview was to elicit students' open-ended thoughts about their decision, their views of Princeton, their identities and previous experiences. We analyzed the general themes of these interviews, and used them to construct a pilot survey that targeted all Princeton sophomores as they considered whether to join an eating club. The objective of the pilot survey was twofold: first, identify the population of sophomore students who forecasted that they would deviate or conform to the eating club norm, and second, test the survey questions and scales that we designed after conducting the interviews on a larger sample of Princeton students. Following this, we conducted the main survey, for which we sampled a representative subset of the original sample, balanced by students who forecast that they would deviate vs. conform to the eating club norm. The main survey took place at the end of the students' semester, after they had been living with their decision to deviate or conform.

## **2.2 Study 1: Qualitative and Quantitative Data Collections on Students Who Decided to Deviate vs. Conform**

### **2.2.1 Study 1a: Interviews of deviant and conformist students**

The objective of this set of interviews was to gain a better understanding of the population of interest and learn about Princeton students' social life. Specifically, interviews with students who did or did not deviate from the eating club norm allowed us to gather crucial insights from the students about their perception of the clubs, the context of their decision to join or not join, and their lives more generally.

#### **Recruitment**

Since the objective of the interviews was to understand the reasons why a minority of students on campus decide to deviate from the eating club norm, we primarily recruited students who decided not to join an eating club. We also interviewed some eating club members in order to get a sense of the common reasons why students join the clubs. We recruited independents and eating club members separately, by advertising the study to different Princeton undergraduate mailing lists. The interviews lasted for 30 minutes on average, and participants were paid at a rate of \$12 per hour for their time.

We interviewed a total of 26 Princeton upperclassmen (8 juniors and 18 seniors). Overall, the students that we interviewed covered 16 different departments and majors (e.g., History of Science, Engineering, Psychology & Neuroscience, Journalism, French, Economics, and Computer Science). Interviewees were White (46%), Asian

(42%), Black (7%) and Hispanic (3%). The great majority of the interviewees were women; only 1 interviewee identified as a man.

## **Procedure**

We set up meetings in the psychology department with the juniors and seniors who agreed to participate. Interviews were semi-structured and conducted by the first author. The interviewer had a pre-determined set of open questions to ask to the participants, but allowed himself to explore some particular themes or responses whenever something relevant came up in the conversation. The interviews were not recorded; the interviewer took written notes during the session that he later typed.

We asked open questions to the independent interviewees about what it feels like to be an independent at Princeton, the reasons why they decided to go independent, their perception of eating clubs and the eating club norm at Princeton, and their identification with Princeton. In addition, we asked a set of questions about their background and experiences with past social groups. In particular, we asked whether they deviated from their group norms in the past, if their past social groups were different from Princeton, and how different they felt from the typical member of their past groups.

The purpose of Study 1a was not to identify systematic differences between the students who deviated vs. conformed to the eating club norm. To do so, we would have needed a much larger sample of Princeton upperclassmen. Instead, we conducted an inductive thematic analysis to identify recurring themes in the qualitative data (Braun and Clarke, 2006). We then used these themes to understand better the context of the eating club norm at Princeton on the one hand, and formulate hypotheses about the antecedents of norm deviance on the other hand.

Our thematic analysis strategy closely followed the guidelines introduced by Braun and Clarke (2006). During data collection, we went through a phase of iterative

reading of the interviews, during which we familiarized ourselves with the data. Every time we had conducted and transcribed two to four interviews, the authors would meet to read the interview transcripts together and discuss initial ideas about the content. By the end of this familiarization phase, our meeting notes included a list of themes that we had identified in the data. Once data collection was over, both authors reviewed the interview transcripts and meeting notes again to build together a list of the primary patterns of responses about: i) the reasons why participants decided to join or not join an eating club (e.g., the different costs associated with deviating from the norm), ii) their past and current social life, and iii) their tendency to deviate from other social norms. Then, we reviewed and named the themes (Braun and Clarke, 2006) by going through the interview transcripts again, and confirmed the prevalence of these themes in the data. Finally, we selected vivid examples that illustrate each of the themes that we identified (Braun and Clarke, 2006), which we report in this article.

## **Interview Findings**

We organize this findings section into three main themes that we identified from an analysis of interview notes. The first theme is about the social and logistical costs of deviating from the eating club norm at Princeton. The second is about the deviants' relationship with their past groups, and the third about their feeling of belongingness and identification to a current group, Princeton University.

**Theme one: The Social and Psychological Costs of Deviance From the Eating Club Norm.** Participants highlighted two types of costs<sup>2</sup> associated with deviating from the eating club norm at Princeton: social and logistical costs.

---

<sup>2</sup>In addition to costs, the participants brought up barriers to conforming to the eating club norm during the interviews, which stem from economic and religious constraints. In particular, students mentioned that they would not join eating clubs for materialistic (expensive membership) or religious reasons (concerns related to religious laws about diet and alcohol consumption).

*The social costs associated with deviating from social norms.* In sixteen of eighteen interviews, participants who were independents talked about feeling isolated, ignored, or left out by friends who belong to clubs. We quote three of these participants as examples of what most of them told us about how it feels to not be an eating club member on campus:

*“People who go independent are kind of marginalized and are perceived as not fitting. There’s also no cohesion between independents. The listserv was created last year, but in any case, you never know who else is going independent.”*

*“There’s a huge culture of eating meals with others here, and when you’re an independent you have to plan a lot more to see your friends. It’s kind of hard, but I manage to see them once a week.”*

*“Independents are ignored. It’s just a general vibe, perceived as less social or doing their own thing.”*

The participants also expressed feeling judged, even ashamed for being independent. One of them said that since she dropped out of her eating club, she feels that her social status is different, *“in particular with underclassmen who perceive independents as not super cool.”* Another student explicitly mentioned that *“people segregate because of eating clubs; they are judging you, judging if you are socially apt.”* On this note, a different interviewee shared the following story:

*“Recently at a Shabbat dinner, the eating club question came up. I had to say that I was an independent. I could see that they were judging me from the face they made. It makes me uncomfortable and angry, but I understand them in a way, because I used to be like that too.”*

***The logistical costs associated with deviating from social norms.*** Going independent not only involves this type of social cost, but also has negative logistical implications. One of the interviewees thought that undergraduates *“have no freedom of choice, because it is really hard to go to grocery stores, there’s not a lot of options and most of us don’t have a car.”* She added, *“the university is forcing the students to take the most conventional route.”*

Another participant shared thoughts about how the current structure makes it difficult to be an independent:

*“Independents are a little off the grid, even housing is on one side of campus. Freshmen and sophomores interact with a lot of people, then there’s no more community. Also, there’s no support on campus, I have to take a bus once a week to go get groceries, and that’s not easy.”*

Finally, a senior confessed:

*“I just want to graduate, leave this place, and be part of a world where you can eat with whoever you wanna eat with anytime.”*

**Theme Two: Independents’ past relationship with membership groups.** We explored how participants described their relationship with their social groups in the past. In contrast to eating club members, independent students emphasized having a history of both deviating from social norms and feeling different from others.

***Past instances of deviance.*** The interviewer asked the students if they felt that, before becoming a Princeton student, they were the type of person who would tend to not follow mainstream behavior. The majority of the independent interviewees responded to this question with ease, in contrast with the reaction of most of the eating club members. The content of the independents’ responses also appeared qualitatively different: three of the seven eating club members said that they mostly

conformed or always tried to fit in. Another one just said that she didn't party too much in high school. On the contrary, independent students expressed much more interest in the question and were eager to share stories. One of them responded: *"my whole life!"* Another student explained:

*"Yes. I was not a typical student in high school. I worked at a grocery store and I really liked the people at the grocery store. But in school, I never wanted to follow the crowd. Also, most people in my hometown get married very young, don't go to college, and don't even understand why I am going to Princeton."*

A third participant responded that she had not followed the mainstream much and told the interviewer:

*"Here at Princeton, I created and tailored my major in Visual and Performing Arts<sup>3</sup> and I am the only one who is doing that."*

A final example comes from a student who voiced out:

*"I don't take into account what other people think and do. I do what I want to do, and I'm not prone to social pressure. People in my band do a lot of stuff that I don't really understand, such as dating. I think that people go on date because of social pressure."*

***Feeling different from the typical ingroup member.*** Interview data also suggested that the independents, on top of reporting deviating from norms in the past, also mentioned feeling different from the typical member of their past groups. One student explained that she never felt like a *"typical person because she is a twin, gender non-conforming, and not the kind of typical Muslim person who fits at the*

---

<sup>3</sup>The real title of this participant's major was modified to protect their privacy; details from other quotes have also been changed for this reason.

*typical religious event.”* Another interviewee mentioned that she was never part of the main crowd. She explained: *“kids were looking at me differently. I performed better, because I was more intelligent.”* A senior said that she *“always felt different”*, and added: *“I had a different childhood, mostly because I lived in developing countries and that makes you grow up maybe too fast. I saw stuff that are hard to process for a 12-year old kid. I feel 5 years older than the students here.”* As a final example, one of the independent interviewees shared:

*“I always felt different. I am Malaysian, but I was born in New Zealand, from Chinese parents. My parents met in New Zealand, but then they moved to Malaysia. I don’t really feel part of any of these cultures really. I also went to a British school and never really felt part of that either. So I was always aware that I was not like everyone. But being at Princeton made it even more salient how unique my life is.”*

**Theme Three: Identification and belonging.** One of our objectives during the interviews was to explore how the independents would respond to questions related to their identification to Princeton. Although a few students responded negatively, the answers were overall positive.

We noticed that the independents tended to emphasize their institutional more than their social identification with Princeton. For instance, an interviewee specified, *“I do identify with my department, anthropology, and the pursuit of knowledge at Princeton.”* Another student expressed her love for architecture and said that she really threw herself into the department. A third independent responded, *“I try to identify with the best parts of what Princeton is. And a part of identifying with a group is to be able to criticize it. I won’t be back for each Reunions, but I identify, although I’m wary and critical.”*

In the surveys we explored further the idea that independent students may separate out the social and institutional dimensions of their identification to resolve the psychological conflict triggered by deviating from a well-established social norm.

### **Who are the deviants? Summary of the interview findings**

Our analysis of the exploratory interviews suggests that deviating from the eating club norm at Princeton may come with substantive costs of two types: social and logistical. Socially, most junior and senior independents reported feeling left out, ignored, and isolated from others. Some of them also expressed feeling ashamed or judged by others, which adds a prescriptive component to the eating club norm at Princeton. These social costs often coexist with logistical costs. For instance, most independents do not have access to a kitchen, and when they do, they need to make extra efforts to find groceries. Also, it is difficult logistically for independent students to have meals with friends who belong to eating clubs.

The interviews also allowed us to explore factors in the deviants' past and present social lives that may be responsible for their decision to deviate from the eating club norm at Princeton. Interviews with deviants, as opposed to eating club members, featured stories in which they deviated from norms in the past, and felt different from the typical member of their past social groups on dimensions such as intelligence, gender identity, and other characteristics. Finally, the interviews suggest that the deviants identified with Princeton as an institution, but described belonging less on a social level, particularly in comparison to conformists.

Overall, we concluded from the interviews that the setting at Princeton would be interesting for a more systematic investigation of the social and psychological characteristics of people who deviate from an ingroup norm. The setting features both costs of deviation (social and logistical), but also a relatively numerous group of people who do choose to deviate, which would make generalization about deviants

more reliable. Using our descriptive findings from the interviews, we designed a survey for Princeton sophomores who were currently deciding whether or not to join an eating club.

### **2.2.2 Study 1b: Students deciding to deviate or conform**

We used ideas generated from the interviews to create a pilot survey, which had two goals. First, we used the pilot survey to identify a large sample of prospective deviants and conformists among students on campus, meaning second year students facing their mid-year decision about whether to join an eating club or not. We planned to invite these students to participate in a subsequent survey at the end of the semester (Study 2) after they had been living with their decision to deviate or conform from the eating club norm. Second, we used the pilot survey to test the relevance of survey questions and scales for assessing some of the themes that arose in our interviews, with a broader population of students.

We fielded the pilot survey in the first week of the Spring semester, when students have their first opportunity to either conform to the eating club norm (“prospective conformists”), or deviate from the norm and become an independent (“prospective deviants”). To recruit participants, we sent an email to all Princeton sophomores from an undergraduate research assistant inviting them to participate in the survey in exchange for a \$5 Amazon gift card. The participants completed the survey online by clicking the link to the Qualtrics survey from the email. In order to encourage participation, we mentioned in the recruitment email that we would possibly contact them back at the end of the spring semester with the opportunity to take part in a \$20 follow-up survey. We collected as many respondents as possible, in an attempt to identify a relatively large proportion of students who planned to deviate. The survey remained open for six days: out of the 1,308 sophomores of this class year, 408 participated (i.e., 31% of the entire population of sophomores).

Although identifying systematic differences between the students who deviated vs. conformed to the eating club norm was not the primary objective of the pilot survey, we briefly describe the main correlates of deviance that appeared in the data. We include a table of item means in Appendix A (Table A.1) and make the data available at <https://osf.io/gxy6v/>.

In addition to asking the students if they were about to join an eating club, the survey asked about their past and present experiences with norms and social groups (e.g., past norm deviance, perception of norms, belongingness and identification, expectations about Princeton), and mental health and demographic characteristics. We pre-registered groups of explanatory variables, but we treated all predictors but belonging and identification as exploratory, given the nature of this survey as a pilot and as a tool for identifying the population of “new” deviants and conformists on campus.

## Main results

**Feeling different from others in the past and present (0 = feel very different 100= feel very similar).** Prospective deviants reported feeling more different from the typical ingroup member than prospective conformists by 10.42 percentage points ( $M_{deviants} = 47.08$ ;  $M_{conformists} = 57.51$ ;  $SD = 18.97$ ;  $p < .001$ ; Appendix A, Table A.1). This analysis suggests a significant positive relationship between feeling different from others and choosing to deviate, which supports qualitative data collected in the context of Study 1a.

**Deviance as a choice in the past and present (1 = never chose to deviate, 5 = always chose to deviate).** Prospective deviants reported deviating in the past from social norms at a greater rate than prospective conformists ( $M_{deviants} = 2.91$ ;  $M_{conformists} = 2.54$ ;  $SD = 1.04$ ;  $p = .002$ ; Appendix A, Table A.1). This find-

ing, in line with the interview results, suggests the existence of a significant positive relationship between choosing to deviate in the past and in the present.

**Belonging and identification (0 = I do not belong / identify at all, 100 = I belong / identify a lot).** Prospective deviants reported lower levels of identification with and belonging to Princeton University than prospective conformists by 6 percentage points ( $M_{bel.deviant} = 61.95$ ;  $M_{bel.conformists} = 68.05$ ;  $SD = 22.01$ ;  $p = .02$ ;  $M_{ident.deviant} = 71.82$ ;  $M_{ident.conformists} = 77.91$ ;  $SD = 21.83$ ;  $p = .02$ ).

## 2.3 Study 2: Main Survey of Deviant and Conformist Students

Using results from our pilot survey and combining those insights with our qualitative data, we launched our main survey toward the end of the spring semester after students had officially made their decisions about whether or not to join an eating club (i.e., to be a conformist or deviant, respectively). The objectives of the survey were descriptive, but we hypothesized several relationships between characteristics and experiences of the students on the one hand, and their decision to deviate on the other. The survey’s pre-registration can be found at <https://osf.io/gxy6v/>

### Theory- and data-driven hypotheses

Based on both theory and our data from Studies 1a and 1b, we formulated the following hypotheses.

**Hypothesis 1: feeling different from others in the past and present.** We expected deviants to report feeling more different from the typical members of their group in the past and in the present. This hypothesis emerged from the findings of both the interviews and survey data. This relationship has been suggested by theoretical work (Monin and O’Connor, 2011), but was never empirically tested.

**Hypothesis 2: deviance as a choice in the past and present.** We expected deviants to report deviating from social norms in the past and in the present at greater rates than conformists.

**Hypothesis 3: belonging and identification.** We expected deviants to have lower levels of identification with Princeton and to feel that they belong to the Princeton community less than conformists. Additionally, based on some hints from the interviews, we added items exploring possible differences between how the deviants' and conformists' identification and belongingness relate to the academic vs. social aspect of Princeton.

**Hypothesis 4: self-monitoring.** We expected deviants to have lower levels of self-monitoring than conformists.

**Hypothesis 5: Big Five personality traits.** We expected extraversion and agreeableness to negatively correlate with deviance, and neuroticism to be negatively associated with deviance.

**Additional measures.** Individual differences and mental health. In order to explore further how deviating from the eating club norm may be related to one's self-perceived willingness and ability to socialize, we asked the participants about the extent to which they see themselves as: socially awkward, a rebel, an independent thinker, and a maverick. In addition, we measured participants' mental health, which is a typical correlate of belongingness (Cialdini and Goldstein, 2004), hypothesized to be a predictor of deviance from the eating club norm in this study.

## **Recruitment strategy**

At the end of spring 2017, we contacted all prospective deviants ( $n = 116$ ) and a representative subsample of the prospective conformists ( $n = 120$ ) from Study 1b, and invited them to participate in a follow-up survey (i.e., Study 2). We offered a \$20 Amazon gift card for their participation in the 30-minute follow-up survey. We

deliberately made this compensation large in order to limit attrition. In the end, a total of 212 “deviating” and “conforming” sophomores took the survey, that is, 90% of those who were recruited. Although attrition is rather small (i.e., 10%), we conducted robustness checks using inverse probability weighting, in which we attribute greater weight to the participants who are most similar to those who are missing in our final follow-up survey sample. Our results are the same with this procedure and so below we present only the unweighted results.

## Materials and Procedure

At the beginning of the survey, the participants were asked if they joined an eating club. This binary variable returns 1 for deviants, 0 for conformists, and constitutes the norm deviation indicator that we use in subsequent analyses. Out of the 116 prospective deviants that we recontacted, 109 participants (i.e., 94%) reported having actually deviated from the norm. Deviants (vs. conformists) constituted 50.4% of our main survey sample.

Subsequently, the participants responded questions exploring different characteristics and experiences we hypothesized would correlate with their decision to deviate from vs. conform to the eating club norm. At the end of the survey, the participants were asked to answer demographic questions (gender, ethnicity, family income and religiosity), which we use as covariates in all regressions. We now turn into describing the survey in more detail.

**Feeling different from others in the past and present.** The objective of this series of questions was to explore the idea that feeling different from the typical ingroup member in the past may increase one’s probability of deviating from norms. We first asked the participants to name their most important past membership groups and asked them: i) “did you usually feel like you fit in?” (0 = not at all; 100 = completely); ii) “how often did people from these groups intentionally make you feel

different?” (0 = never; 100 = all the time); iii) “Before becoming a Princeton student, in your social life, how often did you feel different from other people?” (scales: 1= never; 4 = always); iv) “At what age do you think you started to feel different from other people?”; v) “Do you feel that you have a lot in common with the average Princeton student?” (0 = I have nothing in common; 100 = I have a lot in common).

We also explored the different ways in which respondents have felt different from other members of their social groups. We asked the participants to select all the characteristics that made them feel different from others in the past from the following list: religion, race, values, gender identity, sexual orientation, ideas, personality, skills, habits, and interests.

**Deviance as a choice in the past and present.** We asked participants the following three questions: i) “How often do you not behave in accordance with the mainstream or socially expected behavior at Princeton?” (0 = never; 100 = all the time); ii) “Prior to arriving at Princeton, how often would you behave in opposition to mainstream social behaviors?” (0 = never; 100 = all the time); iii) After asking participants to think about the first time they deliberately deviated from a norm, we asked: “how old do you think you were?”.

**Belonging and identification.** The series of questions about belonging and identification were composed of eight items: four items dedicated to belongingness, and the other four items dedicated to identification. Both subseries of four items were formulated in the exact same way, using the same scale from 0 (not at all) to 100 (a lot). The series of questions about belongingness came first, in the following order: i) “To what extent do you feel that you belong to the Princeton community, in general?”; ii) “To what extent do you feel that you belong to the Princeton community, socially speaking [iii) academically speaking]?”; iv) “Is it more important to you to belong to Princeton socially or academically?”

Following this series of questions about belonging, the participants responded to the same series of question about identification (e.g., “To what extent do you feel that you identify with the Princeton community, in general?”, etc.)

**Personality.** We used the Ten Item Personality Measure (Gosling et al., 2003) to assess the Big five personality traits of the participants, who were asked to rate the extent to which they perceived themselves as, for instance: extroverted and enthusiastic; critical and quarrelsome; or open to experience and complex. To assess the participants’ level of self-monitoring, we used items from the Self-Monitoring Scale (Lennox and Wolfe, 1984,  $\alpha = .65$ ).

**Mental Health.** We assessed the participants’ mental health through self-report of sadness and depression on the one hand (e.g., “Please select the options that most closely describe the level of depression or sadness that you have been feeling in the past 2 weeks: none, mild, moderate, moderately severe, severe”), and happiness on the other hand (i.e., “How happy are you”, scale from 1 = not at all happy to 100 = extremely happy). We created an index of negative mental health ( $\alpha = .81$ ) by averaging three items: the participants’ self-reports of sadness, depression, and happiness (reverse coded).

**Demographics.** We asked the participants for their gender, race, family income, and religiosity. We controlled for these variables in each of the regression analyses presented below.

**Unreported survey items.** We also measured other potential social and psychological correlates of norm deviation, such as levels of anticipatory socialization, items from the Self-Determination Scale (Sheldon, 1995), and emotions about the eating club norm, which were for exploratory purposes and not reported here (but see online materials on the Open Science Framework<sup>4</sup> for the full survey).

---

<sup>4</sup>osf link: <https://osf.io/gxy6v/>

## Study pre-registration

This study was pre-registered on the Open Science Framework: <https://osf.io/gxy6v/>

## Results

Our objective was to understand and describe the relationships between deviance decisions and variables describing students' social memberships, past instances of deviance, and their personality, rather than to build a predictive model. As such, we run multiple linear regressions<sup>5</sup> of the binary deviance indicator (Hellevik, 2009) on each of these variables separately. Bivariate models also recognized the fact that many variables correlated with one another, creating problems of multicollinearity<sup>6</sup>. For each regression, we used robust standard errors and controlled for demographic variables (i.e., gender, ethnicity, family income, and religiosity).

$$y_i = \beta_0 + \gamma P_i + X_i\beta + \epsilon_i \quad (2.1)$$

in which  $y_i$  indicates if the  $i$ 'th participant is a deviant,  $\gamma$  is an estimate of the effect of the predictor of interest on individual  $i$ 's probability to deviate,  $P_i$  is the predictor of interest for the  $i$ 'th participant,  $X_i$  is a vector of preregistered demographic variables,  $\beta$  is a vector of covariate effects, and  $\epsilon_i$  is an error term.

For each outcome variable of interest, we first conduct a multiple regression using the aforementioned model (Hellevik, 2009). When we run separate multiple regressions on variables that are part of the same theme (e.g., "deviance as a status in the past and present"), we use the Bonferroni correction to compute adjusted p-values.

---

<sup>5</sup>The findings described in the results section are robust to logistic regression analyses. Because the objective of these analyses is to estimate the average marginal effects of our predictors of interest in terms of probabilities, we report the results of linear regression analyses (see Gomila, 2019; Woolridge, 2002, pp. 455-456; e.g., Hoffman et al., 2016).

<sup>6</sup>We used Factor Analysis (Appendix A, see Tables A.1, A.2, and A.3; and Figure A.1) to confirm that the items that we used to test each hypothesis contribute together to separate components. This analysis revealed 4 principal components that together, explain 75% of the variance. PC1 can be defined as social belonging and identification, PC2 as academic belonging and identification, PC3 as feeling different in the past and present, and PC4 as deviating in the past and present.

All the significance levels of the predictors are robust to this correction. Significance stars in the Results section reflect adjusted p-values. Finally, we illustrate the relationship between deviance and the independent variable under consideration by computing and displaying the probability of an individual’s deviation for different values of the independent variable, maintaining the pre-specified demographic covariates at their mean.

**Deviance as a pattern.** In line with Hypothesis 1, the set of regressions examining the difference between deviants’ and conformists’ past and present experiences with their group suggests that deviants tend to feel less similar from the “typical Princeton student” than conformists ( $b_1 = -.006$ ,  $SE_{robust} = .001$ ,  $p_{adj} = .002$ ;  $R_{adj}^2 = .17$ ; Figure 2.2).

The data also support the idea that the deviants have a history of feeling more different from other people in their groups compared to the conformists (Hypothesis 1). Self-reports of feeling different from the typical group member are recalled at a younger age by deviants ( $b_1 = -.02$ ,  $SE_{robust} = .009$ ,  $p_{adj} = .035$ ;  $R_{adj}^2 = .14$ ; Figure 2.2). Also predicting deviance are self-reports of perceiving oneself as not fitting well in their past social groups ( $b_1 = -.005$ ,  $SE_{robust} = .002$ ,  $p_{adj} = .035$ ;  $R_{adj}^2 = .14$ ; Figure 2.2) and that former group members made one feel different ( $b_1 = .003$ ,  $SE_{robust} = .001$ ,  $p_{adj} = .035$ ;  $R_{adj}^2 = .14$ ; Figure 2.2).

Although deviants are more likely to report feeling different from others than are conformists, and that they started to feel this way at a younger age compared to conformists, Figure 2.3 reveals that deviants do not have significantly different reasons for feeling different. When asked for the reasons they sometimes felt different from others, Figure 2.3 reveals that deviants and conformist report at similar rates feeling different from other members of their social groups because of their personality, values, gender identity, and religion. Descriptively, those who deviated from the eating club norm picked more dimensions along which they felt dif-

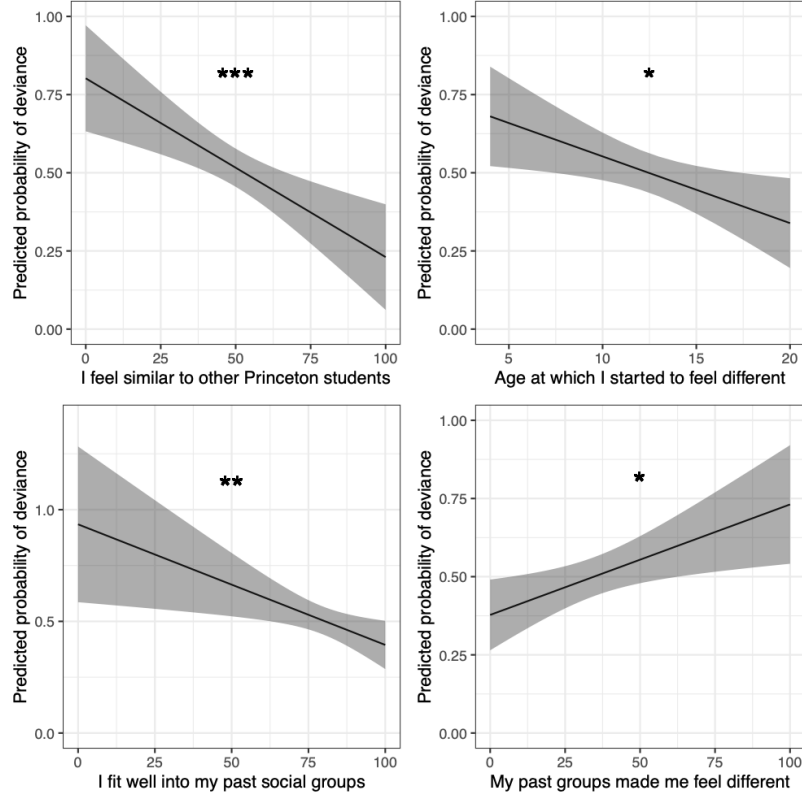


Figure 2.2: Feeling different in the past and present predicts deviance. These predicted probabilities of deviance are derived from separate regressions of the binary eating club deviance item on each self-reported item of deviance as status in the past and present, controlling for demographic variables set at their mean.

ferent than those who conformed, but the difference is not statistically significant ( $M_{deviants} = 4.01; M_{conformists} = 3.63; 1.55; p = .12$ )

In line with Hypothesis 2, deviants reported more active choices to deviate from social norms in the present and past, compared to conformists (see Figure 2.4). The results revealed that deviants were significantly more likely to report that they chose to deviate earlier in their life ( $b_1 = -0.03, SE_{robust} = .008, p_{adj} < .001; R_{adj}^2 = .18$ ). Specifically, deviants reported starting to deviate at a significantly younger age, on average 11.3 years old, than conformists, who reported starting to deviate at 13.4 years old on average. In addition, deviants reported deviating at higher rates than conformists in the past, although this difference does not reach conventional levels of

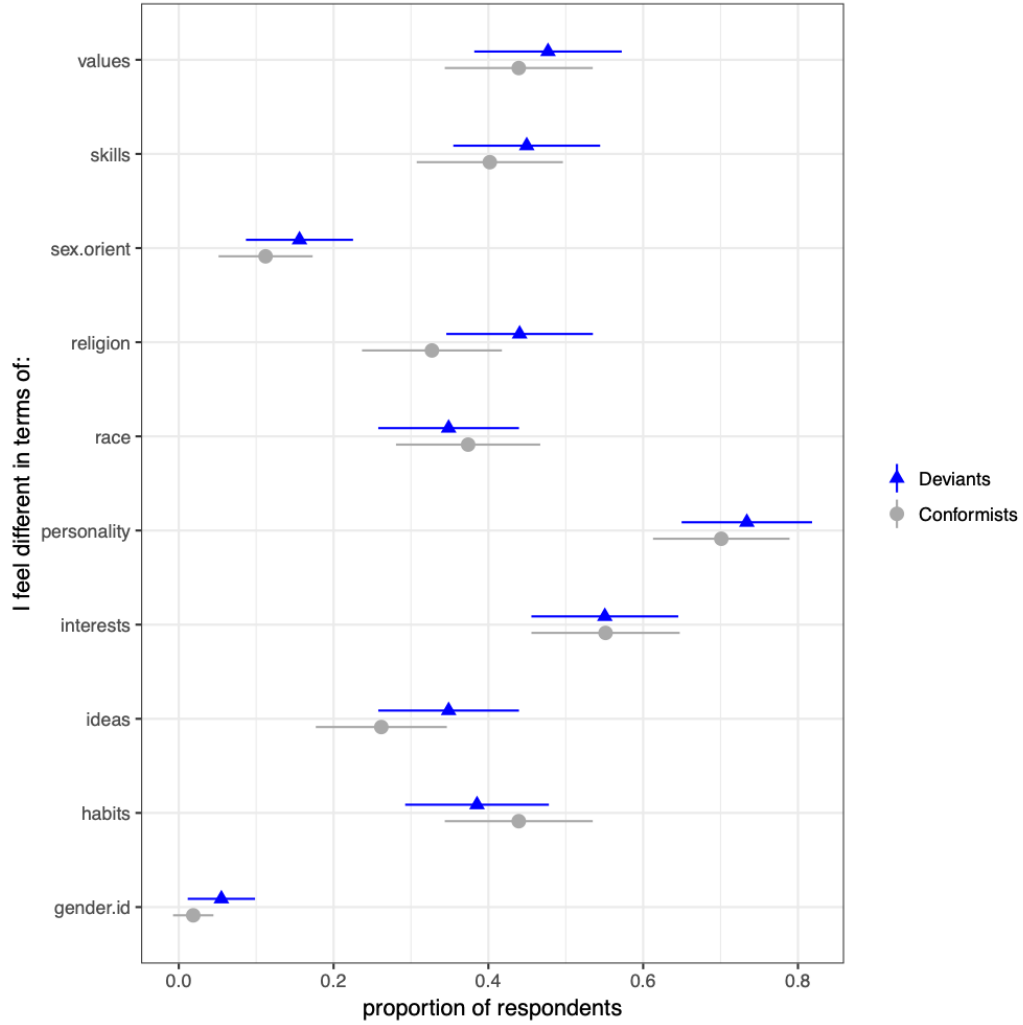


Figure 2.3: Reasons for feeling different. Proportion of deviants and conformists who reported that the items displayed on the y-axis (e.g., personality, values, race, etc.) have made them feel different from other members of their social groups, and 95% confidence interval bars.

statistical significance ( $b_1 = 0.04$ ,  $SE_{robust} = .03$ ,  $p = .12$ ;  $R_{adj}^2 = .14$ ). Finally, deviant students also reported that they currently choose to deviate from social norms other than eating club norms compared to students who conform to the eating club norm ( $b_1 = 0.006$ ,  $SE_{robust} = .001$ ,  $p_{adj} < .001$ ;  $R_{adj}^2 = .19$ ).

**Belonging and identification.** The results also support Hypothesis 3, demonstrating that deviants report lower levels of belongingness and identification than conformists ( $b_{1.belonging} = -0.005$ ,  $SE_{robust} = .001$ ,  $p_{adj} = .001$ ;  $R_{adj}^2 = .16$ ; Figure 2.5;

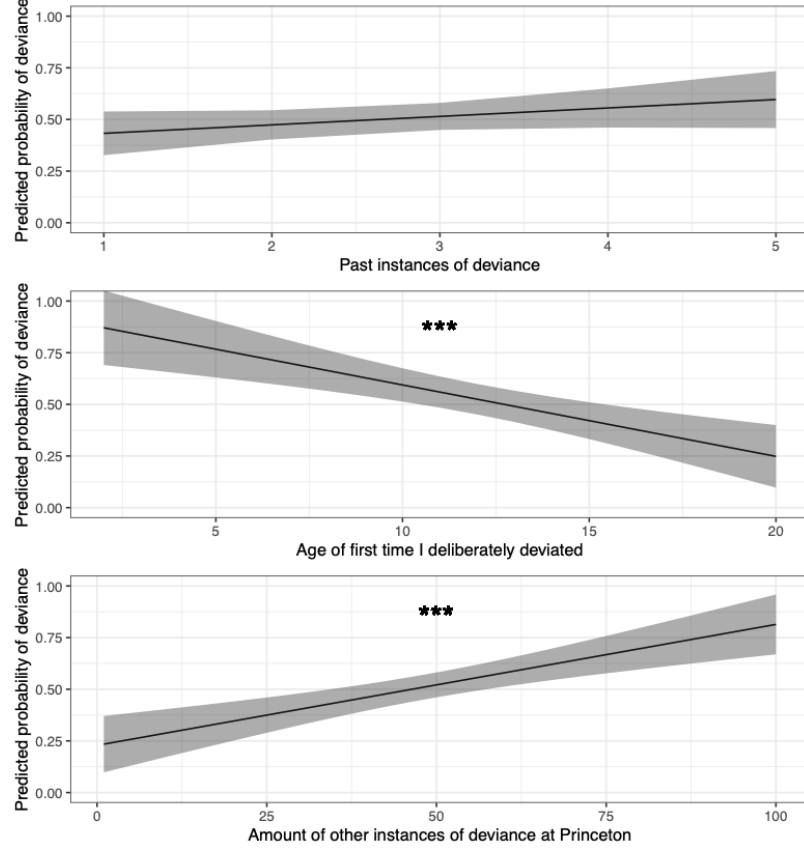


Figure 2.4: Deviating in the past and present predicts deviance. These predicted probabilities of deviance are derived from separate regressions of the binary eating club deviance item on self-reported deviance from other norms in the past and present, controlling for demographic variables set at their mean.

$b_{1.identification} = -0.004, SE_{robust} = .001, p_{adj} = .005; R_{adj}^2 = .15$ ). For instance, the regression model predicts that, compared to those who feel that they belong “a lot” to the Princeton community, the students who feel that they belong “a little” to the Princeton community are more likely to deviate from the eating club norm by around 35 percentage points. The magnitude, direction and significance of the relationship between deviance and identification items mirrors those of belonging, and are displayed in Appendix A (Figure A.2).

Although we did not find significant differences between the deviants and the conformists in terms of academic belonging ( $b_1 = 0.002, SE_{robust} = .002, p = .2; R_{adj}^2 = .11$ ) and identification ( $b_1 = -0.0009, SE_{robust} = .001, p = .55; R_{adj}^2 = .11$ ), de-

viants reported prioritizing their academic belonging over their social belonging ( $b_1 = 0.006, SE_{robust} = .002, p_{adj} = .001; R_{adj}^2 = .20$ ), and their academic identification over their social identification ( $b_1 = 0.007, SE_{robust} = .001, p_{adj} < .001; R_{adj}^2 = .17$ ).

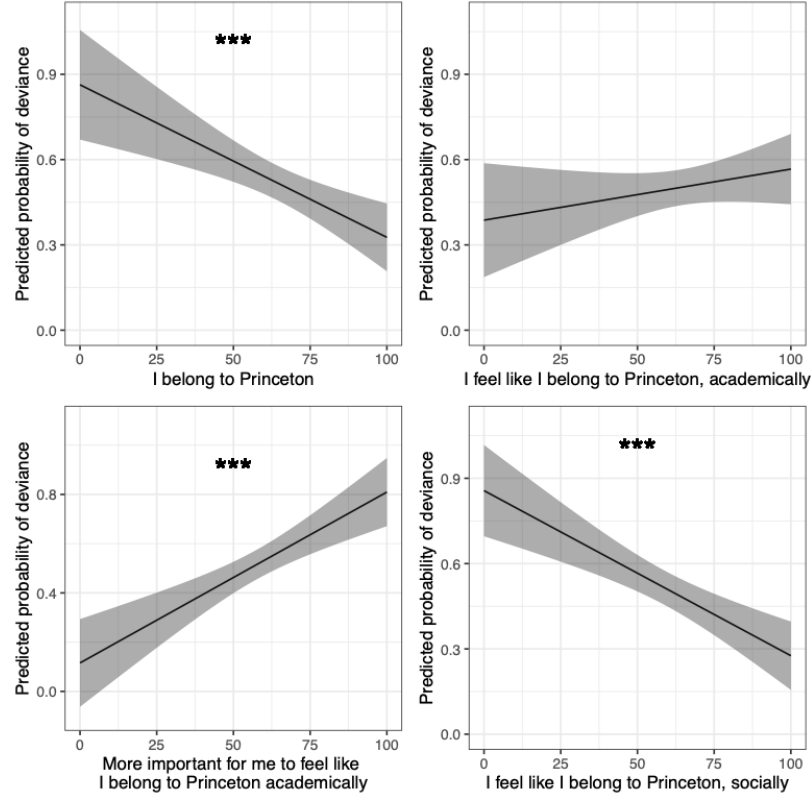


Figure 2.5: Belonging predicts deviance. These predicted probabilities of deviance are derived from separate regressions of the binary eating club deviance item on self-reported belonging items, controlling for demographic variables set at their mean.

**Self-monitoring and personality traits.** The results supported Hypothesis 4, with deviant students scoring significantly lower in self-monitoring than conformist students ( $b_1 = 0.18, SE_{robust} = .056, p = .002; R_{adj}^2 = .15$ ; Figure 2.6). However, contrary to Hypothesis 5, there were no differences between conformists and deviants for any of the items of the Big five, that is, extraversion ( $b_1 = -0.027, SE_{robust} = .02, p = .15; R_{adj}^2 = .12$ ), neuroticism ( $b_1 = 0.00, SE_{robust} = .02, p = .99; R_{adj}^2 = .11$ ), openness to experience ( $b_1 = 0.02, SE_{robust} = .03, p = .48; R_{adj}^2 = .11$ ), conscientiousness ( $b_1 = 0.002, SE_{robust} = .02, p = .92; R_{adj}^2 = .11$ ) and agreeableness ( $b_1 =$

$-0.02$ ,  $SE_{robust} = .03$ ,  $p = .44$ ;  $R_{adj}^2 = .12$ ). Similarly, we found no difference between deviants and conformists in terms of how much of a maverick ( $b_1 = 0.01$ ;  $SE_{robust} = .02$ ,  $p = .75$ ;  $R_{adj}^2 = .11$ ), a rebel ( $b_1 = 0.02$ ,  $SE_{robust} = .02$ ,  $p = .25$ ;  $R_{adj}^2 = .12$ ), an independent thinker ( $b_1 = 0.04$ ,  $SE_{robust} = .02$ ,  $p = .14$ ;  $R_{adj}^2 = .12$ ), or how social awkward they feel ( $b_1 = 0.03$ ,  $SE_{robust} = .02$ ,  $p = .09$ ;  $R_{adj}^2 = .13$ ).

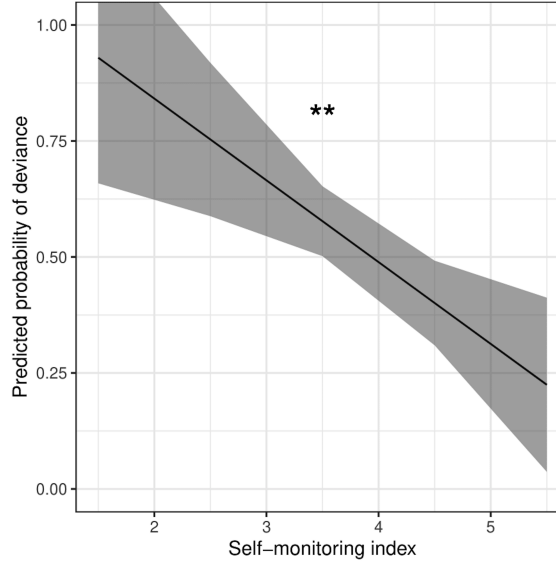


Figure 2.6: Self-monitoring predicts deviance. These predicted probabilities of deviance are derived from separate regressions of the binary eating club deviance item on the self-monitoring index, controlling for demographic variables set at their mean.

**Mental health.** Deviants and conformists did not differ in terms of their mental health ( $b_1 = -0.04$ ,  $SE_{robust} = .03$ ,  $p = .18$ ;  $R_{adj}^2 = .12$ ; Appendix A, Table A.4).

## 2.4 General Discussion

We used an established Princeton University norm (i.e., joining an Eating Club) to which 73% of the undergraduate students adhere, to examine the social and psychological correlates of deviance from social norms. Exploratory interviews of junior and senior undergraduates revealed that deviating from the eating club norm at Princeton by refusing to join an eating club comes with substantive social costs. This

suggests that the norm is prescriptive (students should join eating clubs), and not just descriptive (most students do join clubs).

However, deviance from the eating club norm may have become relatively easier than it used to be because of recent developments. The university offers financial subsidies to students who do not join eating clubs, and increasing numbers of students are choosing to remain independent from clubs. This may make the eating club norm seem less “settled” (Swidler, 1986). In addition, our data suggest this to be the case: students perceive the norm to be weaker (fewer students joining eating clubs) than it actually is. As a result, our survey findings, which highlight the personality and social experiences that differentiate students who deviate from the students who conform, paint a picture of deviance from a social norm that is undergoing change. As a result, while deviance from the eating club norm still comes with costs, the present research focuses on more “ordinary” deviance. This feature of this research addresses a gap in the literature on more “ordinary” deviance, as opposed to previous scholarship on deviation from strong and highly punitive norms (e.g., rescuing Jews during the Holocaust; Monroe, 2008). During “unsettled” times, the norm seems to be on the decline and structural opportunities exist for those who decide to deviate.

Our qualitative and quantitative data consistently suggest three themes regarding those who deviate. First, deviance is a pattern. Individuals who decide to deviate from social norms deviated more in the past, started to deviate earlier, and deviate more in the present. These results are true for two different experiences of deviance: being made to feel different by members of previous social groups, and intentionally deviating from norms of previous and present social groups. Second, individuals who choose to deviate tend to be low identifiers who feel that they belong to their social group less than those who decide to conform. Third, we found that deviants tend to be lower self-monitors (Snyder, 1979) than the conformists, but we found no

differences in personality traits as measured by the Big Five, or in levels of mental health between deviants and conformists.

On the one hand, the finding that people deviating in the present also report deviating in the past seems intuitive and in need of no explanation: individuals strive to maintain a coherent and consistent self (Fiske, 1993; Ross et al., 2010; Swann and Bosson, 2010), which may explain that past behavior is a good predictor of future behavior. On the other hand, patterns of deviance seem to need more explanation. Most research on social norms focuses on the rewards of following social norms and the strong costs of violating them. Previous findings, as well as our own, show that deviance comes with high social and psychological costs (respectively, social sanctions, and feeling odd or guilty for deviating; Almenberg et al., 2010; Asch, 1955; Cialdini and Goldstein, 2004). Related research suggests that social sanctions effectively promotes more conformity and fewer instances of deviance on a collective level (Fehr and Fischbacher, 2004).

On an individual level, there is less research, and a possible question for future experimental research is why individuals choose deviance repeatedly over the course of time. For example, do some people who repeatedly deviate from norms over time have a different way of predicting or dealing with the social and psychological costs? Individuals in general may overestimate social sanctions incurred by deviance. For those who actually deviate, they may discover that they experience lower social costs than expected.

Deviance may become part of one’s identity through the development of narratives around the deviant self. An experience of deviation as a choice (or as something imposed on them, as we observed in our survey), may shape the individual’s own perception of the self as deviant. Individuals may present themselves as deviant by status (e.g., “I’ve always been different from others”), or by choice (e.g., “I’m the type of person who likes to do different things”). These themes have been previously

explored by Monin and O'Connor (2011), who theorized about deviation by choice vs. status; and by Blanton and Christie (2003), who theorized and showed how individuals prefer to define the self as deviants in positive ways from others (“sticking out in good ways”; 2003). These narratives and identities may influence one’s own and others’ expectations for and actual instances of deviant behavior.

Those who deviated from the eating club norm at Princeton feel that they socially belong and identify less to the typical Princeton student. Deviants and conformists both feel that they belong and identify with Princeton academically, but deviants were more likely than conformists to say that they prioritized academics over social life. One possibility is that those who deviate may compensate their decreased level of social belonging by increasing their feelings of academic and institutional belonging. This may be one strategy to minimize the costs of deviation for future investigation.

## **Limitations**

One limitation of this study is that we study deviance from one specific type of norm, which is highly social—membership in a club. Thus, our description of deviants from this norm may only apply to the type of people who do not wish to be as social as others, or who do not get along as well with others. This is a possibility, and is one more reason why we need more research on the topic of deviance. However, we also point out that we did not find personality differences that one might expect if deviance in this specific case had everything to do with degree of sociality—in particular, we do not find differences in terms of personality traits as measured by the Big Five (Bodankin and Tziner, 2009). Furthermore, those who deviated from the eating club norms did not differ from others in their perception of the self as socially awkward, as independent thinker, or as rebellious. Finally, deviants and conformists did not differ either in the way they feel different from other people, and we did not find differences in terms of mental health or academic identification or belonging.

Still, we hope that this research inspires similar descriptive and experimental research exploring other types of norms in other settings.

A limitation to causal inference is that the series of studies described in this paper are not experimental. The objective of the set of studies presented in this article was to explore the social and psychological correlates of deviance, not to make causal or directional claims about the causes of deviance. In this sense, we did not manipulate any treatments and our results rely only on participants' self-reports about their past and present, which implies that they might have erroneously retrieved and reported information from their past. More research, especially experimental research, is needed to understand whether individuals who deviated from a norm build narratives around their past and present self that are coherent with their deviant behavior, such as reporting more acts of deviance in the past, or whether deviance is part of a causal loop in which feeling different from others and deviating begets deviance.

Although we cannot rule out the possibility that demand characteristics may have driven the differences between deviants and conformists, it is unlikely. During the recruitment period, we never mentioned to the students that this study was about deviance from norms or eating clubs. We deliberately called the surveys "the sophomore survey", and made sure that the question asking the participants about their decision to join an eating club was part of a broader set of general questions. As a result, we believe that the participants were not actively seeking to respond to the object of the study.

These descriptive results point the way toward more research on social norm deviance. First, researchers should seek to understand how and why the social and psychological costs that one endures after deviating correlate with their probability to deviate again. Experimental designs in various settings and that involve deviance from different kinds of social norms are necessary to understand what kinds of char-

acteristics and experiences are general drivers of deviance, and which are specific to context. Second, future research should explore if and how feeling different from other members of one's group may increase one's probability to deviate, and why. Third, because acts of deviance often come with negative consequences, individuals may update their perception of the self, or of the costs, after deviating from a social norm. What types of narratives do deviants develop about the self or about the act of deviating, that shape their own and others' expectations about future behavior? Investigations into these and other questions will greatly contribute to our understanding of patterns of deviance.

# 3

## Experiments Testing the Consequences of Deviance in Behavioral Games

### 3.1 Studying the Consequences of Deviance in Causal Terms: Challenges and Approach

In the present chapter, I ask whether a pattern of deviance can emerge from one experience of deviance and why. We designed two studies aiming to identify the causal effects of violating social norms on three main outcomes: i) individuals' behavioral choice to violate social norms again, ii) individuals' self-perception as a deviant, and iii) individuals' expectations of the costs of violating social norms. In the past, researchers who studied deviance typically compared the characteristics or outcomes of individuals who were already more or less likely to violate social norms (judged by some individual difference). The present experiments are, to the best of my knowledge, the first to focus on identifying the consequences of deviance in causal terms using experimental methods. This strategy can be challenging because researchers

need a setting that allows them to randomly assign individuals to violate social norms, which is difficult to operationalize.

### **Logic of the studies**

To study the causal effects of deviance, we conducted two randomized experiments with 3,499 participants (summing both sample sizes) who played a series of behavioral games for real money<sup>1</sup>. We assigned a random half of them to receive an incentive to violate a strong, explicit, and uniformly-followed social norm of cooperation.

Participants first played the Public Goods game, in which cooperation to contribute money to a common pool results in the most evenly distributed money for all players. To make a norm of cooperation clear to all participants, the high rate of conformity to a cooperative norm was displayed on the screen during each game round. In the final round of one game, participants in the treatment condition (but not control condition) received a private message promising a monetary incentive to violate the social norm of cooperation of their group by contributing a sum far beneath the group norm (which would result in a greater monetary benefit to themselves on top of the incentive). We then observe: are participants induced to violate the cooperative norm more likely to break norms in future? After violating a social norm, do individuals update their expectations of the costs of deviance, their perception of the self as deviant, or both? To answer these questions, we compare control and treatment participants' post-treatment contributions to different behavioral games (Public Goods game, Dictator Game, and Trust Game) with new groups and players, as well as their responses to survey questions.

---

<sup>1</sup>Participants played with Monetary Units (MUs). The amount of MUs that they accumulated across studies was converted into U.S. Dollars at a rate of 2,000 MUs per \$1. at the end of the study, and transferred to participants as a bonus.

## 3.2 Method

### 3.2.1 Overview of the design

In both studies<sup>2</sup> participants first played multiple rounds of the Public Goods game (PGG) with three separate groups (G1, G2, and G3) of four players in total (Figure 3.1). We used this game to measure and experimentally manipulate participants' inclination to deviate from clearly advertised norms of cooperation.

Participants were kept blind to the fact that their game partners were bots who always conformed to the social norm by making contributions aligned with the advertised group norm. This design feature ensured that no participants witnessed norm violations from other players, and that the advertised group norm was reinforced by the behavior of the other players.

At the last round of the PGG played with their first group, participants assigned to the treatment condition received a private message promising a monetary incentive of 50 Monetary Units (MUs) to violate the group norm by making a particularly low contribution of 10 MUs (when the norm was to contribute 98 MUs out of an endowment of 100 MUs received at the beginning of the round; Figure 3.1). To control for the additional resources and the experience of receiving a private message regarding special treatment, control condition participants were simply offered to “accept a bonus” by clicking a specific button displayed on the screen. Following this experimental treatment round, participants entered into two new groups, playing two rounds of the PGG each with the new players. After this, they played a one-shot Dictator Game (DG) and a one-shot Trust Game, again with new players (Figure 3.1).

We use the following rounds of the PGG game (with new groups), the Dictator Game, and the Trust Game to explore the impact of the incentive to deviate on participants' propensity to violate new norms in different settings. In particular,

---

<sup>2</sup>Participants who took part in Study 1 were not able to sign up for Study 2.

behavior in the Dictator Game is understood as a measure of adherence to fairness norms (Whitt and Wilson, 2007), and the Trust game as a measure of adherence to trust norms (Berg et al., 1995). Finally, participants completed a short survey, in which we measured their perception of the costs of violating social norms, their perception of the self as deviant, and demographics.

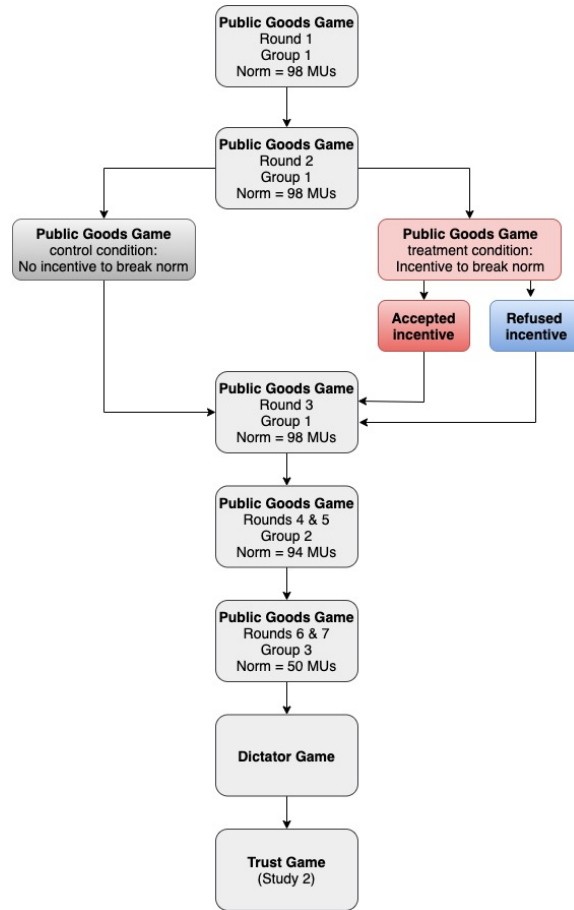


Figure 3.1: Description of the studies. In Studies 1 and 2, participants played the Public Goods game with 3 different groups of players and advertised norms. They played 3 rounds with the first group, 2 rounds with the second group, and 2 rounds with the third group. Before playing the last round with the first group, half of all participants were randomly assigned to an experimental condition that incentivized them to violate the norm of Group 1 at round 3. After playing all seven rounds of the Public Goods game, participants played the Dictator Game with a new player, and in Study 2 only, participants also played the Trust Game with a new player.

### 3.2.2 Pre-treatment setup

**Public Goods game.** At the beginning of each round of the PGG, players received 100 Monetary Units (MUs), and could choose to cooperate or defect by making low vs. high contributions to a common pool. The total amount contributed by all four players was then multiplied by 1.6 and evenly redistributed to all players. The maximum payoff for the group is when all players contribute the maximum amount (i.e., 100 MUs). However, conditional on other players making large contributions, making a low contribution increases the payoff of a single player.<sup>3</sup>

**Mobile and settled players** The objective of these studies was to explore violations of clearly established social norms. For this reason, participants were told that the game involved two different types of players: “mobile” and “settled” players. They learned that mobile players moved from group to group every few rounds, whereas settled players kept playing with the same group the entire time. All participants were assigned to be mobile players, and joined groups composed of settled players (bots) only. This feature of the design made the advertised norm information convincing to the participants.

**Norm advertisement.** To make the norm of each group clear to all participants, we provided them with explicit normative information. The instructions explained that different groups had developed different collective strategies, that is, the typical contribution were different in different groups. When they joined a new group and before making their contribution decisions, participants were explicitly told about: i) the typical contribution in the group, ii) the minimum contribution made in the past N rounds in their group, in which N is a number varying between 65 and 130. For

---

<sup>3</sup>In the case of *repeated* PGG (ie., when players play multiple rounds together), defecting is often not in one’s best material interest: when a player defects, the perceived equilibrium changes, and other players usually start defecting too in subsequent rounds. However, defecting in the last round is always in one’s best material interest, as long as the players know that they will never interact with each other again. In our studies, players are always told whether the round that they are about to play is the last round.

instance, participants were told that G1’s (settled) players typically contributed 98 MUs, G2’s (settled) players typically contributed 94 MUs, and G3 (settled) players typically contributed 50 MUs (Figure 3.1).

**Deception.** Unknowingly to participants, they played all games with fictitious players whose contributions were preset to be consistently aligned with the group norm. This feature of the study allowed us to ensure that participants were exposed to conformity only. To prevent participants from suspecting deception, we made bots’ behaviors look like human behavior. For instance, participants would sometimes have to wait up to 30 seconds until other players (bots) made their contribution decisions in the PGG.

### 3.2.3 Experimental round of the Public Goods game

Before playing their final round with Group 1 (G1), participants were randomly assigned to either the control or treatment condition. Participants assigned to the treatment condition were told that they were selected for “the low-contribution deal”: they would receive a 50 MU bonus if they accepted to make a particularly low contribution of 10 MUs at their last round with G1. The participants who accepted the low-contribution deal (compliers) were then forced to make a 10 MU contribution at this round. To control for the possible effect of being offered a deal with extra money, participants in the control condition were also offered a bonus, but were never asked to make a low contribution with G1.

### 3.2.4 Outcome measures

**Post-treatment rounds of the PGG.** After playing the experimental round with G1, participants played 2 rounds of the PGG with G2, and 2 rounds with G3. We moved participants to a different group right after the treatment round to ensure

that at any post-treatment rounds, participants had the same perception of the equilibrium of their current group independently of their assignment to an experimental condition. In other words, because they moved to a new group, participants' expectations of the contributions of their group's players was independent of the amount that they contributed in the previous round. For this reason, any average difference between treatment and control participants cannot be attributed to a change in one's expectations of other players' contributions.

**Dictator Game.** After playing the Public Goods game with these three groups, participants played a one-shot Dictator Game. In the Dictator Game, all participants were paired with a new player, endowed with 100 Monetary Units, and offered to share some of this endowment with their new game partner. Put differently, participants were all assigned to be the “giver”, and never assigned to be the “receiver”.

**Identification of compliance in the control condition.** Participants assigned to the treatment condition were not forced to accept the incentive to deviate. Those who refused the incentive to deviate (i.e., “non-compliers”) represent 35% of the treatment group from each study. To account for noncompliance, we conducted: i) an intent-to-treat (ITT) analysis, which allows us to identify the causal effects of being assigned to an incentive to deviate (Gerber and Green, 2012), and ii) an instrumental variable (IV) analysis, which allows us to identify the causal effects of deviating after accepting the incentive to deviate (Angrist and Pischke, 2009). To aid the ITT analysis, we directly measured noncompliance within the control group as well, with a question placed in an additional round of public goods games for control group participants that incentivizes to deviate from the group norm in a fashion identical to the treatment group.

In Study 1 (but not Study 2), after playing the DG participants from the control group were led to believe that they would play the PGG again with a new group (G4), which had the same advertised norm as G1 (98 MU average contribution).

Immediately after they joined G4, participants from the control condition were offered the same incentive to deviate of 50 MUs as treatment participants received with G1. Right after making their decision to accept or refuse this incentive to deviate, participants were told that the game was canceled because a player from G4 had left the study, and were directed to the survey questions. This procedure allowed us to successfully identify the control participant who would have complied to the treatment, and directly compare their post-treatment decisions to treatment condition compliers (Figure 3.3).

**Trust Game.** In Study 2 (but not in Study 1), participants also played a one-shot version of the Trust Game with two successive players. In the Trust Game, participants were paired with another player (bot) and assigned to the role of the “receiver”. In this context, participants always started a game with 0 MUs. The other player, commonly referred to as the sender, started with an endowment of 100 MUs. The first mover is always the sender, who decides how much of their endowment to send to the receiver. This amount is tripled and sent to the receiver, who then decides how much of this amount to send back to the sender. In this game, if the sender believes that the receiver is not trustworthy, the maximum payoff is to not send any MUs to the receiver. Like in the Public Goods game, participants were informed about the norm. A message indicating the typical behavior of senders and receivers (i.e., the norm) was displayed on the screen: participants were told that senders typically send between 65 and 75 MUs, and receivers typically send between 85 and 95 MUs back.

**Perceptions of the costs of violating norms: Hypothetical Trust game deal.** To compare the price of deviance for treatment and control participants, we asked: how large a monetary incentive would have to be for participants to violate a norm. At the end of Study 2, we asked participants: “if we were to offer you a

low-reciprocation deal, what is the minimum number of MUs that you would need to accept the deal?”

**Identity questions.** Participants were also asked to answer a total of 9 questions about their perception of the self as a norm violator. These questions were on a scale from 1 to 7 (1 = strongly disagree, 7 = strongly agree) and included: “I see myself as someone who stands out of the crowd”, “I see myself as a bit of a deviant”, or “I see myself as someone who doesn’t like to follow social conventions”.

### 3.2.5 Analytic strategy

For each outcome, we use two analytic strategies, which we preregistered on the Open Science Framework (OSF)<sup>4</sup>. Statistical significance is robust to both analytic strategies. First, we report intent-to-treat (ITT) effects, estimated using linear regression (OLS). This analysis yields an estimate of the effect of being incentivized to violate a social norm on the outcome of interest, which are more conservative than the results of the Instrumental Variable analysis.

The main model can be written:

$$y_i = \beta_0 + \tau Z_i + \epsilon_i \quad (3.1)$$

in which  $y_i$  is the outcome for the  $i$ ’th participant,  $\tau$  is the intent-to-treat effect,  $Z_i$  is a treatment assignment indicator for the  $i$ ’th participant, and  $\epsilon_i$  is an error term. In Study 1, we include gender ( $X_i$ ) into the model to correct for covariate imbalance between the treatment and control conditions (Angrist and Pischke, 2009). The conclusions of these analyses are the same regardless of the presence of covariates in the model.

---

<sup>4</sup>Preregistration files can be found on the following link: <https://osf.io/yextr/> We also posted the data and code onto this repository.

All the findings reported in the Results section are robust to the following model:

$$y_i = \beta_0 + \tau Z_i + X_i\beta + \epsilon_i \quad (3.2)$$

in which  $y_i$  is the outcome for the  $i$ 'th participant,  $\tau$  is the intent-to-treat effect,  $Z_i$  is a treatment assignment indicator for the  $i$ 'th participant,  $X_i$  is a vector of preregistered covariates,  $\beta$  is a vector of covariate effects, and  $\epsilon_i$  is an error term.

In addition, because a share of the treatment condition participants refused to comply to the treatment, we also report estimates of the treatment on treated effect (TOT) using instrumental variable (IV) regression analysis. The results of the IV analysis is illustrated in Figure 3.3 and Figure 3.4b, which compare i) the treatment condition compliers (incentivized to deviate with G1c) to the control condition participants who would have complied to the treatment (incentivized to deviate with G4), ii) the treatment condition non-compliers to the control condition participants who would not have complied to the treatment.

### Study pre-registration

This study was pre-registered on the Open Science Framework: <https://osf.io/yextr/>

## 3.3 Results

### 3.3.1 Description of the sample

A total of 3,499 U.S. residents from the online participant pool Prolific participated in Study 1 or Study 2. Control and treatment participants across studies were comparable on all demographic variables except for gender. Specifically, the treatment condition included fewer women than control (Table 3.1). This imbalance is entirely driven by a randomly-generated gender imbalance between conditions in Study 1

(Appendix B: Table B.1 and Table B.2). As preregistered, we control for covariates that are not equally represented across conditions in the regression analyses; inclusion and exclusion of this covariate in Study 1 analyses does not impact the statistical or substantive significance of the experimental treatment.

Table 3.1: Description of the sample

|                              | Range | All participants | Control | Treatment | Difference | P-value |
|------------------------------|-------|------------------|---------|-----------|------------|---------|
| Gender                       | 0-1   | 0.47             | 0.48    | 0.45      | -0.03      | 0.05    |
| Age                          | 18-81 | 34.00            | 33.61   | 34.08     | 0.47       | 0.25    |
| Income                       | 1-5   | 2.59             | 2.57    | 2.60      | 0.03       | 0.46    |
| White                        | 0-1   | 0.74             | 0.74    | 0.74      | 0.00       | 0.81    |
| Black                        | 0-1   | 0.07             | 0.07    | 0.08      | 0.01       | 0.15    |
| Asian                        | 0-1   | 0.08             | 0.08    | 0.08      | 0.00       | 0.99    |
| Hispanic                     | 0-1   | 0.06             | 0.06    | 0.07      | 0.00       | 0.55    |
| Experience wt games          | 1-4   | 1.55             | 1.54    | 1.56      | 0.02       | 0.48    |
| Pre-treatment contribution 1 | 0-100 | 88.90            | 89.31   | 88.49     | -0.82      | 0.3     |
| Pre-treatment contribution 2 | 0-100 | 88.26            | 88.86   | 87.65     | -1.21      | 0.15    |

Note: Table 3.1 displays descriptive statistics for the entire pool of participants that took part in either Study 1 ( $N = 1,439$ ) or Study 2 ( $N = 2,060$ ). The total list includes the average value of preregistered covariates and pre-treatment contributions to the Public Goods game for the entire sample and for each experimental condition. The difference in means between the treatment and control condition are listed along with the p-values from separate t-tests for these differences in means.

### 3.3.2 Pre-treatment contributions

The results reveal that most participants followed the advertised norm of cooperation of their Public Goods game group pre-treatment. In the two rounds that preceded the treatment incentive to deviate with G1, the median contribution to the common pool across the two studies was 97 MUs, with an average contribution of 89 MUs. Further explorations of levels of cooperation in the first two rounds of the Public Goods game revealed that 75% of the participants contributed more than 90 MUs. As displayed in Table 3.1, before the treatment incentive to deviate, participants from the treatment

and control conditions made comparable contributions to the common pool in both studies.

### **3.3.3 Noncompliance to treatment**

In the final round with G1, treatment participants received their incentive to deviate from the highly cooperative norm. In both studies, 35% of the participants who received the incentive to deviate decided to not comply: on average, these non-compliers donated 95 MUs in this round, compared to the average donation of 10 MUs among those who complied to the treatment (Figure 3.3).

### **3.3.4 Effect of deviance incentive on violations of subsequent norms in different settings**

Figure 3.2 and Figure 3.3 show the average contribution of participants from each condition in the subsequent rounds of their new Public Goods games (see also Appendix B). The intent-to-treat analysis reveals that the incentive to deviate causally increased participants' propensity to violate social norms in subsequent Public Goods games featuring different groups, group norms, and players. Compared to participants assigned to the control condition, participants who received the incentive to deviate with Group 1 violated the norms of both Group 2 and Group 3 at higher rates. Specifically, participants from the treatment condition contributed 78.60 MUs on average, which is significantly less than the average contribution of 82.40 MUs made by the participants from the control condition. This difference of 3.80 MUs represents a 4.5% decrease in contributions and is statistically significant ( $SE = 1.52$ ,  $p < .01$ ; Figure 3.2). Similarly, participants from the treatment condition contributed less than the participants from the control condition with their third group. Although all participants were informed that the norm with Group 3 was to contribute 50

MUs, participants from the treatment condition contributed 46.81 MUs on average, whereas participants from the control condition contributed 51.15 MUs on average. This statistically significant difference of 4.34 MUs ( $SE = 1.35$ ,  $p = 0.001$ ; Figure 3.2) represents a 8.5% decrease in contributions due to random assignment to the incentive to deviate.

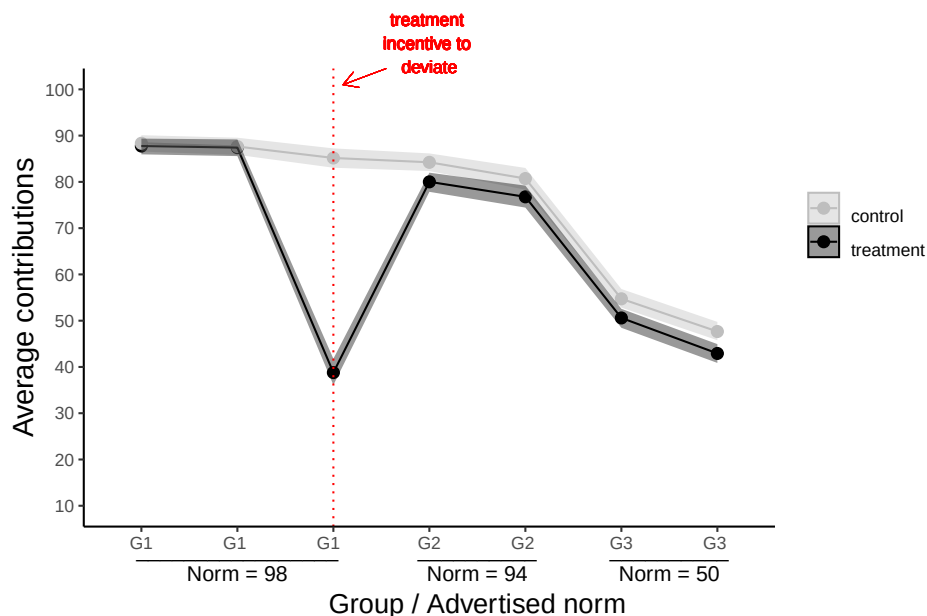


Figure 3.2: Average contributions in the Public Goods game. This figure illustrates the difference in average contributions of treatment and control participants in the Public Goods game at each round. Pre-treatment, participants randomly assigned to treatment and control condition made comparable contributions to the common pool. Post-treatment, however, participants who received the incentive to deviate violated norms at higher rates. They made significantly lower contributions to the common pool than participants in the control condition at every post-treatment round. Each of these differences in means represents the intent-to-treat effect, that is, the causal effect of being randomly assigned to receiving the incentive to deviate.

For the intent-to-treat analysis, we identify treatment compliers and control compliers, who are the participants who were randomly assigned to receive an incentive to deviate (treatment, at the end of Game 1, and control, in a final additional Public Goods game) and treatment and control non-compliers, who refused the incentive to deviate when it was offered at those times. Figure 3.3 shows that control and

treatment non-compliers (76% and 66% of their groups, respectively) are similar in their relatively high and norm-conforming contributions to the Public Goods game pre-treatment, whereas control and treatment non-compliers were also comparable prior to treatment in their significantly lower and less norm-conforming contributions prior to treatment. This figure shows the broad comparability of our two groups of compliers and non-compliers in treatment and control; it also illustrates how individual differences play a role in responses to normative information. The ITT analysis using these groups, based on random assignment to treatment and control, does not require the comparability of compliers in each condition in order to show the causal effect of the invitation to deviate on the treatment participants who accepted it in the third round of Group 1.

Consistent with the ITT analysis, the IV analysis also reveals a significant effect of the invitation to deviate, among treatment participants who accepted it (the TOT): the effect represents a 7 to 12% average decrease in contributions to the common pool after participants had deviated one time (Figure 3.3). Specifically, the analyses reveal that complying to the treatment decreased participants' contributions to the common pool by 5.68 MUs with Group 2 ( $SE = 2.27$ ,  $p = 0.01$ ) and 6.56 MUs with Group 3 ( $SE = 2.00$ ,  $p = 0.001$ ), on average<sup>5</sup>. The impact of the incentive to deviate on future deviations was not driven by a group of participants who learned to give the same low amount each time (10 MUs), nor by a group of outliers who simply adjusted their playing strategy to include low amounts each time (close to 10 MUs; see Appendix B.4).

---

<sup>5</sup>These estimates were derived from a pre-registered instrumental variable analysis strategy, described in the method section. The results of the instrumental variable analysis are equivalent to direct comparisons of the contributions of treatment and control compliers.

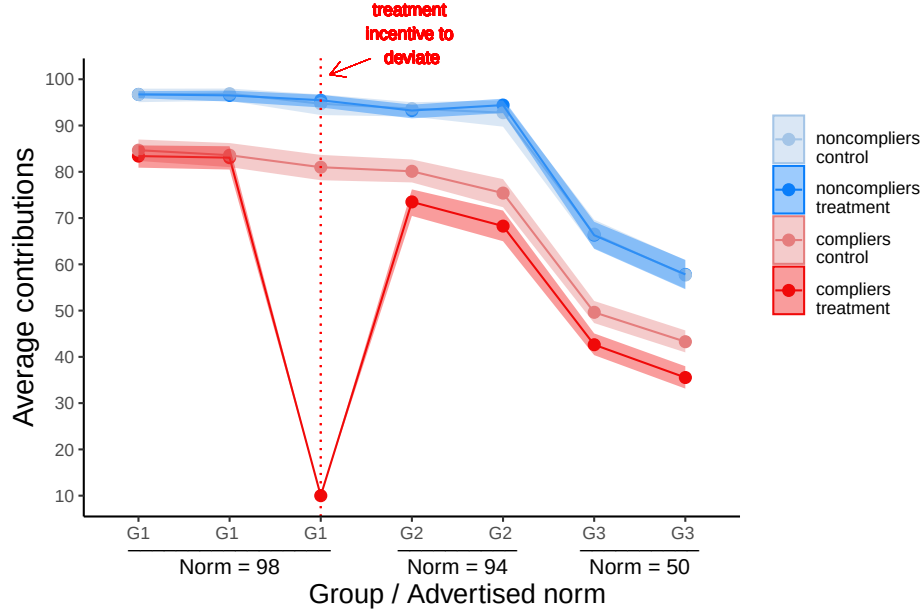


Figure 3.3: Illustration of the treatment on the treated effect on Public Goods game contributions. This figure illustrates the treatment on the treated effect (TOT) for each post-treatment round, that is, the causal effect of taking up the treatment on one's contributions to the common pool in the Public Goods game. As displayed by the blue lines, we found no difference between the contributions of treatment non-compliers and control non-compliers (those who refused an incentive to deviate in a later round), pre- and post-treatment. As depicted by the red shades, we found a statistically significant difference between treatment and control compliers (those who accepted an incentive to deviate in a later round) at each post-treatment round.

### 3.3.5 Effect of treatment incentive to deviate on future violations of norms of fairness and reciprocation

The results from participants' play in subsequent behavioral games reveal that the treatment incentive to deviate also increased participants' violation of different social norms of fairness and reciprocity in entirely new settings with different rules and players. The extension of their deviant behavior to these new settings also argues against a simple learning explanation for our previous findings, that the players violated norms because they were taught by experimenters how to earn the most money in one particular (Public Goods) game.

**Dictator Game.** As displayed in Figures 3.4a and 3.4b, both the intent-to-treat and the instrumental variable analyses reveal that incentivizing and successfully inducing participants to deviate from a norm of cooperation in the Public Goods game increased participants' rates of violation of an implicit (i.e., not advertised) norm of fairness in the Dictator Game. Participants from the treatment condition sent an average amount of 37.81 MUs to their partner, which is significantly lower than the average 41.18 MUs that control condition participants sent to their partner (averaging across compliers and non-compliers). This difference of 3.37 MUs is statistically significant ( $SE = 1.36$ ,  $p = .01$ ) and represents an 8% decrease in real money donated due to the experimental incentive to deviate.

**Trust Game.** As displayed in Figure 3.4c, participants who received the incentive to deviate during the Public Goods game also violated a norm of reciprocity (which was prominently advertised) at higher rates than the participants from the control condition who never received the incentive to deviate. Specifically, whereas participants assigned to the control condition sent 77.05 MUs back to the sender, participants from the treatment condition sent 74.32 MUs back (averaging across compliers and non-compliers). This effect of 2.73 MUs represents a 3.5% decrease in reciprocity and is marginally significant in the conservative ITT analysis ( $SE = 1.59$ ,  $p = 0.09$ ).

**Perceptions of the costs of violating norms.** At the end of the study, both treatment and control participants were asked to imagine that they were offered a monetary incentive (in MUs) to violate a strong norm of reciprocity in a hypothetical Trust Game. Participants were asked for the price of their deviance: how large this monetary incentive would have to be for them to violate the norm.

On average, participants who received the treatment incentive to deviate during the Public Goods game estimated that they would need a lower monetary incentive to violate a norm of reciprocity in the Trust Game than those who never received an incentive to deviate. This suggests that inducing participants to violate a norm

depreciated their expectation of the cost of violating a social norm. Specifically, participants from the treatment condition estimated that they would need a 258 MU bonus and control participants estimated 288, a 30 MU difference, in order to violate a social norm in the context of this hypothetical Trust Game. The difference between control and treatment estimates of the cost of deviating from a norm was statistically significant ( $SE = 15.00$ ,  $p = .04$ ; Figure 3.4d).

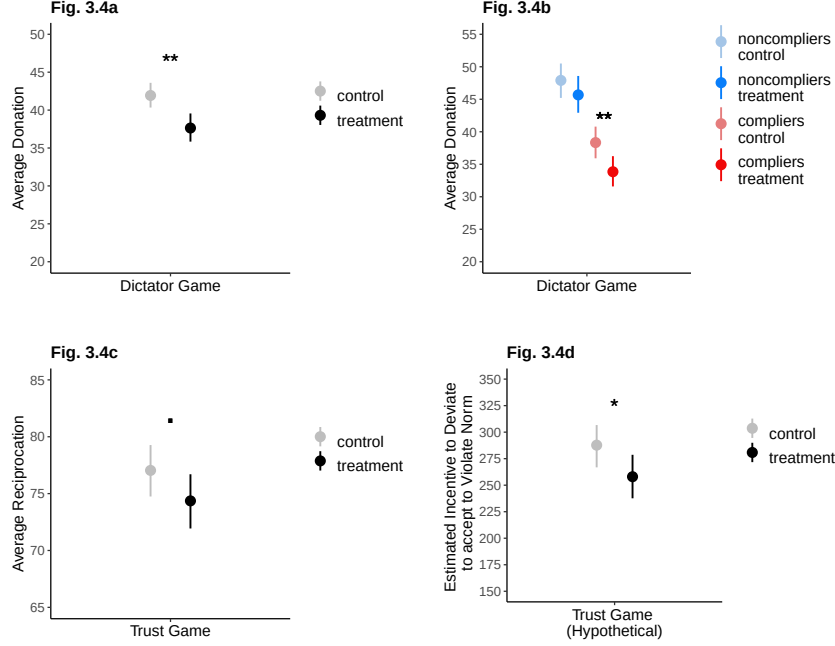


Figure 3.4: Causal effect of the incentive to deviate in different games. **Fig. 3.4a** illustrates the difference in average donation between treatment and control participants in the Dictator Game. This difference in means represents the intent-to-treat effect—the causal effect of being assigned to the treatment incentive to deviate (in a previous Public Goods game) on deviation from the Dictator game’s implicit (non-advertised) norm of fairness. **Fig. 3.4b** illustrates the treatment on treated effect (TOT), that is, the causal effect of complying to the treatment incentive to deviate in the previous Public Goods game on behavior in the Dictator game. Here we found no difference between the donations of treatment and control non-compliers (in blue) and a statistically significant difference between the donations of treatment and control compliers (in red). **Fig. 3.4c** illustrates the difference between treatment and control participants’ reciprocation levels in the Trust Game (Study 2). This figure suggests that the previous treatment incentive to deviate made participants more likely to violate the Trust game’s explicitly-advertised norm of reciprocation. The difference between the two experimental conditions is marginally significant. Finally, **Fig. 3.4d** illustrates the difference in treatment and control participants’ perceptions of the costs of violating social norms. Specifically, this figure displays participants’ average estimate of the amount that they would require as an incentive to violate a norm of reciprocation in a hypothetical Trust Game.

**Self-perception as a deviant.** The results suggest no effect of the treatment incentive to deviate on participants’ perception of the self, among all treatment participants and among compliers to the treatment.

### 3.4 General discussion

The results of these experiments ( $N = 3,499$ ) reveal that inducing people to violate a social norm increased their propensity to violate the norm again, and to violate different social norms in new settings with new players and groups. Traditional theories of deviance have considered the tendency to conform or deviate as a personality trait or individual difference (e.g., Snyder, 1979; Becker, 1963; Gino, 2018; Bennett and Robinson, 2000; DeBono et al., 2011). Consistent with this literature, our evidence shows that a minority of individuals cannot be induced to deviate, even under favorable conditions of “no punishment” from group members and real monetary reward. However, we also find that the majority of individuals can be induced to deviate under these circumstances, and that this experience results in a behavioral pattern of deviation over a period of time, in response to new types of norms. This finding that social experience causally impacts an individual’s tendency to deviate represents a new finding in the norms and deviance literature.

Why do individuals who deviated from one norm do so again? The experimental results suggest one psychological mechanism explaining the emergence of a pattern of deviance. Violating a social norm caused individuals to depreciate their perception of the costs of deviation. When asked how large of a monetary incentive they would need to violate a norm of reciprocity, participants who previously had been incentivized to deviate required a significantly lower amount of incentive than those who were not. We did not find support for a psychological mechanism involving the self –treatment participants did not consider themselves to be “deviants” to a greater extent than control participants, even after they deviated on average more across several different behavioral games.

Particular features of the behavioral games used in these experiments may explain why we found that individuals depreciated the perceived costs of deviance but did not change their self-perception. In these games, we did not allow other players

to punish deviants, either with social or monetary sanctioning through messages or fines. This design stands in contrast to many behavioral game paradigms aimed at testing the evolution of norms, which feature punishment (Fehr and Schurtenberger, 2018; Fehr and Fischbacher, 2004; Fehr and Gächter, 2000). In our context, it is perhaps unsurprising that treatment individuals depreciated the anticipated costs of deviation. They found deviation to be a rewarding monetary experience, and possibly also found that it wasn't as difficult a psychological experience as they may have originally anticipated. Note, however, that no treatment participant offered to deviate for free –participants who anticipated a monetary payoff for deviating in the face of no punishment still required on average 258 MUs to deviate from a group norm. The absence of punishment in this game may limit the generality of our findings, but this feature usefully shows how psychologically compelling norms can be, even when violation of the norms results in material gain and zero social sanctioning.

We find that individual differences matter for deviation, but also that experience with a single act of deviation can cause a behavioral pattern that is suggestive of an individual difference. Both the self-monitoring scale (Snyder, 1979) measured at the end of Study 1 and the deviant identity index measured at the end of Study 2 were strong and statistically significant predictors of participants' decision to accept the incentive to violate the group social norm. Specifically, being a low self-monitor or perceiving one's identity as deviant was associated with increased levels of norm violations in the Public Goods game ( $\beta_{sm} = 3.45, SE = 0.99, p < .001$ ,  $\beta_{id} = -3.87, SE = 0.64, p < .001$ ). Taken together, the correlation between individual differences and acts of deviance on the one hand, and the causal impact of participants' randomly-incentivized experiences with deviance on future acts of deviance on the other hand are consistent with accounts of the malleability of personality traits across the life span (Damian et al., 2019). The present research indicates that

individuals may be more or less inclined to conform or deviate, but that experiences can causally shape or re-shape these personal inclinations.

This experimental paradigm offers many possible permutations that can test the boundaries of these effects in future research. As already discussed, future research can include punishment in the game, as a way to investigate whether self-perceptions may change more and perceptions of costs may increase when players punish deviant teammates. Additionally, deviance could be randomly invited through non-monetary incentives. The type of deviance that we model may parallel real world deviant acts like corruption that are fueled by monetary rewards. Other acts of deviance are based on principle rather than material gain, such as extreme acts of activism on behalf of an oppressed group. Future research could attempt to examine the emergence of patterns of deviance borne of deviance motivated by principle as opposed to material interest.

## 4

# Public Acts of Deviance in Settled Cultures: A Field Experiment in the Hasidic Community

How does the experience of publicly violating norms that are meaningful to our friends, relatives, and neighbors, shape us? Scant research has examined how individuals' public acts of deviance within their social groups impact their private views, perception of their community, and future behaviors. The present study focuses on extreme cases of deviance in the context of settled cultures, in which norms are strong and widely followed, and acts of deviance are rare (Swidler, 1986). Specific questions include: Does the experience of violating a norm increase individuals' willingness to support deviants? Does it make them more inclined to deviating from norms again? After violating a norm, do individuals become more critical of their community?

These questions are important to address because we might not expect a single act of deviance to create social change—it is more likely that sustained acts of deviance over time, potentially growing in size or spreading to others, help to shift community norms (Centola et al., 2018; Moscovici, 1976, 1994)

In the present research, we deployed a randomized field experiment in the Hasidic community in New York, which is a subgroup within the ultra-Orthodox Jewish community. The Hasidic community is distinct for its homogeneity and strong and injunctive social norms that stand alongside its strict religious codes and regulations. A core principle guiding the Hasidic Jews is the idea of *preservation* of the Jewish civilization, which implies that any form of individual or collective change is unwelcome. To achieve this ideal of preservation, the Hasidic follow strict religious and social (i.e., non-religious) rules. For instance, community members tend to not engage with any outside groups, including non-orthodox or secular Jews. They also follow a strict religious dress code that was established centuries ago. For example, men have to wear *Payies*, which are long and curly side burns. In the Torah portion called Parshat Kdushim, one can read: “You shall not round off the side-growth on your head, or destroy the side-growth of your beard” (Leviticus: 19, 27).

Non-religious social norms about how to dress on non-religious occasions are also clear, strong, and prescriptive, although they vary depending on the location of the community. For example, among Hasidic Jews in Brooklyn, U.S., women wear fancy clothing, preferably designer brands, at great personal cost and even when their budgets can barely accommodate the expense. Both men and women in these communities look down upon women who do not dress *balebatish*, a Yiddish term commonly used in the Brooklyn Hasidic community to describe high-end, fancy and elegant clothes and accessories.

For these reasons, the Hasidic community constitutes what sociologists call a “settled” culture, defined as a culture that constrains individuals into a limited repertoire of practices, such that “alternative ways of actions seem unimaginable or at least implausible” (Swidler, 1986, p. 284). This fact about the Hasidic community makes a study of deviance in that context highly unique. To date, most studies in settled cultural periods are only able to examine deviance and cultural change post-hoc, that

is, after cultural change already happened. For instance, political scientist Kristen Monroe, who studied the heroes who violated extremely strong norms to rescue genocide and war victims, conducted her studies in the years following atrocities, once the norms and culture had already changed (Monroe, 1996, 2008, 2011). The present study of the consequences of deviating from a social norm in the Hasidic community while the community is resistant to deviance and change forwards our understanding of how individual and social changes occur in settled cultural periods.

## 4.1 The Satmar Hasidic Community of Brooklyn

### 4.1.1 Study context: A settled culture

The Satmar Hasidic community in New York is a conservative division of the broader ultra-Orthodox Jewish community. Originally from Hungary, the Satmar Hasidic Jews arrived in New York after the Holocaust and settled down in two main neighborhoods of Brooklyn: Williamsburg and Borough Park. Community members follow the principles of Hasidism established by Rabbi Joel Teitelbaum, which emphasizes devotion in the performance of public and private rituals, adherence to tradition, charity, and the rejection of modernity. To protect their conservative lifestyle, the Satmar Hasidic community has distanced itself not only from non-Jews and secular Jews, but also from Orthodox Jews, who adhere to less strict religious ideology (Deutsch, 2009; Poll, 1969). For instance, contrary to other ultra-Orthodox Jewish subcommunities, the Satmar Hasidic Jews remain strongly opposed to the establishment of the secular state of Israel.

Most Hasidic Jews' behaviors and activities are guided by both religious laws (*Halach*) and non-religious social norms. In the context of this study, we draw a clear line between behaviors governed by religious laws (*Halach*) vs. non-religious social norms, defined as community members' impression of the average or desirable

views, attitudes and behavior of the group. To illustrate this distinction, let’s consider clothing rules for Satmar Hasidic women, who constitute the population of interest of the present research. Religious texts (*Halach*) requires women to dress “modestly”, that is, women have to cover their knees, elbows, and if they are married, their hair. This leads them to dress with long-sleeve tops, wear knee-length skirts over opaque leggings, and a *sheitel* (i.e., wig) or *Tichel* (i.e., scarf). Over time, a non-religious social norm developed in the community: Satmar Hasidic women wear fancy clothing, particularly designer brands, often at great personal cost and sacrifice (given that most Satmar are not wealthy). Both men and women in these communities look down upon women who do not dress *balebatish*, or fancy.

We conducted a field experiment to study Satmar Hasidic women’s personal reactions to being invited to violate this high-end clothing norm. Specifically, we recruited Satmar Hasidic women and offered them to participate in our study called *Einer Tug Programa*, which could be translated from Yiddish to English as “Positive Action Day”. A random half of the participants was randomly assigned to be invited to violate the high-end clothing norm for a day whereas the other half was randomly assigned to be invited to carry a Siddur (prayer book) for a day, a neutral activity that they normally engage in.

#### **4.1.2 Ethics of studying the high-end clothing norm in the Hasidic community**

We justify studying the consequences of violating the high-end clothing norm in the Satmar Hasidic community for three main reasons. First, the Satmar Hasidic community is predominantly poor (Roberts, 2011; Yaffe et al., 2018), which makes the social expectation of buying expensive clothing a social issue that puts many households in debt or difficulties. For instance, qualitative interviews of community members revealed that some women “do not fully heat their apartment in the winter so that

*they are able to afford \$700 dresses for themselves and their daughters”* (Yaffe, in prep.) Second, institutions and religious authorities within the community condemn the high-end clothing norm and have previously taken action to attempt to change the norm. For example, a high-school implemented a one-year long intervention aiming to reduce the perceived importance of clothing brands and prices in students and mothers. Third, the social expectation of high-end clothing for women in the community has no religious grounds. This social norm is germane to the Satmar Hasidic Jews in Brooklyn (NY), such that dressing in high-end clothes in Hasidic communities in Europe or Israel would be considered a negative and counter-normative behavior. As a result, violating the high-end clothing norm as part of the study would not carry religious condemnation.

Furthermore, one mode of appeal used by religious leaders in their attempts to change the norm is that ultra-Orthodox Jews generally value poverty and its expressions more than they value wealth (Yaffe et al., 2018), which makes this norm inconsistent with core values of the community. Poverty is viewed as a sign of spirituality, and community members consider poverty a necessary condition to focusing on the study of religious texts (Yaffe, 2020). Thus, our study’s intervention was perceived to be in line with the aims of community leaders and with important aspects of individual community members’ well-being. Naturally, whether community members complied with the intervention’s request to violate the norm for one day was entirely voluntary, and once accepted individuals could vary in their adherence to the study as they were comfortable.

Finally, one of the principal investigators of this project, Dr. Yaffe, belongs to the Hasidic community. At the time of data collection, she lived in one of the neighborhoods in which we deployed the experiment with her family. Living in the community allowed Dr. Yaffe to informally discuss with other community members, such as her neighbors, family, and friends. This led her to identify that the high-end clothing

issue was of major concern in the community. In addition, living in the neighborhood with other community members enabled Dr. Yaffe to monitor individual and community-wide reactions to the study, and debrief with individuals as the study unfolded.

### **4.1.3 Hypotheses**

First, consistent with classic theories of conformity to social norms (Cialdini and Goldstein, 2004; Schwartz, 1973), we hypothesized that participants experienced social and psychological costs for publicly violating the high-end clothing norm. Specifically, we expected participants to feel judged by others for violating the norm or experience negative affect. Second, consistent with theories of the self and results from Chapter 3 of the present dissertation, we hypothesized that violating the high-end clothing norm leads individuals to become more prone to deviance. Specifically, we hypothesized that violating this norm leads participants to be more supportive of deviants, more inclined to publicly deviate from social norms in future, and develop deviant private views of their community.

In addition, we formulated hypotheses about the psychological mechanisms at play. We hypothesized that violating the high-end norm causes participants to depreciate their perception of the costs of violating social norms and update the self-concept by perceiving themselves as a more of a deviant.

Finally, we explored the idea that deviating from the high-end clothing norm could lead individuals to become more sensitive to social issues. We asked, for instance, if participants would be willing to volunteer for ultra-Orthodox Jewish organizations.

## 4.2 Method

### 4.2.1 Experimental design

#### Recruitment

A team of field officers, composed exclusively of Hasidic Jewish women, built a convenience sample using a referral sampling strategy. First, field officers identified highly connected members of the Hasidic community in Brooklyn, who helped build an initial list of 293 names and phone numbers of Hasidic women who lived in Brooklyn at the time of the study. Field officers contacted each of these women on the phone and introduced the study. During their phone calls, all field officers read the exact same script to prospective participants. Specifically, field officers first introduced themselves and explained the terms of the study:

*“The positive action that we ask participants to do are usually quite simple and require no training or particular effort. However, some actions were chosen in collaboration with Rebbetzins and community members to address some pervasive issues in the community. In this sense, you may be assigned to do something that will make you feel some discomfort, like social discomfort. Remember that you can decide to withdraw at any time.”*

Field officers subsequently informed prospective participants that they would not receive any compensation for completing the activity, but that they would receive \$20 if they agreed to complete a questionnaire within 24 hours of the activity.

After providing these general details about the study (without specifying which activity they would participate in) and responding to prospective participants’ questions, field officers obtained consent from participants who agreed to take part in the study by asking them the following two questions: 1) “Do you understand all that I just said to you?”, and 2) “Given all that I just said, would you like to participate?”

Women who agreed to participate were then randomly assigned to and informed about their experimental activity. Prospective participants who decided not to participate in the study after learning about the details of their activity were thanked for their time and encouraged to take the survey regardless.

## **Experimental conditions**

### **Treatment condition: the “plain dress”**

The objective of the treatment was to experimentally induce an experience of deviance in a random half of the participants. We asked participants assigned to the treatment condition to wear an experimental dress, the “plain dress”, which we provided them in their size. This dress was much less fancy than dresses that the average Hasidic woman typically wears in Brooklyn. As a result, by wearing the plain dress in public, participants assigned to the treatment condition experienced publicly deviating from the high-end clothing norm.

The exact dress that participants would wear for their activity was not imposed on them. Instead, participants were offered to choose one of three possible dresses. We piloted these dresses to ensure that most participants would select the same dress (they all did), that is, the least inelegant one. Piloting these dresses also provided evidence that wearing the plain dress in public would not be so ungraceful or clownish that participants would be punished by others (e.g., laughed at or scorned) for wearing it.

A day before a participant’s activity, a field officer brought her the plain dress in her size, asked her to wear it for a day, and encouraged her to: meet with at least two close friends, meet with at least two extended family members (i.e., family members who are not their parents or siblings), and go to the market or another public space.

### **Control condition: the “Siddur”**

The objective of the control condition was to make participants complete a neutral activity. That is, the control activity was meant to make participants engage in a behavior that they were familiar with and that did not make them feel particularly positive or negative. Participants in the control condition were asked to carry a Siddur (i.e., prayer book) for a day. Religious acts are part of community members everyday life, and most participants already owned a Siddur. As a result, we did not expect the control activity to have any effects on participants' outcomes.

A day before her activity, a field officer visited the participant to give her the Siddur that she would carry for a day.

#### **4.2.2 Data collection**

During their visit to the participants' home to give them the plain dress or Siddur, field officers handed out an envelope to the participants. This envelope contained the paper survey that the participants could complete for a \$20 cash compensation within 24 hours of completing their activity. The questionnaire included questions about the participants' experience with the activity, their perception of the Hasidic community, their perception social norms (including the high-end clothing norm), their willingness to violate social norms in future, as well as their willingness to help innovators or people facing adversities.

The survey involved both open-ended and likert-scale questions. For all survey questions, participants were asked to circle a number between 1 and 7, which they used to indicate low levels of agreement (= 1), medium levels of agreement (= 4), or high levels of agreement (= 7) with various statements.

To discourage participants to open the envelope and respond to questions before completing their activity, the envelope had a wax seal. We assume that participants did not start filling out the questionnaire before they completed their activity, but if some of them did, this would work against statistical significance of the results.

## Measure of the psychological and social costs of the one-day activity

**Description of the one-day activity.** The survey began with a series of 4 open-ended questions asking participants to describe their activity. Specifically, we asked participants to write about what they did, whom they met, where they went, and whether they were able to follow their activity’s guidelines.

**Social and psychological discomfort.** To measure the psychological and social costs of violating the high-end clothing norm in a more systematic way, we asked participants to indicate on a scale from 1 (no discomfort) to 7 (high discomfort) the extent to which they “felt weird”, “felt good” (reversed), or “felt that they were standing out” while completing their activity. Additional items asked participants if they thought that people: “noticed something unusual” or “were judgmental” towards them during the activity day.

## Adherence with the community’s views

**Views on poverty.** In the Hasidic community, poverty is viewed as a necessary condition to being able to dedicate oneself to the study of religious texts. For this reason, expressing positive views towards poverty is expected and desirable (Yaffe et al., 2018). We asked participants to indicate the extent to which they agreed with two statements about poverty: “I think that poverty gives a better perspective on life”, and “I think that the least fancy people have a better perspective on life.”

**Views on authority figures.** Asokanims are authority figures of the Hasidic Jewish community who assume political (non-religious) roles. We asked participants about the extent to which they agree with the following statement: “There is a good reason for every act that the asokanim in the community does.”

**Views on own local community.** Participants were asked to indicate the extent to which they agree with the following statement: “My community in Williams-

burg/Borough Park is the best Jewish community in the world.” This framing was suggested by Hasidic women because the Hasidic Jews in Brooklyn commonly refer to their local community as “the best Jewish community in the world!”

### **Inclination to deviate from social norms**

**Expressing willingness to help innovators.** We told participants that recently, two women entrepreneurs informed us about their projects aiming to make the community happier and healthier. One of the projects was about teaching a class on healthy food whereas the other project was about promoting and encouraging women to go to the gym. For each of these (counter-normative) projects, we asked participants to indicate the extent to which they were willing to help these entrepreneurs by talking with them on the phone.

**Expressing willingness to advocate against the high-end clothing norm.** Participants were asked to report the extent to which they would like to sensitize community members to the problems that come with wearing expensive clothes.

**Expressing willingness to publicly violate social norms again.** We measure participants’ willingness to violate social norms in future in the context of the following five increasingly counter-normative behaviors in the Hasidic community: using a kosher smartphone with access to the filtered internet to purchase clothes, having a kosher smartphone with an access to the filtered internet, dressing casually, driving, and volunteering for non-Hasidic organizations that help people outside of the community.

### **Psychological mechanisms leading to shifts in individual’s inclination to deviate**

**Perception of the self as deviant.** Two items adapted from the self-monitoring scale (Snyder, 1979) aimed to estimate the extent to which participants perceived

themselves as a deviant: “I see myself as someone who is strong enough to stand out of the crowd” and “I see myself as someone who will always try to fit with the crowd.”

**Perception of the costs of violating social norms.** Two items aimed to estimate participants perception of the costs associated with violating social norms: “I would never do something different from my friends because it feels too bad.” and “It’s easy for me to do something different from my community in order to improve it.”

**Perception of high-end clothing.** We asked participants to indicate the extent to which they agreed with four different statements about high-end clothing. Two of these statements aimed to assess the participants’ personal beliefs about the overpriced clothing norm, whereas the other two statements aimed to assess participants’ perception of the strength of the overpriced clothing norm. These statements read: “I believe [in general, women think] that buying and wearing expensive clothes is important”, and “If I [in general, women think that if they] don’t wear expensive clothes, my [their] family and friends will not like me [them] as much.” In addition, we asked participants whether they would like to purchase and wear less expensive clothes in the future.

**Sensitivity to social issues in the community.** We measured participants’ sensitivity to the community’s social issues in two different ways. First, we asked participants to report the extent to which they would like to volunteer for ultra-Orthodox *chesed* (i.e., charity) organization. Second, we asked participants about whether they hope that change will occur in the community in two different survey items. Specifically, we measured the extent to which they agreed or disagreed with the following statement: “When I think of myself and other women, I hope that some things will change.” Then, participants were offered to tell us more about what they wish will change in an open-ended question.

### 4.2.3 Analytic strategy

#### Noncompliance

During the consent process, that is, before learning about their assignment to an experimental condition, participants were informed that they could withdraw from the study at any time. Most of the participants who withdrew from the study after discovering their assignment had been assigned to the plain dress condition, which highlights the counter-normative and thus challenging nature of this experimental activity.

An overall 3% of participants assigned to the Siddur (control) condition refused to complete their one-day activity after learning about their assignment. All of these participants filled out an endline survey. Among the participants assigned to the plain dress (treatment) condition, 28% refused to complete the one-day activity, and we collected endline survey responses from 82% of them. Importantly, we collected demographic data (i.e., marriage status, number of siblings, and age) for all 155 participants who took part in this study.

We use all of the survey respondents in the analyses, whether or not they complied with the treatment. This intent-to-treat (ITT) analysis provides an estimate of the causal effect of *being assigned to* deviate from the high-end clothing norm and constitutes a conservative test of the effect, since not all of the individuals in the analysis chose to experience the randomly assigned treatment. In addition, to account for noncompliance in our analyses, in Appendix C we use a technique called instrumental variables regression analysis (IV analysis; Angrist and Pischke, 2009), which derives the treatment on treated effect (TOT), that is, the causal effect of violating the high-end clothing norm among those who experienced the randomly-assigned experimental condition. These two analytic strategies yield comparable statistical significance re-

sults; however, because the IV analysis is a more liberal analysis it typically yields larger effect sizes (Appendix C).

To account for missing values from the share of noncompliers who did not take the survey, we conduct additional intent-to-treat analyses, which we report in the “weighted” column of Tables 4.1 to 4.3. These additional regression analyses employ a method called Inverse Probability Weighting (IPW; Gerber and Green, 2012; Gomila and Clark, 2019). The objective of IPW is to render the treatment and control groups relatively more comparable given that some individuals selected out of the survey, by upweighting the responses of respondents who are most similar to non-respondents. IPW makes the assumption that missing values are missing completely at random conditional on observed covariates (Gerber and Green, 2012; Gomila and Clark, 2019). Given the strong similarity in the life courses of women in the Hasidic Jewish community, it seems realistic to assume that their demographic covariates that we collected (i.e., marriage status, number of siblings, and age) are sufficient to explain patterns of missingness in the data.

We present estimates of the causal effects of deviating from the norm from both the unweighted ( $\tau_u$ , without IPW) and weighted ( $\tau_{ipw}$ , with IPW) analyses in the main text and tables—primarily to contrast the two analytic strategies because we did not anticipate this type of missingness in our preregistration and thus did not register the use of IPW. In large part, the weighted and unweighted analysis result in the same conclusions (consistent also with the IV analyses in Appendix C), although the weighted (IPW) analyses yield more statistically significant findings and larger effect sizes. While we present both weighted and unweighted ITT estimates, we primarily focus our results discussion to those analyses that were pre-registered and arguably provide the most conservative estimates: unweighted ITT estimates.

## Model

For each outcome variable, we estimate the intent-to-treat (ITT) effect of being assigned to the treatment condition using the following simple regression analysis:

$$y_i = \beta_0 + \tau Z_i + \epsilon_i \quad (4.1)$$

in which  $y_i$  is the outcome for the  $i$ 'th participant,  $\tau$  is the intent-to-treat effect (i.e., difference in means between the control and treatment conditions),  $Z_i$  is a treatment assignment indicator for the  $i$ 'th participant, and  $\epsilon_i$  is an error term.

## Study pre-registration

This study was pre-registered on the Open Science Framework: <https://osf.io/g5q73/>

## 4.3 Results

### 4.3.1 Description of the sample

A total of 155 women from the Hasidic community participated in this study, 52% of them were assigned to the Siddur (control) condition, and 48% of them were in the plain dress (treatment) condition. At the time of the study participants were 23 years old, had 8 siblings on average, and 60% of them were married.

### 4.3.2 Measure of the psychological and social costs of the one-day activity

Consistent with our hypotheses, participants invited to wear the plain dress (treatment) experienced psychological and social costs, which was not the case of participants assigned to carry a Siddur (control). For each dependent variable reported

below, the difference between treatment and control participants is large, statistically significant, and appears in both open-ended descriptions of their activity and responses to Likert scale items (1 = low, 7 = high; Table 4.1).

**Open-ended description of the one-day activity.** All participants assigned to carry a Siddur (control) confirmed, in the open-ended questions asking them about their day, following their activity’s guidelines. On the contrary, 15% of the participants assigned to wear the plain dress did not entirely follow their activity’s guidelines. Some participants from the treatment condition mentioned that they encountered resistance from family members or avoided wearing the dress in public settings because they anticipated negative reactions from others. These statements underline the counter-normative nature of wearing the plain dress in the Hasidic community in Brooklyn. For instance, a participant wrote: *“I definitely encountered obstacles and my mother made me change outfits.”* Another participant mentioned: *“I mostly stayed at home but I went to take the garbage out and I waited for my daughter outside.”* Although these participants ended up not wearing the plain dress in public for a significant amount of time, their feedback and answers to questions about their experience indicate that they did psychologically experience violating the high-end clothing norm on that day.

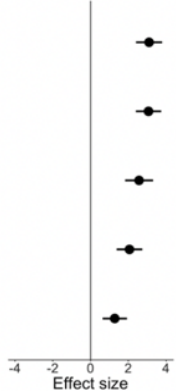
**Social and psychological discomfort.** As hypothesized, the experience of carrying a Siddur for a day was neutral. Participants assigned to the Siddur (control) condition did not experience any discomfort during their activity, with average ratings below 2 on a scale from 1 (no discomfort) to 7 (a lot of discomfort) for all items (Table 4.1). In contrast, participants assigned to the plain dress (treatment) condition experienced both social and psychological discomfort, with average responses between 4 and 5 on the same 1-7 scale (Table 4.1). Specifically, in addition to experiencing negative affect for deviating, participants assigned to wearing the plain dress felt that people noticed something unusual about them and judged them for that. These

differences between control and treatment participants are large and statistically significant, with effect sizes ( $\tau_u$ ) going from 1.28 to 3.10 points (Table 4.1). They indicate that violating the high-end clothing norm comes with significant psychological and social costs in the Hasidic community in Brooklyn.

Table 4.1: Psychological and social discomfort with the activity.

| Item [hypothesized direction of the effect $\tau$ ] | <i>Condition</i> |       | $\tau$            |                   |
|-----------------------------------------------------|------------------|-------|-------------------|-------------------|
|                                                     | Siddur           | Dress | unweighted        | weighted          |
| <b>Social and psychological costs [ + ]</b>         |                  |       |                   |                   |
| I felt like I stood out [ + ]                       | 1.62             | 4.72  | 3.10***<br>(0.35) | 1.50***<br>(0.30) |
| I felt weird [ + ]                                  | 1.80             | 4.87  | 3.07***<br>(0.34) | 0.54<br>(0.64)    |
| People noticed something unusual [ + ]              | 1.52             | 4.09  | 2.57***<br>(0.37) | 3.11***<br>(0.73) |
| People were a little judgmental [ + ]               | 1.51             | 3.56  | 2.06***<br>(0.34) | 1.45***<br>(0.21) |
| I felt good (reversed) [ + ]                        | 2.99             | 4.26  | 1.28***<br>(0.33) | -0.29<br>(0.40)   |

<sup>†</sup>p<0.1; \*p<0.05; \*\*p<0.01; \*\*\*p<0.001



Note: Each row of Table 4.1 includes: the average response for the participants in the Siddur and Dress conditions, the effect size ( $\tau$ ) of the treatment derived from unweighted ( $\tau_u$ ) and weighted ( $\tau_{ipw}$ ) regression analyses, and the standard error of the effect size of the treatment. Statistical significance is indicated using stars. The figure appended to the right side of the table provides a visualization of each unweighted effect sizes.

### 4.3.3 Adherence with the community's views

The results suggest that deviating from the high-end clothing norm led participants to adhere less with the mainstream views of the Hasidic community and become more critical of some aspects of the community.

**Views on poverty.** Participants assigned to violate the high-end clothing norm agreed significantly less with traditional views on poverty than participants assigned to carry a Siddur (Table 4.2). As displayed in Table 4.2, participants' beliefs that

poverty improves individuals' perspective on life was less strong in the plain dress condition ( $M_T = 3.55, SD = 1.63$ ) than in the Siddur condition ( $M_C = 4.05, SD = 1.49$ ;  $\tau_u = -0.50, SE = .29, p < .1$ ). Similarly, participants' belief that the least fancy generally have a better perspective on life was less strong in the treatment ( $M_T = 4.07, SD = 1.81$ ) than in the control condition ( $M_C = 4.76, SD = 1.70$ ;  $\tau_u = -0.70, SE = .33, p < .05$ ). These differences between treatment and control condition represent a 12 to 14% difference in one's views on poverty.

**Views on authority figures.** Participants assigned to deviate from the high-end clothing norm believed that the acts of the Asokonim (local authority) are generally good to a lesser extent ( $M_T = 5.57, SD = 1.41$ ) than the participants assigned to carry a Siddur ( $M_C = 6.03, SD = 1.16$ ). The effect size of the unweighted intent-to-treat analysis was 0.46 points in absolute value ( $SE = .23, p < .05$ ; Table 4.2) on a scale from 1 to 7, corresponding to an estimated 7.6% change due to treatment assignment only.

**Views on own local community.** Participants in the plain dress condition ( $M_T = 4.71, SD = 1.81$ ) agreed less with the statement that their community is the best Jewish community in the world than participants in the Siddur condition ( $M_C = 4.92, SD = 1.73$ ). Although this effect of 0.22 points on a 1 to 7 scale does not reach statistical significance in the unweighted intent-to-treat analyses ( $SE = .31, p = .49$ ), we note that estimates from weighted analyses (IPW) yield larger and statistically significant effect size of 0.79 points in absolute value ( $SE = .31, p < .01$ ).

#### 4.3.4 Inclination to deviate from social norms

Inconsistent with our hypotheses, the primary –unweighted– analyses suggest that deviating from the high-end clothing norm did not lead participants to be more inclined to publicly violate norms in future. If anything, results from weighted regression

(IPW) suggest the opposite. However, deviating from the high-end clothing norm seems to have caused tolerance or support for deviance, and increased one’s inclination to deviate in private.

**Expressing willingness to help innovators.** Although the positive effect of deviating on participants’ willingness to help innovators does not reach statistical significance in the unweighted intent-to-treat analyses (Table 4.2), weighted intent-to-treat analyses provide some evidence of a positive effect of the treatment on participants’ willingness to help innovators by discussing their project with them on the phone (e.g.,  $\tau_{ipw} = 1.59, SE = .41, p < .001$ ; Table 4.2). This estimated effect of deviance is large, corresponding to a 40% increase in expressed willingness to privately support deviants in the community.

**Expressing willingness to advocate against the high-end clothing norm.** Contrary to our hypotheses, we did not find that participants assigned to wearing the plain dress were more willing to sensitize others to issues surrounding fancy clothing in the community. This suggests that after violating a norm in public, individuals are not increasingly ready to do it again. In fact, although it does not reach statistical significance, the effect seems to go in the opposite direction: descriptively, both unweighted and weighted regression analyses suggest that deviating from the high-end clothing norm makes participants less willing to publicly advocate against that norm ( $\tau_u = -0.36, SE = .31, p = .25$ ;  $\tau_{ipw} = -1.43, SE = .88, p = .11$ ; Table 4.2).

**Expressing willingness to publicly violate social norms again.** Similarly, participants assigned to wearing the plain dress were not more willing to publicly violate unrelated social norms in future. This null finding appeared using a pre-registered 5-item index and held when we analyze the data item by item (Table 4.2). If anything, some of the results from other analytic strategies suggest the opposite: that participants in the treatment condition were less likely to be willing to publicly

violate social norms. For instance, weighted regression estimates suggest a large and statistically significant negative effect of violating the high-end clothing norm on participants' willingness to drive ( $\tau_{ipw} = -1.97, SE = .59, p < .001$ ; Table 4.2), a highly counter-normative behavior for women in the Hasidic community.

Table 4.2: Effect of treatment on inclination to deviate

| Item [hypothesized direction of the effect $\tau$ ]               | Condition |       | $\tau$                       |                             |
|-------------------------------------------------------------------|-----------|-------|------------------------------|-----------------------------|
|                                                                   | Siddur    | Dress | unweighted                   | weighted                    |
| <b>Adherence with the community's views [-]</b>                   |           |       |                              |                             |
| Poverty improves perspective on life [-]                          | 4.05      | 3.55  | -0.50 <sup>†</sup><br>(0.29) | -1.89***<br>(0.33)          |
| The least fancy people have a better perspective on life [-]      | 4.76      | 4.07  | -0.70*<br>(0.33)             | -0.76***<br>(0.20)          |
| Good reason for every act the asikonim does [-]                   | 6.03      | 5.57  | -0.46*<br>(0.23)             | -1.74***<br>(0.54)          |
| My community is the best Jewish community in the world [-]        | 4.92      | 4.71  | -0.22<br>(0.31)              | -0.79**<br>(0.31)           |
| <b>Inclination to deviate from social norm [+]</b>                |           |       |                              |                             |
| I would like to help innovator by phone (gym) [+]                 | 4.19      | 4.54  | 0.35<br>(0.31)               | 1.59***<br>(0.41)           |
| I would like to help innovator by phone (healthy food) [+]        | 3.95      | 3.92  | -0.03<br>(0.31)              | 1.63**<br>(0.65)            |
| I want to sensitize others to clothing norm issues [+]            | 3.92      | 3.57  | -0.36<br>(0.31)              | -1.43<br>(0.88)             |
| I would like to own a kosher smartphone to buy clothes online [+] | 2.79      | 3.05  | 0.26<br>(0.37)               | -0.58<br>(0.38)             |
| I would like to own a kosher smartphone to buy clothes [+]        | 2.46      | 2.65  | 0.19<br>(0.34)               | 1.08<br>(0.73)              |
| I would like to dress more casually [+]                           | 3.07      | 3.03  | -0.04<br>(0.30)              | 0.60<br>(0.54)              |
| I would like to drive [+]                                         | 4.28      | 4.33  | 0.06<br>(0.43)               | -1.97***<br>(0.59)          |
| I would like to volunteer outside of the community [+]            | 4.05      | 3.53  | -0.52<br>(0.32)              | 1.43 <sup>†</sup><br>(0.79) |

† p<0.1; \* p<0.05; \*\* p<0.01; \*\*\* p<0.001

Note: Each row of Table 4.2 includes: the average response for the participants in the Siddur and Dress conditions, the effect size ( $\tau$ ) of the treatment derived unweighted ( $\tau_u$ ) and weighted ( $\tau_{ipw}$ ) regression analyses, and the standard error of the effect size of the treatment. Statistical significance is indicated using stars. The figure appended to the right side of the table provides a visualization of each unweighted effect sizes.

### 4.3.5 Psychological outcomes associated with shifts in individuals' inclination to deviate

**Perception of the self as deviant.** Contrary to our hypotheses, violating the high-end clothing norms did not lead participants to integrate deviance into their self-concept. If anything, the results provide some evidence of the opposite, that is, deviating increase individuals' awareness of their acts of conformity. Although the average responses of treatment and control participants to both items measuring self-concept were not statistically significant, estimates from weighted regression analyses (IPW) suggest that wearing the plain dress led participants to consider themselves as the type of person who "always tries to fit". The estimated effect of the treatment of 0.82 ( $SE = .27, p < .001$ ; Table 4.3) on a scale from 1 to 7 is statistically significant and corresponds to a 20% change in one's perception of the self as deviant.

**Perception of the costs of violating social norms.** The results suggest that, consistent with our hypotheses, wearing the plain dress led participants to depreciate their perception of the costs of violating norms in general. Specifically, compared to those assigned to the treatment condition, participants who violated the high-end clothing norm expressed lower levels of agreement with the statement: "I never do something different from friends because it feels too bad." As displayed in Table 4.3, this causal effect of deviance is marginally significant based on our primary unweighted analyses ( $\tau_u = -0.43, SE = .26, p < .1$ ) and statistically significant at the 1% significance level based on weighted regression analyses ( $\tau_{ipw} = -0.96, SE = .21, p < .001$ ). Unexpectedly, although participants perceived the costs of deviance as lower than others after wearing the plain dress, they did not become more inclined to violate norms again (Table 4.2). This is unexpected and inconsistent with past findings that decreased perceptions of the costs of deviance is associated with increased propensity to deviate, an issue that we explore further in the discussion. Regardless, note that

control and treatment participants did not differ in their responses to the other item aiming to assess the costs of deviating: “It’s easy for me to do something different from my community in order to improve it” (Table 4.3). The reasons for this null effect may be that this item closely mirrored the plain dress activity, which participants did not find easy (Table 4.1).

**Perception of high-end clothing.** Violating the high-end clothing norm led treatment condition participants to perceive this norm as significantly weaker (Table 4.3). Participants indicated their level of agreement with the idea that Hasidic women generally believe that buying expensive clothes is important. Participants assigned to the treatment condition agreed significantly less with this statement ( $M_T = 4.61, SD = 1.46$ ) than participants in the control condition ( $M_C = 5.11, SD = 1.30$ ). This difference of 0.51 ( $SE = .24, p < .05$ ) on a scale from 1 to 7 corresponds to a 10% change in treatment condition participants’ perception of the strength of the norm that they just violated. The reason for this shift could be that participants had anticipated more social and psychological pain than they actually experienced for violating the high-end clothing norm. This interpretation, which we consider in depth in the discussion, finds support in the weighted analysis of participants’ responses to the item: “If I don’t wear expensive clothes, my family and friends will not like me as much.” The weighted regression analysis suggests a negative causal effect of violating the high-end clothing norm on participants’ belief that others’ won’t like them if they deviate ( $\tau_{ipw} = -0.90, SE = .26, p < .001$ ; Table 4.3). The weighted analysis (but not the unweighted analysis) also reveal that wearing the plain dress had a negative impact on how much participants value wearing high-end clothes. Specifically, wearing the plain dress decreased participants’ perceived importance of buying and wearing expensive clothes ( $\tau_{ipw} = -0.96, SE = .28, p < .001$ ; Table 4.3). However, consistent with the null findings that violating the high-end clothing did not make participants

more inclined to publicly deviating in future, treatment and control groups did not differ in their willingness to buy and wear less expensive clothes (Table 4.3).

Table 4.3: Effect of treatment on possible psychological mechanisms

| Item [hypothesized direction of the effect $\tau$ ]             | <i>Condition</i> |       | $\tau$                       |                    |
|-----------------------------------------------------------------|------------------|-------|------------------------------|--------------------|
|                                                                 | Siddur           | Dress | unweighted                   | weighted           |
| <b>Perception of the self as deviant</b> [+]                    |                  |       |                              |                    |
| I always try to fit (reversed) [+]                              | 3.93             | 4.09  | 0.16<br>(0.28)               | -0.82***<br>(0.27) |
| I feel strong enough to deviate [+]                             | 4.77             | 4.87  | 0.10<br>(0.26)               | 0.28<br>(0.19)     |
| <b>Perception of the costs of violating social norms</b> [-]    |                  |       |                              |                    |
| Easy for me to deviate (reversed) [-]                           | 3.67             | 2.59  | -0.08<br>(0.25)              | 0.11<br>(0.38)     |
| Deviating makes me feel too bad [-]                             | 3.01             | 2.58  | -0.43 <sup>†</sup><br>(0.26) | -0.96***<br>(0.21) |
| <b>Perception of high-end clothing</b> [0]                      |                  |       |                              |                    |
| Women believe that wearing expensive clothes is important [0]   | 5.11             | 4.61  | -0.51*<br>(0.24)             | -0.95***<br>(0.29) |
| Women believe that close others won't like them otherwise [0]   | 3.29             | 3.37  | 0.07<br>(0.34)               | -1.16***<br>(0.31) |
| I believe that wearing expensive clothes is important [0]       | 3.09             | 2.87  | -0.22<br>(0.29)              | -0.96***<br>(0.28) |
| I believe that close others won't like them otherwise [0]       | 2.04             | 2.03  | -0.01<br>(0.27)              | -0.90***<br>(0.26) |
| I want to purchase and wear cheaper clothes [0]                 | 4.03             | 3.95  | -0.08<br>(0.32)              | -0.01<br>(0.23)    |
| <b>Sensitivity to social issues</b> [+]                         |                  |       |                              |                    |
| I hope things will change [+]                                   | 4.90             | 4.29  | -0.61 <sup>†</sup><br>(0.32) | -0.09<br>(0.36)    |
| I want to volunteer for Ultra-Orthodox Jewish organizations [+] | 5.71             | 5.36  | -0.35<br>(0.23)              | 0.55<br>(0.37)     |

<sup>†</sup>p<0.1; \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

Note: Each line includes: the average response for the participants in the Siddur and plain dress conditions, the effect size ( $\tau$ ) of the treatment derived unweighted ( $\tau_u$ ) and weighted ( $\tau_{ipw}$ ) regression analyses, and the standard error of the effect size of the treatment. Statistical significance is indicated using stars.

**Sensitivity to social issues.** Our exploratory analyses of participants' sensitivity to social issues suggests that violating the high-end clothing norm did not increase participants' willingness to volunteer for Hasidic organizations or their perception of how much change is needed in the community (Table 4.3). However, the experience of violating a norm seems to have shaped participants' perception of what needs to change in the community. An overall 40% of the participants from each experimental condition answered an open-ended question asking what specifically needs to change

in the community. First, about 20% of the control condition participants' responses focused on spiritual changes of the inner self whereas none of the treatment condition participants wrote about this theme . Instead, the majority (i.e., 63%) of the participants assigned to violate the high-end clothing norm complained about social pressure in the community and explicitly stated that this is what needs to change. In contrasts, less than half of the control group (48%) raised issues related to social pressure.

## 4.4 General Discussion

The present study provides some of the first evidence of how the experience of publicly violating a social norm in settled cultures, in which norms are strong, widely followed, and acts of deviance are rare (Swidler, 1986), impacts human psychology. Deviating from a strong norm of clothing in the Hasidic community in Brooklyn did not cause dwellers to become more inclined to publicly violate norms in future. However, this experience made them more critical of their community and more supportive of deviance. Specifically, violating a norm in public caused participants to agree less with core views of their community, express lower levels of endorsement for local authority figures, and be more inclined to supporting innovators' counter-normative projects. Taken together, these results suggest that publicly deviating from a social norm may constitute a first step towards growing tolerance for deviance and developing deviant views and private (but not public) behaviors.

Further analyses of possible psychological mechanisms underlying these findings revealed that the experience of deviance affected participants' perception of social norms and deviance. Although treatment participants reported that violating the high-end clothing norm was socially and psychologically challenging (e.g., feeling judged, negative affect), it led them to perceive this norm as weaker. Furthermore,

the experience of violating the clothing norm caused participants to mitigate their perception of the costs of violating social norms in general.

Why did this experience of deviating from a norm decrease participants' perception of both the strength of the norm that they broke, and the costs associated with deviating from norms in general? The studies described in Chapter 2 of the present dissertation suggest that participants might have anticipated disproportionately strong reactions from their family, friends, and acquaintances. In these behavioral game studies, when the experience of deviance was perceived as less costly than anticipated, individuals depreciated their perception of the costs of violating norms.

However, although treatment participants did depreciate their perception of the costs of publicly violating norms, they were unexpectedly not more inclined to do it again. A possible explanation is that deviating from the high-end clothing norm did not mitigate participants' perception of the costs of violating social norms *enough* to make them want to repeat the experience. After deviating, individuals may still anticipate that the social or psychological costs of deviating will outweigh the possible benefits of deviating. We expect this to be especially the case in settled cultures such as the Hasidic community, in which norms are so internalized that the benefits of violating norms may not be obvious to individuals (Schwartz, 1973). As a result, individuals from settled cultures may generally not be willing to violate a norm enough to actually do it. More field research should explore how individuals' perception the costs and benefits of violating social norms compare in settled and unsettled social groups, and the impact of these perceptions on individuals' behaviors.

In the present study, contrary to predictions from theories of the self, deviating from a norm did not lead participants to incorporate deviance into their self-concept (Markus, 1977; Swann and Bosson, 2010). If anything, the results provide some evidence that the opposite happened: Deviating enhanced participant's perception of the self as conformist, generally speaking. A possible reason for this effect

is that deviance is so rare in the Hasidic community that participants assigned to the treatment condition had never violated a norm before. Instead of feeding deviance into their self-concept, this novel and challenging psychological experience could have made its uniqueness salient, leading participants to realize that they generally conform, not deviate. More research should explore how challenging, first-time social experiences impact one's perception of the self and future behaviors.

Pre-registered exploratory analyses also suggested that publicly violating a norm led participants to argue against social pressure in their community. In addition, whereas 20% of the control group focused on the importance of spiritual change in the inner self, no participants from the deviance condition mentioned spiritual change or the self. Instead, deviating seemed to channel participants' attention towards group dynamics and questions related to social change. Future research should explore further whether encouraging individuals to deviate makes them more open to transitioning to a more unsettled cultural setting, in which norms are weaker and deviance is more frequent. If so, the experience of violating norms could constitute a first step towards the development of social change in settled cultural periods.

### **Potential Limitations**

Culture, social status, and social power represent possible limitations to the present findings. In this study, we aimed to make inferences about the consequences of violating a norm in settled cultures. For one, the Hasidic Jews in Brooklyn (New York) constitute one model of a settled cultures, but others settled social groups differ on aspects that may moderate the consequences of violating norms. For instance, acts of deviance can be more severely punished in other settled communities, such as prisons, gangs, or rebel groups (Decker and Pyrooz, 2015; Liebling and Maruna, 2013; Littman, 2018). In more punitive settings, where abiding by group norms can be essential to one's safety, the overall experience of violating a norm may significantly

differ, and its consequences deserve further exploration. For two, some evidence suggests that power and high-status within one's social group facilitates deviance and mitigates the costs associated with counter-normative behaviors (Pike and Galinsky, 2020; van Kleef et al., 2015). In the present study, participants were primarily young Hasidic women, a subgroup within the female ultra-Orthodox population with lower power and social status. The experience of violating a norm for women in positions of power, such as Rebbetzins in the Hasidic community, may have been milder, more rewarding, and perceived differently by other community members. In turn, this qualitatively different experience could have led powerful individuals to be more inclined to violating norms in front of others, and perhaps promote social change. For these reasons, we encourage conceptual replication of the present research in different communities including participants from subgroups with varying social status and power.

### **Future research**

The present study calls for research on individual and social change in settled cultures such as, but not limited to, prisons, religious communities, rebel groups, or gangs. Conducting naturalistic field studies in these social groups can be difficult, but it is not impossible. Researchers can contact local organizations or community members who are already working on producing change in their community, and start a program of research with them (e.g., Blair et al., 2019). Alternatively, researchers could use existing data to explore the causes of change that has already happened using, for instance, natural experiments or quasi-experiments (Keele and Titiunik, 2016; Sekhon, 2009).

# 5

## Conclusion

The present dissertation provides consistent evidence that the experience of deviating from a social norm enhances individuals' inclination to be a deviant. Together, the findings from a series of studies conducted in the lab and the field converge towards one possible mechanism underlying this effect: violating norms causes people to depreciate their expectation of the costs associated with violation. However, publicly deviating from a norm did not lead participants from our studies to update their views of the self. As discussed below, the present thesis challenges and informs theories of deviance, theories of conformity, and theories of the self.

### 5.1 Emergence of Patterns of Deviance in the Life Span

#### 5.1.1 What have we learned about theories of deviance, conformity, and the self?

**Rethinking the emergence of deviance in the life span.** The finding that patterns of deviance can emerge from a public act of deviance at any point in one's

life is consistent with the idea that deviance can be chronic, but challenges traditional views of deviance as personality or a stable trait that develops early in the life span (e.g., Bernberg, 1954; Diefendorff and Mehta, 2007; Gelfand, 2018; O’Neill and Hastings, 2011; Grant, 2016; Snyder, 1979). The present studies suggest that when they deviate from a social norm, individuals tend to reappraise their perception of the costs of deviance. Specifically, their experience of deviance seems to make them realize that deviating did not feel as bad as they expected, or that others did not react as negatively as they anticipated. This depreciation of the costs of deviance does make individuals more likely to deviate in future, but not because of their personality or traits. Instead, people deviate because of a new perception of how mild—as opposed to painful—the experience will be, which makes deviance more appealing.

**Deviance does not motivate conformity.** Classic theories of conformity to norms suggest that deviating makes individuals more inclined to conform in future, as a way to emend their self-concept (e.g., alleviate feelings of guilt or shame, enhance self-esteem) or reestablish positive relationships with their community (Cialdini and Goldstein, 2004; Williams et al., 2000). None of the present studies provide support to these theoretical arguments. Even public acts of deviance in an extremely “settled” culture such as the Hasidic community, which led individuals to feel negatively and judged by others, did not increase their inclination to conformity. If anything, the results suggest the opposite: after violating a strong norm of their community, developed deviant private views, became more critical of their community, and exhibited more support for the counter-normative projects of Hasidic innovators.

**When public act of deviance does not impact the self-concept.** Although deviating from a social norm increased individuals’ propensity to deviate, it did not reshape their perception of the self as a deviant. This was unexpected in light of theories of the self-concept (Swann and Bosson, 2010) or self-schemata (Markus, 1977), which suggest that individuals observe and rationalize their behaviors to make them

part of a coherent narrative about the self. These theories predict that individuals subsequently behave in ways that are consistent with their updated perception of the self.

Different reasons may explain this discrepancy between theory and findings across studies. First, the context of behavioral games played online (Chapter 3) may have not been naturalistic enough for participants to reshape their perception of the self based on their behavior in these games. Second, deviance in the Hasidic community (Chapter 4) is so rare that participants assigned to the deviance condition might have been violating a social norm for the first time in their life. Instead of feeding deviance into their self-concept, this novel and challenging psychological experience could have made its uniqueness salient, leading participants to realize that they never deviate.

### **5.1.2 Who are the “robust conformers”?**

Although the present dissertation focuses on the consequences of violating norms, participants’ behavior highlighted the power of social norms for producing conformity. Perhaps most striking is Chapter 3, which demonstrates that even norms that belong to groups of strangers playing games online and that are not policed by social or monetary punishment are followed by the great majority of the participants. In this context, the great majority of participants conformed to the advertised and implicit (non-advertised) social norms at every round even though it was not in their personal material interest, and around 35% of participants refused to violate their group norm in exchange for real money. This result contributes to a large body of work establishing social norms as powerful drivers of human behavior, and suggests the existence of a particular type of robust norm conformers, who are more relatively more reluctant to violate social norms compared to other people. Researchers may further investigate how these high-conformers react to more appealing offers to violate norms.

## **5.2 The Possible Consequences of Punishment**

### **5.2.1 Misperceived costs of deviance.**

In the present research, participants did not get punished by other members of their social groups for deviating from social norms, which may have contributed to their depreciation of the costs of deviance and increased inclination to be a deviant. This was the case when participants deviated not only in groups of strangers who were not able to punish them (Chapter 3), but also when they violated meaningful norms of their community in front of relatives, friends, and neighbors (Chapter 4). Combined, these results suggest that individuals may generally misperceive the social costs deviance that come with deviance, and that deviating from a social norm leads them to get a more accurate sense of how others react when they observe someone violating a social norm. Future research should explore the prevalence and magnitude of the discrepancy between people's expectation of the social costs of deviance, and the actual costs of violating different norms. In addition, researchers should also examine how perceptions of the costs are affected by one or several instance of punishment.

### **5.2.2 Punishment and the self**

Punishment for violating norms may lead individuals to update their perception of the self as a “deviant.” Theories of the looking-glass self (McIntyre, 2013; Cooley, 1922) would predict that if other group members ostracised individuals or otherwise call attention to their wrongdoing, norm violators may update their perception of the self as deviants, and this self-perception may also drive future deviance. In the present studies, the encouragement or incentive to deviate, paired with the lack of punishment from others, may have given participants sufficient external justification for their behavior, preventing them from generating justifications based on the self.

### **5.2.3 Avoiding punishment: A learning process?**

While a large body of work has demonstrated that the possibility of being punished for violating social norms increases levels of conformity (Fehr and Fischbacher, 2004), scant research has investigated the effects of being punished for violating a social norm. The experience of being punished by others for deviating may deter future acts of deviance, or lead individuals to develop new strategies for minimizing the negative social consequences of deviating in future, which would lead them to feel more capable about deviating again. Strategies may include the ability to justify their deviant behavior to others, which is for future work to explore.

## **5.3 Deviance in Context: Perceived Meaning of Deviance and Social Psychological Outcomes**

### **5.3.1 Intrinsically vs. extrinsically motivated deviance**

Although using encouragement designs with random assignment to being invited (vs. not) to deviate from a norm allows us to make causal claims, this design bounds our inferences to a certain type of deviance: deviance that is mostly extrinsically motivated, that is, driven by contextual factors. In contrast to intrinsically motivated deviance (e.g., Monin and O'Connor, 2011), extrinsically motivated deviance may be interpreted as unintentional, undesired, or as a temporary condition by the individual who deviates. For these reasons, extrinsically motivated deviance may be more difficult to justify to others, rationalize, and perhaps integrate into one's self-concept or identity. The experience of intrinsically motivated deviance, on the flip side, may be more transformative and reshape individuals and their future behaviors. Future research should explore how these different sources of motivations to deviate impact the consequences of violating a social norms.

### 5.3.2 Degree of naturalism

The present dissertation includes studies of deviance in diverse contexts and using different methods with varying degrees of naturalism (Paluck and Cialdini, 2014). Whereas Chapter 2 and Chapter 4 present studies of deviance from existing norms of participants' communities, Chapter 3 uses behavioral games that participants played online, with strangers. Studying the seemingly exact same phenomenon (i.e., deviance) using these radically different approaches led me to reflect upon the idea that the psychology of deviance in games and in one's everyday life may actually be different phenomena.

Game studies are minimal group studies: participants are assigned to play with groups of other players on the moment and learned about the group norm right after joining a given group. In these game settings, the group norm is not internalized and the consequences of individuals' behaviors are constrained to the experiment. Participants have never interacted with each other prior to the study and will not connect outside of the study. In contrast, the stakes of publicly violating norms in one's community, independently of its level of "settledness" (Swidler, 1986), are larger. Norms are meaningful and internalized, generally endorsed by family members, friends, and neighbors, and widely followed and prescriptive. In these more meaningful contexts, community members may expect that public acts of deviance will have both immediate and long-term negative consequences. Individuals may anticipate stronger negative emotions such as guilt and shame for violating a widely followed norm in front of meaningful others. In future, studies of deviance should differentiate between public acts of deviance that are experienced in meaningful vs. less relevant contexts by the norm-violator.

### **5.3.3 Deviating with others**

Future studies should explore the consequences of deviating from norms with others, as a group. Researchers could encourage participants to violate social norms together with their friends, colleagues, relatives, or even with strangers. Studies could explore the content of pre- and post-deviance group discussions among deviants or with conformists, and measure their effects on future behaviors and the formation of minority groups. Theories of social change suggest that with time, minority groups who consistently engage in counter-normative behaviors produce social change in their community (Moscovici, 1976). This channel to social change would be particularly relevant for settled communities, in which persuasion strategies may not be as effective at changing minds and behaviors (Swidler, 1986).

### **5.3.4 Long-term effects of deviance**

Researchers should deploy longitudinal field studies to tackle the ever-important question of long-term effects: How do public acts of deviance shape individuals and their social groups over time? The way individuals process their experience of deviance in the following days or weeks, on their own or with others, may have important psychological, social, and behavioral implications. After violating a norm, individuals may build positive vs. negative narratives around their act of deviance or successfully vs. unsuccessfully justify their behavior to their friends and relatives. These different outcomes may impact individuals' tolerance and support for deviance, their inclination to violate norms in future, and their ability to foster deviance and social change in their community.

# Bibliography

- Almenberg, J., Dreber, A., Apicella, C., and Rand, D. G. (2010). Third Party Reward and Punishment: Group Size, Efficiency and Public Goods. SSRN Scholarly Paper ID 1715305, Social Science Research Network, Rochester, NY.
- Angrist, J. D. and Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, Princeton, New Jersey.
- Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, 193(5):31–35.
- Becker, H. S. (1963). *Outsiders: Studies in the sociology of deviance*. Outsiders: Studies in the sociology of deviance. Free Press Glencoe, Oxford, England.
- Bennett, R. J. and Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85(3):349–360.
- Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(1):122–142.
- Bernberg, R. E. (1954). Personality correlates of social conformity. *Journal of Applied Psychology*, 38(3):148–149.
- Bicchieri, C. (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford University Press.
- Blair, G., Littman, R., and Paluck, E. L. (2019). Motivating the adoption of new community-minded behaviors: An empirical test in Nigeria. *Science Advances*, 5(3):eaau5175.

- Blanton, H. and Christie, C. O. (2003). Deviance regulation: A theory of action and identity.
- Bodankin, M. and Tziner, A. (2009). Constructive Deviance, Destructive Deviance and Personality: How do they interrelate? *The AMFITEATRU ECONOMIC journal*, 11(26):549–564.
- Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101.
- Briggs, S. R., Cheek, J. M., and Buss, A. H. (1980). An analysis of the Self-Monitoring Scale. *Journal of Personality and Social Psychology*, 38(4):679–686.
- Bruner, J. S. (1990). *Acts of Meaning*. Harvard University Press. Google-Books-ID: YHt\_M41uIuUC.
- Burger, J. M. (1987). Desire for control and conformity to a perceived norm. *Journal of Personality and Social Psychology*, 53(2):355–360.
- Centola, D., Becker, J., Brackbill, D., and Baronchelli, A. (2018). Experimental evidence for tipping points in social convention. *Science*, 360(6393):1116–1119.
- Chaiken, S., Wood, W., and Eagly, A. H. (1996). Principles of persuasion. In *Social psychology: Handbook of basic principles*, pages 702–742. Guilford Press, New York, NY, US.
- Chang, L. J., Smith, A., Dufwenberg, M., and Sanfey, A. G. (2011). Triangulating the Neural, Psychological, and Economic Bases of Guilt Aversion. *Neuron*, 70(3):560–572.
- Cialdini, R. B. and Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55(1):591–621.

- Cialdini, R. B. and Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In *The handbook of social psychology, Vols. 1-2, 4th ed*, pages 151–192. McGraw-Hill, New York, NY, US.
- Cooley, C. H. (1922). *Human Nature and the Social Order*. Scribner.
- Crocker, J., Major, B., and Steele, C. (1998). Social stigma. In *The handbook of social psychology, Vols. 1-2, 4th ed*, pages 504–553. McGraw-Hill, New York, NY, US.
- Damian, R. I., Spengler, M., Sutu, A., and Roberts, B. W. (2019). Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years. *Journal of Personality and Social Psychology*, 117(3):674–695.
- DeBono, A., Shmueli, D., and Muraven, M. (2011). Rude and Inappropriate: The Role of Self-Control in Following Social Norms. *Personality and Social Psychology Bulletin*, 37(1):136–146.
- Decker, S. H. and Pyrooz, D. C. (2015). *The Handbook of Gangs*. John Wiley & Sons. Google-Books-ID: 4zmkBgAAQBAJ.
- Deutsch, N. (2009). The Forbidden Fork, the Cell Phone Holocaust, and Other Haredi Encounters with Technology. *Contemporary Jewry*, 29(1):3–19.
- Di Palma, G. and McClosky, H. (1970). Personality and Conformity: The Learning of Political Attitudes. *The American Political Science Review*, 64(4):1054–1073.
- Diefendorff, J. M. and Mehta, K. (2007). The relations of motivational traits with workplace deviance. *Journal of Applied Psychology*, 92(4):967–977.
- Durkheim, E. (1938). *The rules of sociological method*. The rules of sociological method, 8th ed. Chicago, University of Chicago Press.

- Elms, A. C. and Milgram, S. (1966). Personality characteristics associated with obedience and defiance toward authoritative command. *Journal of Experimental Research in Personality*, 1(4):282–289.
- Erikson, K. (2005). *Wayward Puritans: A Study in the Sociology of Deviance*. Pearson/Allyn and Bacon. Google-Books-ID: Mu2PAAAAMAAJ.
- Fehr, E. and Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2):63–87.
- Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994.
- Fehr, E. and Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, 2(7):458.
- Festinger, L. (1950). Informal social communication. *Psychological Review*, 57(5):271–282.
- Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology; Palo Alto*, 44:155.
- Gantman, A., Gomila, R., Martinez, J. E., Matias, J. N., Paluck, E. L., Starck, J., Wu, S., and Yaffe, N. (2018). A pragmatist philosophy of psychological science and its implications for replication. *Behavioral and Brain Sciences*, 41.
- Gelfand, M. J. (2018). *Rule Makers, Rule Breakers: How Tight and Loose Cultures Wire Our World*. Little, Brown Book Group Limited.
- Gerber, A. S. and Green, D. P. (2012). *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton. Google-Books-ID: yxEGywAACAAJ.
- Gilbert, D. T. and Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117(1):21–38.

- Gino, F. (2018). *Rebel Talent: Why it Pays to Break the Rules at Work and in Life*. Pan Macmillan UK.
- Gomila, R. (2019). Logistic or linear? Estimating causal effects of treatments on binary outcomes using regression analysis. Publisher: PsyArXiv.
- Gomila, R. and Clark, C. S. (2019). Missing Data in Experiments: Challenges and Solutions.
- Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6):504–528.
- Grant, A. (2016). *Originals: How Non-Conformists Move the World*. Penguin.
- Haney, C., Banks, C., and Zimbardo, P. (1972). Interpersonal Dynamics in a Simulated Prison:. Technical report, Defense Technical Information Center, Fort Belvoir, VA.
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43(1):59–74.
- Hoffman, K. M., Trawalter, S., Axt, J. R., and Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, 113(16):4296–4301.
- Hornsey, M. J., Majkut, L., Terry, D. J., and McKimmie, B. M. (2003). On being loud and proud: Non-conformity and counter-conformity to group norms. *British Journal of Social Psychology*, 42(3):319–335.
- James, W. (1907). Pragmatism’s Conception of Truth. *The Journal of Philosophy, Psychology and Scientific Methods*, 4(6):141–155.

- Jetten, J. and Hornsey, M. J. (2010). *Rebels in Groups: Dissent, Deviance, Difference, and Defiance*. John Wiley & Sons.
- Keele, L. and Titunik, R. (2016). Natural Experiments Based on Geography. *Political Science Research and Methods*, 4(1):65–95.
- Kirsch, C., Lubart, T., and Houssemand, C. (2015). Creative Personality Profile in Social Sciences: The Leading Role of Autonomy. *Creativity. Theories – Research - Applications*, 2(2):180–211.
- Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior*. Self-presentation: Impression management and interpersonal behavior. Brown & Benchmark Publishers, Madison, WI, US.
- Lennox, R. D. and Wolfe, R. N. (1984). Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology*, 46(6):1349–1364.
- Liebling, A. and Maruna, S. (2013). *The Effects of Imprisonment*. Routledge.
- Link, B. G. and Phelan, J. C. (2001). Conceptualizing Stigma. *Annual Review of Sociology*, 27(1):363–385.
- Littman, R. (2018). Perpetrating Violence Increases Identification With Violent Groups: Survey Evidence From Former Combatants. *Personality and Social Psychology Bulletin*, 44(7):1077–1089.
- Major, B. and O’Brien, L. T. (2005). The Social Psychology of Stigma. *Annual Review of Psychology*, 56(1):393–421.
- Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, 35(2):63–78.
- McIntyre, L. (2013). *The Practical Skeptic: Core Concepts in Sociology*. McGraw-Hill Higher Education. Google-Books-ID: 0RwiAAAAQBAJ.

- Milgram, S. (1963). Behavioral Study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4):371–378.
- Miller, D. T. and Prentice, D. A. (1996). The construction of social norms and standards. In *Social psychology: Handbook of basic principles*, pages 799–829. Guilford Press, New York, NY, US.
- Monin, B. and O’Connor, K. (2011). Reactions to defiant deviants: Deliverance or defensiveness? In *Rebels in groups: Dissent, deviance, difference and defiance*, pages 261–280. Wiley-Blackwell.
- Monroe, K. R. (2008). Cracking the Code of Genocide: The Moral Psychology of Rescuers, Bystanders, and Nazis during the Holocaust. *Political Psychology*, 29(5):699–736.
- Morrison, K. R. and Miller, D. T. (2012). Explaining Differences in Opinion Expression. In *Rebels in Groups*, pages 219–237. John Wiley & Sons, Ltd.
- Moscovici, S. (1976). *Social influence and social change*. published in cooperation with European Association of Experimental Social Psychology by Academic Press. Google-Books-ID: Tce4AAAAIAAJ.
- Moscovici, S. (1994). Three concepts: Minority, conflict, and behavioral style. In *Minority influence*, Nelson-Hall series in psychology, pages 233–251. Nelson-Hall Publishers, Chicago, IL, US.
- O’Neill, T. A. and Hastings, S. E. (2011). Explaining workplace deviance behavior with more than just the “Big Five”. *Personality and Individual Differences*, 50(2):268–273.
- Packer, D. J. (2012). The Dissenter’s Dilemma, and a Social Identity Solution. In *Rebels in Groups*, pages 281–301. John Wiley & Sons, Ltd.
- Packer, D. J. and Chasteen, A. L. (2010). Loyal Deviance: Testing the Normative Conflict Model of Dissent in Social Groups. *Personality and Social Psychol-*

*ogy Bulletin*, 36(1):5–18.

- Paluck, E. L. and Cialdini, R. B. (2014). Field Research Methods. In Reis, H. T. and Judd, C. M., editors, *Handbook of Research Methods in Social and Personality Psychology*, pages 81–97. Cambridge University Press, New York, 2 edition.
- Paluck, E. L. and Green, D. P. (2009). Deference, Dissent, and Dispute Resolution: An Experimental Intervention Using Mass Media to Change Norms and Behavior in Rwanda. *American Political Science Review*, 103(4):622–644.
- Paluck, E. L. and Shepherd, H. (2012). The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network. *Journal of Personality and Social Psychology*, 103(6):899–915.
- Pike, B. E. and Galinsky, A. D. (2020). Power leads to action because it releases the psychological brakes on action. *Current Opinion in Psychology*, 33:91–94.
- Poll, S. (1969). *The Hasidic Community of Williamsburg*. Transaction Publishers. Google-Books-ID: vyMPyLmEJigC.
- Prentice, D. A. (2012). The Psychology of Social Norms and the Promotion of Human Rights. In Goodman, R., Jinks, D., and Woods, A. K., editors, *Understanding Social Action, Promoting Human Rights*, pages 23–43. Oxford University Press.
- Prentice, D. A. and Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2):243–256.
- Prentice, D. A. and Miller, D. T. (1996). Pluralistic Ignorance and the Perpetuation of Social Norms by Unwitting Actors. In Zanna, M. P., editor, *Advances in Experimental Social Psychology*, volume 28, pages 161–209. Academic Press.

- Roberts, S. (2011). Kiryas Joel, N.Y., Lands Distinction as Nation's Poorest Place. *The New York Times*.
- Ross, L. (1977). The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process. In Berkowitz, L., editor, *Advances in Experimental Social Psychology*, volume 10, pages 173–220. Academic Press.
- Ross, L., Lepper, M., and Ward, A. (2010). History of Social Psychology: Insights, Challenges, and Contributions to Theory and Application. In *Handbook of Social Psychology*. American Cancer Society.
- Rozin, P. (2001). Social Psychology and Science: Some Lessons From Solomon Asch. *Personality and Social Psychology Review*, 5(1):2–14.
- Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental Social Psychology*, 9(4):349–364.
- Sekhon, J. S. (2009). Opiates for the Matches: Matching Methods for Causal Inference. *Annual Review of Political Science*, 12(1):487–508.
- Sewell, W. H. (2004). The concept (s) of culture. In Spiegel, G. M., editor, *Practicing history. New directions in historical writing after the linguistic turn*, pages 90–110. Routledge.
- Shashkini, N. (2018). ICC presidents are combating declining membership at sign-in eating clubs. *The Princetonian*. Retrieved from <https://www.dailyprincetonian.com/article/2018/05/icc-presidents-are-combating-declining-membership-at-sign-in-eating-clubs>.
- Sheldon, K. M. (1995). Creativity and self-determination in personality. *Creativity Research Journal*, 8(1):25–36.
- Smith, C. A., Organ, D. W., and Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, 68(4):653–

- Snyder, M. (1979). Self-Monitoring Processes. In Berkowitz, L., editor, *Advances in Experimental Social Psychology*, volume 12, pages 85–128. Academic Press.
- Steele, C. M. (1988). The Psychology of Self-Affirmation: Sustaining the Integrity of the Self. In Berkowitz, L., editor, *Advances in Experimental Social Psychology*, volume 21, pages 261–302. Academic Press.
- Swann, W. B. and Bosson, J. K. (2010). Self and identity. In *Handbook of social psychology, Vol. 1, 5th ed*, pages 589–628. John Wiley & Sons Inc, Hoboken, NJ, US.
- Swann, W. B. and Jetten, J. (2017). Restoring Agency to the Human Actor. *Perspectives on Psychological Science*, 12(3):382–399.
- Swidler, A. (1986). Culture in Action: Symbols and Strategies. *American Sociological Review*, 51(2):273–286.
- Tankard, M. E. and Paluck, E. L. (2016). Norm Perception as a Vehicle for Social Change: Vehicle for Social Change. *Social Issues and Policy Review*, 10(1):181–211.
- Turner, J. C. (2006). Tyranny, freedom and social structure: Escaping our theoretical prisons. *British Journal of Social Psychology*, 45(1):41–46. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1348/014466605X79840>.
- van Kleef, G. A., Wanders, F., Stamkou, E., and Homan, A. C. (2015). The social dynamics of breaking the rules: antecedents and consequences of norm-violating behavior. *Current Opinion in Psychology*, 6:25–31.
- Whitt, S. and Wilson, R. K. (2007). The Dictator Game, Fairness and Ethnicity in Postwar Bosnia. *American Journal of Political Science*, 51(3):655–668.

- Williams, K. D., Cheung, C. K. T., and Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, 79(5):748–762.
- Woolridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, Mass. woolridge2001.
- Yaffe, N. M. (2020). Capabilities and Universalism – An Empirical Examination: The Case of the Ultra-Orthodox Community. *Journal of Human Development and Capabilities*, 21(1):84–98.
- Yaffe, N. M., Solak, N., Halperin, E., and Saguy, T. (2018). “Poor is pious”: Distinctiveness threat increases glorification of poverty among the poor. *European Journal of Social Psychology*, 48(4):460–471.

# Appendix A

## Supplementary materials for Chapter 2

### A.1 Summary statistics for the pilot survey

Table A.1: Description of the sample

| Label                                                              | Range   | Mean<br>deviants | Mean<br>conformists | Mean<br>difference | t-value | p-value |
|--------------------------------------------------------------------|---------|------------------|---------------------|--------------------|---------|---------|
| Deviant                                                            | [0-1]   | 1.00             | 0.00                | 1.00               |         |         |
| Freshman expectations:<br>excited about going to princeton         | [0-100] | 81.36            | 85.43               | -4.06              | -1.76   | 0.08    |
| Freshman expectations:<br>excited about meeting princeton students | [0-100] | 78.12            | 80.88               | -2.76              | -1.19   | 0.24    |
| Freshman expectations: finding similar others                      | [0-100] | 58.80            | 65.26               | -6.46              | -2.50   | 0.01    |
| Similarity high school vs. princeton                               | [0-100] | 36.74            | 40.32               | -3.58              | -1.47   | 0.14    |
| Moved a lot before Princeton                                       | [0-10]  | 2.12             | 2.30                | -0.19              | -1.03   | 0.31    |
| Identification with high school                                    | [0-100] | 56.64            | 61.44               | -4.81              | -1.49   | 0.14    |
| Contact with important past reference groups                       | [0-6]   | 4.19             | 4.57                | -0.39              | -2.00   | 0.05    |
| Past instances of deviance                                         | [0-3]   | 2.91             | 2.54                | 0.36               | 3.02    | 0.00    |
| Disliked high school main norms                                    | [0-3]   | 0.16             | 0.16                | -0.00              | -0.07   | 0.94    |
| Strength of past norm                                              | [0-100] | 81.16            | 81.76               | -0.60              | -0.35   | 0.73    |
| Found friends in Princeton                                         | [0-3]   | 2.21             | 2.29                | -0.08              | -0.96   | 0.34    |
| Belonging to Princeton                                             | [0-100] | 61.95            | 68.05               | -6.11              | -2.41   | 0.02    |
| Identification with Princeton                                      | [0-100] | 71.82            | 77.91               | -6.09              | -2.35   | 0.02    |
| Complexity of identification with Princeton                        | [0-100] | 79.31            | 76.10               | 3.21               | 1.22    | 0.22    |
| Dislike Princeton strongest norms                                  | [0-3]   | 0.50             | 0.31                | 0.19               | 2.87    | 0.00    |
| Time it took to identify with Princeton                            | [1-7]   | 3.19             | 2.93                | 0.26               | 1.06    | 0.29    |
| Princeton norm homogeneity                                         | [0-100] | 82.04            | 84.62               | -2.58              | -1.55   | 0.12    |
| Feel similar to the typical princeton student                      | [0-100] | 47.09            | 57.51               | -10.42             | -5.17   | 0.00    |
| Homogeneity of princeton students                                  | [0-100] | 60.18            | 62.63               | -2.45              | -1.49   | 0.14    |
| Eating club norm strength                                          | [0-100] | 62.33            | 64.06               | -1.73              | -1.36   | 0.17    |
| Mental health: happiness                                           | [0-100] | 68.53            | 72.39               | -3.86              | -1.70   | 0.09    |
| Mental health: depression                                          | [0-4]   | 0.81             | 0.61                | 0.19               | 1.71    | 0.09    |
| Mental health: sadness                                             | [0-4]   | 1.28             | 1.11                | 0.17               | 1.59    | 0.11    |
| Thought about transferring                                         | [0-2]   | 0.44             | 0.45                | -0.01              | -0.13   | 0.90    |
| Ethnicity white                                                    | [0-1]   | 0.38             | 0.41                | -0.03              | -0.50   | 0.62    |
| Ethnicity black                                                    | [0-1]   | 0.10             | 0.09                | 0.02               | 0.49    | 0.63    |
| Ethnicity hispanic                                                 | [0-1]   | 0.07             | 0.11                | -0.04              | -1.36   | 0.17    |
| Ethnicity asian                                                    | [0-1]   | 0.33             | 0.32                | 0.01               | 0.29    | 0.77    |
| Family income                                                      | [1-5]   | 3.18             | 3.57                | -0.39              | -3.49   | 0.00    |
| Religiosity                                                        | [1-4]   | 2.82             | 2.23                | 0.59               | 3.63    | 0.00    |
| Gender                                                             | [0-1]   | 0.41             | 0.42                | -0.01              | -0.17   | 0.86    |
| Age                                                                | [18-23] | 19.58            | 19.42               | 0.16               | 1.93    | 0.05    |

## A.2 Factor analysis

The factor analysis revealed that four main components explain about 75% of the variance.

Table A.2: Factor analysis

|                        | PC1      | PC2     | PC3     | PC4      | PC5       | PC6       | PC7      |
|------------------------|----------|---------|---------|----------|-----------|-----------|----------|
| Standard deviation     | 2.303949 | 1.42002 | 1.14301 | 1.071893 | 0.9828263 | 0.7231126 | 0.652195 |
| Proportion of Variance | 0.408320 | 0.15511 | 0.10050 | 0.088380 | 0.0743000 | 0.0402200 | 0.032720 |
| Cumulative Proportion  | 0.408320 | 0.56343 | 0.66393 | 0.752310 | 0.8266200 | 0.8668400 | 0.899560 |

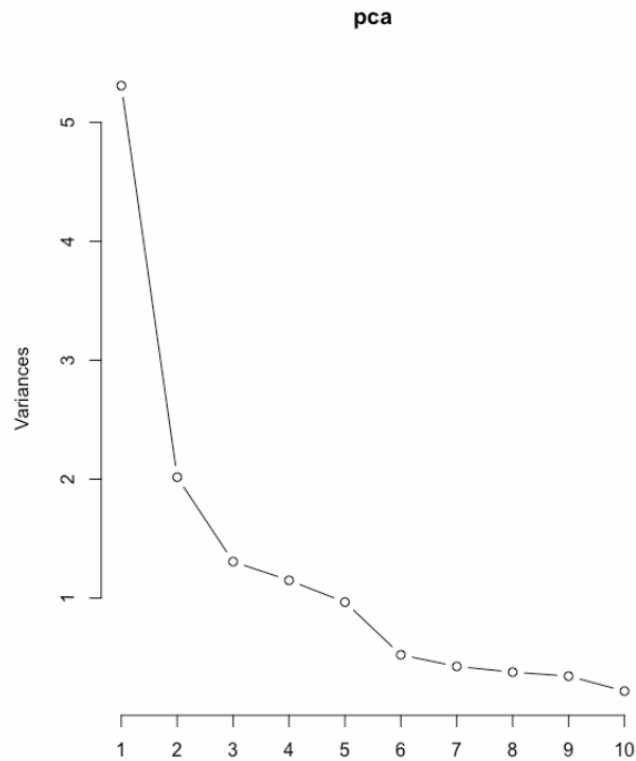


Figure A.1: Scree plot indicating variance captured by each component.

Table A.3: Relative contributions to each component of the PCA

|                            | PC1       | PC2       | PC3       | PC4       | PC5       | PC6       | PC7       |
|----------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Feel similar to others     | 0.1016947 | 0.0026755 | 0.0682956 | 0.0474329 | 0.0859878 | 0.0512825 | 0.1275700 |
| Felt different in the past | 0.0656366 | 0.0252809 | 0.2644480 | 0.0180578 | 0.0572514 | 0.1027039 | 0.1495600 |
| Fit into past groups       | 0.0539765 | 0.0186811 | 0.2866548 | 0.0745909 | 0.0467856 | 0.0632748 | 0.1798644 |
| Past deviance              | 0.0336207 | 0.0287059 | 0.0038136 | 0.2768106 | 0.1416376 | 0.1809029 | 0.0107622 |
| Other deviance             | 0.0702032 | 0.1034584 | 0.0032054 | 0.1895281 | 0.0301694 | 0.2519228 | 0.0372593 |
| Belonging                  | 0.1186815 | 0.0017099 | 0.0304368 | 0.0149730 | 0.0367559 | 0.0027386 | 0.0216740 |
| Social belonging           | 0.1159403 | 0.0088810 | 0.0054453 | 0.0084992 | 0.0701716 | 0.0083322 | 0.0049022 |
| Academic belonging         | 0.0767163 | 0.1308893 | 0.1239353 | 0.0512854 | 0.1391625 | 0.0921665 | 0.0971683 |
| Preference ac. belonging   | 0.0280663 | 0.2322779 | 0.0429276 | 0.1053075 | 0.0871357 | 0.0168044 | 0.0665009 |
| Identification             | 0.1145855 | 0.0187554 | 0.0366659 | 0.0438173 | 0.0126994 | 0.1036491 | 0.1310125 |
| Social identification      | 0.1143521 | 0.0506383 | 0.0097086 | 0.0214184 | 0.0954126 | 0.0690268 | 0.0532782 |
| Academic identification    | 0.0866839 | 0.1340514 | 0.0939101 | 0.0677737 | 0.1145095 | 0.0402652 | 0.0481845 |
| Preference ac. identif.    | 0.0198423 | 0.2439949 | 0.0305532 | 0.0805051 | 0.0823209 | 0.0169304 | 0.0722634 |

Note: This table displays the proportion that each variable contributes to each component of the PCA. Put differently, each value in the table represents the proportional contribution to each principal component.

### A.3 Identification with Princeton

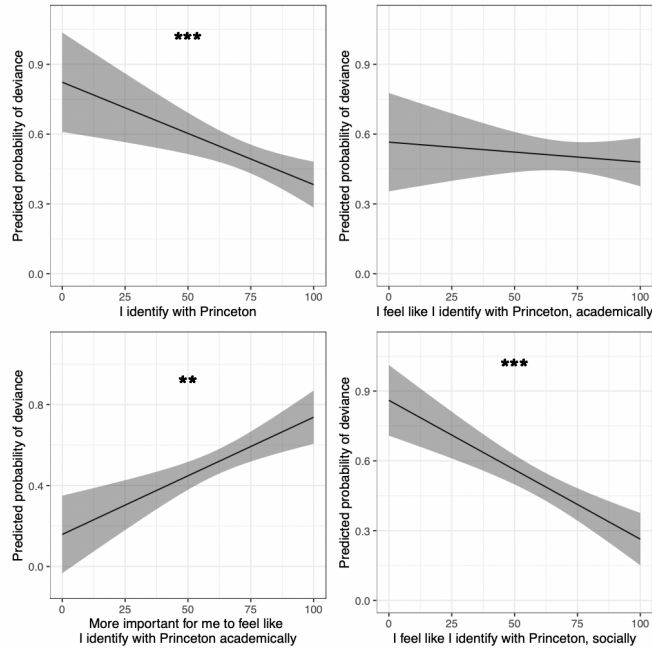


Figure A.2: Identification predicts deviance. These predicted probabilities of deviance are derived from separate regressions of the binary eating club deviance item on self-reported identification items, controlling for demographic variables set at their mean.

## A.4 Summary of the main survey analyses

Table A.4: Table of coefficients for Study 2

| Predictor                                           | $\beta_0$ | $\beta_1$ | SE    | t-value  | $R^2_{adj}$ |
|-----------------------------------------------------|-----------|-----------|-------|----------|-------------|
| I feel similar to other Princeton students          | 0.83      | -0.006    | 0.002 | -3.57*** | 0.166       |
| My past groups made me feel different               | 0.49      | 0.004     | 0.001 | 2.54**   | 0.140       |
| Age at which I started to feel different            | 0.85      | -0.021    | 0.009 | -2.45**  | 0.135       |
| I fit well into my past social groups               | 1.05      | -0.005    | 0.002 | -2.47**  | 0.138       |
| Past instances of deviance                          | 0.49      | 0.041     | 0.026 | 1.55     | 0.153       |
| Age of first time I deliberately deviated           | 1.08      | -0.035    | 0.008 | -4.14*** | 0.219       |
| Amount of other instances of deviance at Princeton  | 0.24      | 0.006     | 0.001 | 4.49***  | 0.223       |
| I belong to Princeton                               | 0.90      | -0.005    | 0.001 | -3.70*** | 0.190       |
| I feel like I belong to Princeton, socially         | 0.91      | -0.006    | 0.001 | -4.53*** | 0.215       |
| I feel like I belong to Princeton, academically     | 0.56      | 0.002     | 0.002 | 1.19     | 0.147       |
| More important for me to belong academically        | 0.22      | 0.007     | 0.001 | 4.67***  | 0.223       |
| I identify with Princeton                           | 0.88      | -0.004    | 0.001 | -3.08*** | 0.181       |
| I feel like I identify with Princeton, socially     | 0.90      | -0.006    | 0.001 | -4.99*** | 0.244       |
| I feel like I identify with Princeton, academically | 0.68      | -0.001    | 0.001 | -0.59    | 0.143       |
| More important for me to I identify academically    | 0.24      | 0.006     | 0.002 | 3.82***  | 0.201       |
| Positive mental health index                        | 0.66      | -0.043    | 0.032 | -1.33    | 0.149       |
| Social monitoring index                             | 1.27      | -0.176    | 0.056 | -3.15*** | 0.180       |
| Extraversion                                        | 0.75      | -0.027    | 0.019 | -1.44    | 0.151       |
| Agreeableness                                       | 0.74      | -0.022    | 0.028 | -0.77    | 0.145       |
| Conscientiousness                                   | 0.62      | 0.003     | 0.027 | 0.09     | 0.142       |
| Emotional stability                                 | 0.64      | -0.000    | 0.023 | -0.003   | 0.142       |
| Openness to experience                              | 0.51      | 0.023     | 0.032 | 0.71     | 0.144       |
| I am a maverick                                     | 0.60      | 0.008     | 0.025 | 0.31     | 0.142       |
| I am an independent thinker                         | 0.40      | 0.041     | 1.49  | 0.136    | 0.151       |
| I am socially awkward                               | 0.47      | 0.033     | 0.019 | 1.670    | 0.154       |
| I am a rebel                                        | 0.54      | 0.023     | 0.020 | 1.15     | 0.147       |

Note: This table displays (respectively) the intercept, unstandardized coefficients, standardized coefficients, standard errors, t-values, and adjusted  $R^2$  for each multiple regression analysis conducted in the context of Study 2. Each of these regression analyses included the following covariates (in parentheses, we report the range of each covariate's coefficients across regressions): ethnicity (0.00; 0.14), gender (0.05; 0.12), income (-0.12; -0.08), and religiosity (0.05; 0.08).

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

# Appendix B

## Supplementary materials for Chapter 3

### B.1 Description of Study 1 and Study 2 samples

Table B.1: Description of Study 1 sample

|                              | Range | All participants | Control | Treatment | Difference | P-value |
|------------------------------|-------|------------------|---------|-----------|------------|---------|
| Gender                       | 0-1   | 0.48             | 0.51    | 0.46      | -0.05      | 0.04*   |
| Age                          | 18-76 | 35.00            | 34.68   | 34.58     | -0.10      | 0.87    |
| Income                       | 1-5   | 2.61             | 2.60    | 2.62      | 0.01       | 0.79    |
| White                        | 0-1   | 0.77             | 0.78    | 0.77      | 0.00       | 0.88    |
| Black                        | 0-1   | 0.07             | 0.07    | 0.07      | 0.00       | 0.8     |
| Asian                        | 0-1   | 0.06             | 0.06    | 0.06      | 0.00       | 0.96    |
| Hispanic                     | 0-1   | 0.05             | 0.05    | 0.06      | 0.01       | 0.67    |
| Experience wt games          | 1-4   | 1.58             | 1.60    | 1.56      | -0.04      | 0.33    |
| Pre-treatment contribution 1 | 0-100 | 88.10            | 88.31   | 87.89     | -0.43      | 0.74    |
| Pre-treatment contribution 2 | 0-100 | 87.58            | 87.58   | 87.59     | 0.01       | 1       |

Note: Table B.1 displays descriptive statistics for the entire pool of participants that took part in Study 1 ( $N = 1,439$ ). Specifically, Table B.1 indicates the average value of pre-registered covariates and pretreatment contributions at the Public Goods Game for the entire sample, as well as for each experimental condition separately. Table S1 also displays the difference in means between the treatment and control condition, and the p-values from separate statistical tests (t-tests) for these differences in means.

Table B.2: Description of Study 2 sample

|                              | Range | All participants | Control | Treatment | Difference | P-value |
|------------------------------|-------|------------------|---------|-----------|------------|---------|
| Gender                       | 0-1   | 0.45             | 0.46    | 0.45      | -0.02      | 0.41    |
| Age                          | 18-81 | 33.00            | 32.86   | 33.72     | 0.85       | 0.12    |
| Income                       | 1-5   | 2.57             | 2.55    | 2.59      | 0.03       | 0.47    |
| White                        | 0-1   | 0.72             | 0.72    | 0.72      | 0.00       | 0.82    |
| Black                        | 0-1   | 0.07             | 0.06    | 0.08      | 0.02       | 0.1     |
| Asian                        | 0-1   | 0.09             | 0.09    | 0.09      | 0.00       | 0.96    |
| Hispanic                     | 0-1   | 0.07             | 0.07    | 0.07      | 0.01       | 0.65    |
| Experience wt games          | 1-4   | 1.53             | 1.50    | 1.56      | 0.06       | 0.08.   |
| Pre-treatment contribution 1 | 0-100 | 89.46            | 89.99   | 88.92     | -1.07      | 0.28    |
| Pre-treatment contribution 2 | 0-100 | 88.74            | 89.74   | 87.70     | -2.04      | 0.05.   |

Note: Table B.2 displays descriptive statistics for the entire pool of participants that took part in Study 2 ( $N = 2,060$ ). Specifically, Table B.2 indicates the average value of pre-registered covariates and pretreatment contributions at the Public Goods Game for the entire sample, as well as for each experimental condition separately. Table S2 also displays the difference in means between the treatment and control condition, and the p-values from separate statistical tests (t-tests) for these differences in means.

## B.2 Results of the Public Goods Game by Round

Table B.3: Intent to Treat Analysis

|                             | <i>Condition</i> |           |                     |
|-----------------------------|------------------|-----------|---------------------|
|                             | Control          | Treatment | $\tau$              |
| <b>Playing with Group 1</b> |                  |           |                     |
| Round a                     | 88.31            | 87.89     | -0.43<br>(1.30)     |
| Round b                     | 87.58            | 87.59     | 0.01<br>(1.37)      |
| Treatment Round (Round c)   | 85.15            | 38.97     | -46.18***<br>(1.88) |
| <b>Playing with Group 2</b> |                  |           |                     |
| Round a                     | 84.16            | 80.14     | -4.02**<br>(1.49)   |
| Round b                     | 80.56            | 77.07     | -3.49**<br>(1.77)   |
| <b>Playing with Group 3</b> |                  |           |                     |
| Round a                     | 54.72            | 50.59     | -4.12**<br>(1.48)   |
| Round b                     | 47.59            | 43.03     | -4.56***<br>(1.47)  |

Note: Table B.3 displays participants' contributions to the Public Goods Game by round, broken down by experimental condition. The last column of the table reveals the size of intent-to-treat effect ( $\tau$ ) for each round and its standard errors.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

Table B.4: Intent to Treat Analysis

|                             | Control | $\tau$              |
|-----------------------------|---------|---------------------|
| <b>Playing with Group 1</b> |         |                     |
| Round a                     | 88.31   | -0.64<br>(1.96)     |
| Round b                     | 87.58   | 0.01<br>(2.07)      |
| Treatment Round (Round c)   | 85.15   | -69.85***<br>(1.75) |
| <b>Playing with Group 2</b> |         |                     |
| Round a                     | 84.16   | -6.09**<br>(2.21)   |
| Round b                     | 80.56   | -5.28*<br>(2.64)    |
| <b>Playing with Group 3</b> |         |                     |
| Round a                     | 54.72   | -6.24***<br>(2.20)  |
| Round b                     | 47.59   | -6.90***<br>(2.18)  |

Note: Table B.4 displays the contributions of control condition participants to the Public Goods Game and the treatment on treated effect ( $\tau$ ) broken down by round. Effect sizes ( $\tau$ ) was calculated using Instrumental Variable analysis (IV). Standard errors are indicated in parentheses and statistical significance is signaled with stars. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

## B.3 Study 2 Results

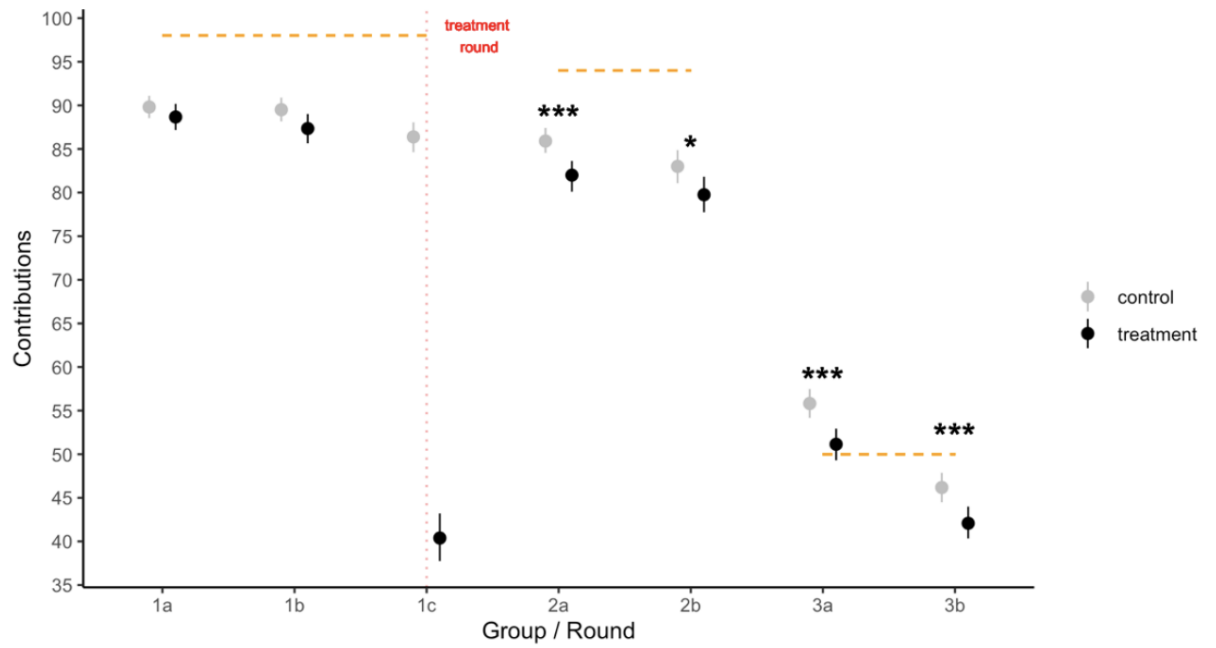


Figure B.1: Public Goods Game by round (ITT).

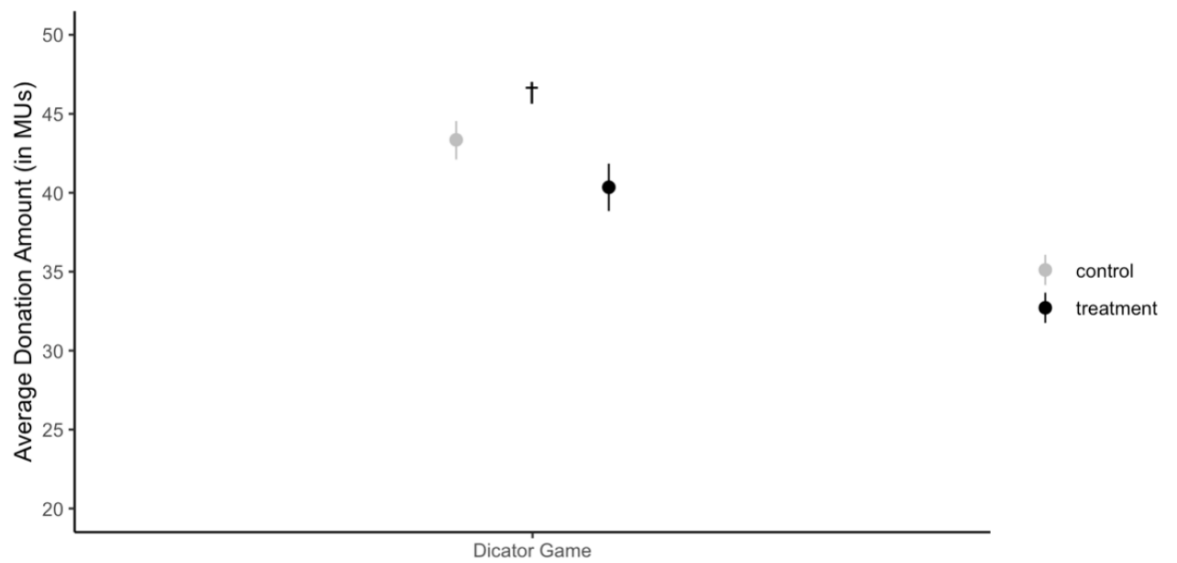


Figure B.2: Dictator Game (ITT).

Table B.5: Intent-to-Treat Analyses (Study 2)

|                                       | <i>Condition</i> |           |                              |
|---------------------------------------|------------------|-----------|------------------------------|
|                                       | Control          | Treatment | $\tau$                       |
| <b>Playing with Group 1</b>           |                  |           |                              |
| Round a                               | 89.99            | 88.92     | -1.07<br>(0.98)              |
| Round b                               | 89.74            | 87.70     | -2.04 <sup>†</sup><br>(1.06) |
| Treatment Round (Round c)             | 86.54            | 40.71     | -45.83***<br>(1.56)          |
| <b>Playing with Group 2</b>           |                  |           |                              |
| Round a                               | 86.12            | 82.35     | -3.77***<br>(1.12)           |
| Round b                               | 83.30            | 80.24     | -3.06**<br>(1.36)            |
| <b>Playing with Group 3</b>           |                  |           |                              |
| Round a                               | 55.84            | 51.21     | -4.63***<br>(1.16)           |
| Round b                               | 46.46            | 42.51     | -3.95***<br>(1.19)           |
| <b>Playing the Dictator Game (DG)</b> |                  |           |                              |
| One-Shot DG (donation)                | 42.34            | 40.55     | -1.80 <sup>†</sup><br>(1.04) |

Note: Table B.5 displays participants' contributions to the Public Goods Game by round and at the Dictator Game (Study 2), broken down by experimental condition. The last column of the table reveals the size of intent-to-treat effect ( $\tau$ ) for each round and its standard errors.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

Table B.6: Instrumental Variable Analyses (Study 2)

|                                       | Control | $\tau$                       |
|---------------------------------------|---------|------------------------------|
| <b>Playing with Group 1</b>           |         |                              |
| Round a                               | 89.99   | -1.66<br>(1.52)              |
| Round b                               | 89.74   | -3.17 <sup>†</sup><br>(1.64) |
| Treatment Round (Round c)             | 86.54   | -71.32***<br>(1.40)          |
| <b>Playing with Group 2</b>           |         |                              |
| Round a                               | 86.12   | -5.87***<br>(1.71)           |
| Round b                               | 83.30   | -4.75**<br>(2.09)            |
| <b>Playing with Group 3</b>           |         |                              |
| Round a                               | 55.84   | -7.21***<br>(1.77)           |
| Round b                               | 46.46   | -6.15***<br>(1.82)           |
| <b>Playing the Dictator Game (DG)</b> |         |                              |
| One-Shot DG (donation)                | 42.34   | -2.79 <sup>†</sup><br>(1.61) |

Note: Table B.6 displays the contributions of control condition participants to the Public Goods Game and at the Dictator Game, and the treatment on treated effect ( $\tau$ ) broken down by round. Effect sizes ( $\tau$ ) were calculated using Instrumental Variable analysis (IV). Standard errors are indicated in parentheses and statistical significance is signaled with stars.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

## B.4 Effects Are Not Due to Learning Processes

Two different analytic strategies demonstrate that the impact of the incentive to deviate on future deviations was not driven by a group of participants who learned to

give the same low amount each time (10 MUs) nor by a group of outliers who simply adjusted their playing strategy to include low amounts each time (close to 10 MUs).

**Strategy 1:** We created a binary variable that indicates whether participants contributed between 0 and 10 MUs vs. more than 10 MUs post-treatment. We found that the probability that a treatment participant makes a contribution between 0 and 10 MUs at round 2a or 2b is not significantly larger than the probability that a control participant makes a contribution between 0 and 10 MUs at round 2a or 2b.

**Strategy 2:** We created a binary variable which returns:

- 0 if the difference between the contribution at the last pre-treatment round and first post-treatment round is greater than 50 (threshold). For instance, a participant who contributed 98 pre-treatment and less than 48 post-treatment returns 0.
- 1 otherwise (e.g., participants who went from contributing 98 pre-treatment to contributing between 48 and 93 return 1).

We find no difference between treatment and control condition regarding the proportion of participants who went from making a large contribution before the treatment round to an extremely low contribution after the treatment. This analysis is robust to modifications to the threshold. That is, the results hold if the threshold is any number between 20 and 80.

## Appendix C

### Supplementary materials for Chapter 4

#### C.1 Instrumental Variable analyses

Table C.1: Effects wearing the plain dress on discomfort with the activity and inclination to deviate.

| Item [expected direction of average treatment effect $\tau$ ]     | Condition |                                       | $\tau$            |
|-------------------------------------------------------------------|-----------|---------------------------------------|-------------------|
|                                                                   | Siddur    | Dress                                 |                   |
| <b>Negative experience with the activity</b> [+]                  |           |                                       |                   |
| I felt like I stood out [+]                                       | 1.62      | 4.80                                  | 3.18***<br>(0.35) |
| I felt good (reversed) [+]                                        | 2.98      | 4.30                                  | 1.32***<br>(0.33) |
| I felt weird [+]                                                  | 1.80      | 4.95                                  | 3.15***<br>(0.35) |
| People noticed something unusual [+]                              | 1.52      | 4.16                                  | 2.64***<br>(0.40) |
| People were a little judgmental [+]                               | 1.51      | 3.62                                  | 2.11***<br>(0.35) |
| <b>Adherence with the community's views</b> [−]                   |           |                                       |                   |
| Poverty gives better perspective [−]                              | 4.05      | 3.45                                  | −0.60†<br>(0.36)  |
| The least fancy people have a better perspective [−]              | 4.76      | 3.92                                  | −0.84*<br>(0.41)  |
| Good reason for every act the asokonim does [−]                   | 6.03      | 5.29                                  | −0.75*<br>(0.36)  |
| My community is the best Jewish community in the world [−]        | 4.92      | 4.56                                  | −0.37<br>(0.52)   |
| <b>Inclination to deviate from social norms</b> [+]               |           |                                       |                   |
| I would like to help innovator by phone (gym) [+]                 | 4.19      | 4.76                                  | 0.57<br>(0.49)    |
| I would like to help innovator by phone (healthy food) [+]        | 3.95      | 3.90                                  | −0.05<br>(0.50)   |
| I want to sensitize others to clothing norm issues [+]            | 3.92      | 3.49                                  | −0.43<br>(0.38)   |
| I would like to own a kosher smartphone to buy clothes online [+] | 2.79      | 3.21                                  | 0.43<br>(0.58)    |
| I would like to own a kosher smartphone to buy clothes [+]        | 2.46      | 2.77                                  | 0.30<br>(0.55)    |
| I would like to dress more casually [+]                           | 3.07      | 3.01                                  | −0.06<br>(0.49)   |
| I would like to drive [+]                                         | 4.28      | 4.34                                  | 0.06<br>(0.49)    |
| I would like to volunteer outside of the community [+]            | 4.05      | 3.21                                  | −0.84<br>(0.54)   |
| Note:                                                             |           | †p<0.1; *p<0.05; **p<0.01; ***p<0.001 |                   |

Note: Each row of Table C.1 includes: the predicted average response for the participants in the Siddur and Dress conditions, the effect size ( $\tau$ ) of being treated (TOT), and the standard error of the effect size of being treated. Statistical significance is indicated using stars.

Table C.2: Effects wearing the plain dress on possible psychological mechanisms.

| Item [expected direction of average treatment effect $\tau$ ]   | <i>Condition</i> |       | $\tau$                       |
|-----------------------------------------------------------------|------------------|-------|------------------------------|
|                                                                 | Siddur           | Dress |                              |
| <b>Perception of the self as deviant</b> [+]                    |                  |       |                              |
| I always try to fit (reversed) [+]                              | 3.93             | 4.12  | 0.19<br>(0.34)               |
| I feel strong enough to deviate [+]                             | 4.77             | 4.89  | 0.12<br>(0.31)               |
| <b>Perception of costs of violating social norms</b> [-]        |                  |       |                              |
| Easy for me to deviate (reversed) [-]                           | 3.67             | 3.55  | -0.12<br>(0.39)              |
| Deviating makes me feel too bad [-]                             | 3.01             | 2.33  | -0.68 <sup>†</sup><br>(0.41) |
| <b>Perception of high-end clothing</b> [∅]                      |                  |       |                              |
| Women believe that wearing expensive clothes is important [∅]   | 5.11             | 4.30  | -0.81*<br>(0.38)             |
| Women believe that close others won't like them otherwise [∅]   | 3.29             | 3.39  | 0.09<br>(0.42)               |
| I believe that wearing expensive clothes is important [∅]       | 3.09             | 2.74  | -0.35<br>(0.46)              |
| I believe that close others won't like them otherwise [∅]       | 2.04             | 2.03  | -0.01<br>(0.32)              |
| I want to purchase and wear cheaper clothes [∅]                 | 4.03             | 3.90  | -0.13<br>(0.53)              |
| <b>Sensitivity to social issues</b> [+]                         |                  |       |                              |
| I hope things will change [+]                                   | 4.90             | 3.97  | -0.93 <sup>†</sup><br>(0.51) |
| I want to volunteer for Ultra-Orthodox Jewish organizations [+] | 5.71             | 5.14  | -0.57<br>(0.38)              |
| <i>Note:</i>                                                    |                  |       |                              |
| <sup>†</sup> p<0.1; *p<0.05; **p<0.01; ***p<0.001               |                  |       |                              |

Note: Each row of Table C.2 includes: the predicted average response for the participants in the Siddur and Dress conditions, the effect size ( $\tau$ ) of being treated (TOT), and the standard error of the effect size of being treated. Statistical significance is indicated using stars.