NOT MY THING

NOT MY THING

How We Teach Cruelty to Cure It,
from the Armband to the Diversity Seminar

DAVID BOLES

David Boles Books · 2026

Not My Thing: How We Teach Cruelty to Cure It, from the Armband to the Diversity Seminar

Copyright © 2026 by David Boles

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the author, except for brief quotations embodied in critical articles and reviews.

This is a work of nonfiction. The cases it describes are drawn from the public record, from published scholarship, and from the author's own experience. Where events are reconstructed from memory, they are presented as memory. The interpretations and conclusions are the author's own.

Published by David Boles Books

BolesBooks.com

First edition, 2026

For Janna Sweenie, and for the ones who would not play, and paid for it.

If there is any fixed star in our constitutional constellation, it is that no official, high or petty, can prescribe what shall be orthodox in politics, nationalism, religion, or other matters of opinion, or force citizens to confess by word or act their faith therein.

Robert H. Jackson, *West Virginia State Board of Education v. Barnette* (1943)

The only obligation which I have a right to assume is to do at any time what I think right.

Henry David Thoreau, “Civil Disobedience” (1849)

Power concedes nothing without a demand. It never did and it never will.

Frederick Douglass, “West India Emancipation” (1857)

Contents

Overture: Not My Thing 1

Part One: The Specimen 7

Chapter One: The Armbands 8

Chapter Two: April 5, 1968 19

Chapter Three: The Anatomy of the Machine 30

Part Two: The Family 42

Chapter Four: StarPower 43

Chapter Five: The Privilege Walk and the Tunnel 55

Chapter Six: The Auction 67

Part Three: The Law 78

Chapter Seven: The Pedagogy of the Oppressor 79

Chapter Eight: Iatrogenic Empathy 91

Chapter Nine: The Absence of the Subject 103

Part Four: The Industry 115

Chapter Ten: From Stunt to Standard 116

Chapter Eleven: The Rented Conscience 128

Part Five: The Science 137

Chapter Twelve: What the Evidence Says 138

Chapter Thirteen: The Collapse of the Dramatic Paradigm 152

Part Six: The Cure and the Refusal 166

Chapter Fourteen: The Vulnerable Days 167

Chapter Fifteen: What Works 181

Chapter Sixteen: Not My Thing 194

Coda: After the Armband 203

Appendix: Where to Start 206

Acknowledgments 209

A Note on Method and Sources 210

Notes on Sources 211

List of Terms 213

Bibliography 215

Questions for Discussion 220
About the Author 223
Also by the Author 224

OVERTURE

Not My Thing

In the spring of 1973, in a public school in Lincoln, Nebraska, a teacher handed me a paper armband and told me to wear it for the rest of the day. Mine was blue, and it carried a single printed letter, an S. The children with brown eyes wore brown bands lettered M. No one explained the letters to a room of eight-year-olds. I have spent the rest of my life arriving at what they had to mean. M for Master. S for Slave. The letters were large enough that any adult in the building could read a child's assigned rank from across a hallway, and they traded sides the next morning when the roles reversed.

I remember the strangeness of it more than the fear, the way a morning that had begun like any other turned, without warning, into something with rules I had not agreed to. Children I had played with at recess were suddenly above me or beneath me by a feature of their faces none of us had chosen. The teacher I trusted was the one doing it, which is the detail that has stayed with me longest, because the cruelty came from the calm adult at the front of the room rather than from a stranger or a bully, the one who smiled while she sorted us and called the sorting a lesson. I did not have the words then for what was wrong. I only knew that the ground had shifted and that the person who should have steadied it was the one tilting the floor.

We were never told why we had been sorted. We were lined up and sent to one side of the room or the other, and the reason was withheld the way the rules of a rigged game are always withheld from the people who are about to lose it. The brown bands could give orders. The blue bands had to obey. When a brown-eyed boy ordered me to fetch him a drink of water, I stayed in my seat, and when he warned me that he would have to whip me, I invited him to try. He went to the teacher instead, and the teacher announced that I would be punished with the shame of history, a phrase that meant nothing to me then and means too much to me now. Around me the other blue bands brayed like donkeys and ground through forced jumping jacks while the school looked on, because the exercise did not stay in the classroom. It followed us into the cafeteria and onto the playground, where the older grades watched us wear our ranks. The humiliation was

built to have an audience.

Those adults believed they were teaching us about racism. The sentence the teacher used, the one I have never forgotten, was that now we knew how it felt to be Black. There were, at most, one or two Black children in the whole school, and none in that room, which was the precondition the lesson required and never admitted. A single Black child standing there would have collapsed the pretense in an instant, because for that child the game would have stopped being a game. The exercise needed an audience that could be told what to think and what to feel and would comply, and we complied.

The presumption in that sentence has only grown larger to me with the years. A roomful of white children was told that a morning in a colored armband had taught them a life, and the teacher believed it, and the belief was the whole error in miniature. A costume is not a life. The bad afternoon a child knows will end at three o'clock bears no honest resemblance to a condition that does not end, that was not chosen, and that no one in that room would carry home. The lesson borrowed someone else's experience without their leave and without their presence, and then congratulated the borrowers on their new understanding.

It did not survive contact with anything real. The moment we carried our armbands into the hall, the fourth, fifth, and sixth graders looked at every third grader as a stupid little kid and treated us that way, M and S alike. What the teachers had built with such care meant nothing the instant it left the one room where an adult stood over it, because a real pecking order already ran the building and it did not care about the color of our eyes. The manufactured cruelty was swallowed whole by the ordinary cruelty of older children, which the exercise had neither created nor addressed.

Then the parents found out. Every family in that neighborhood was white, working-class, and conservative, most of the fathers on the line in the city's plants, and not one of them had been asked before their children were labeled and marched through it. They were furious, and their fury proved the lesson had failed, because the exercise had just shown them through their own sons and daughters what it feels like to be ranked at the bottom by an accident of birth, and their answer was to demand that white children be spared it, voiced by at least one father in the exact slur the lesson was supposed to make unsayable. The protest did not stop the school. On schedule the next morning, the roles reversed, and the second day ran like the first. The teachers tried it on us once, in 1973, and never again, and I have come to believe the parents are the reason. The thing that killed a lesson sold as a cure for racism was racism itself.

I was the only child in my classroom punished for any of it, because I refused both roles. I would not obey on the first day and I would not give the orders on the second, and when the teacher pressed me to command the others so I could feel my power, I told her it was not for me, not my thing. My mother taught fourth grade in the same building. Embarrassed in front of her colleagues, she took my television for a week to teach me

what it feels like to lose a privilege, never once seeing that the stripping of privileges was the exact engine of the exercise she was defending. With the set gone, I read instead, and I have never run out of that particular freedom since.

I have carried this for half a century as a private grievance, the way a person carries an old injustice that no one else seems to remember. What changed, recently, was the recognition that my classroom had been one specimen of a species, and that the species has a long and respectable-looking history. The teacher in Lincoln was copying a famous original, an exercise built in 1968 by an Iowa schoolteacher named Jane Elliott the morning after Martin Luther King Jr. was murdered, an exercise that ran for more than a decade on the children of one farm town, that won a documentary and a place on late-night television, and that grew into a model copied in a thousand classrooms and, eventually, in the training rooms of corporations and governments around the world.

The recognition came as a strange relief and a deeper alarm at once. Part of it was relief, in learning that the thing done to me had a name and a literature, that I had not imagined its wrongness or nursed a child's exaggeration for decades. The alarm was in seeing how respectable the genre had made itself, how a stunt that failed in my classroom had been filmed and praised and franchised and taught to teachers as a model, how the same flawed machine had been installed in the institutions that train the country's workers and soldiers and civil servants. My private grievance turned out to be a public fact, and the public fact was worse than the grievance, because it had escaped the single room and gone to work on millions.

Let me be plain about how I am using that memory, because a reader is right to be wary of it. It is one man's recollection of a single morning fifty-three years gone, reconstructed by the same fallible machinery that reconstructs everyone's childhood, and I would not ask anyone to convict a method on the strength of it. I do not. The memory is the goad and the starting point, the thing that sent me looking, and it carries none of the weight of proof. Everything that decides the case in the pages ahead is drawn from the documented record, the exercise Jane Elliott ran in front of cameras, the games published and sold by name, the trainings measured in peer-reviewed trials, the scandals recorded in court filings and news reports a reader can check for himself. Where I lean on my own classroom I say so, and I lean lightly, for the feel of the thing and the shape of the question. The verdict rests on what can be verified, and I have built it so that a reader who doubts every word of my own story can still follow the evidence to the same end and find it sound.

Elliott's exercise is only the most famous member of a large family. The same machine runs through the simulation games that swept American schools at the end of the 1960s, through the privilege walks and poverty role-plays and oppression tunnels that followed, through the documented scandals in which teachers staged mock slave auctions and Gestapo round-ups and assigned real Black and Jewish children to play the

despised, through the scared-straight prison tours sold to frightened parents, and through the mandatory sensitivity trainings that turned the whole approach into a profession. The family has a name, and this book gives it one. Call it coercive moral pedagogy, the practice of trying to manufacture empathy by manufacturing suffering.

A word on empathy, since the whole quarrel turns on it. The thing has two faces that ordinary speech runs together. One is the sharing of a feeling, the catch in the throat at another person's pain, involuntary and quick and often mistaken for understanding. The other is the harder accomplishment, the accurate grasp of what another life is actually like, built from attention and testimony and time, the kind that can correct a person's conduct because it has corrected his picture of the world. The dramatic exercise targets the first and calls it the second. It manufactures a strong feeling in an afternoon and presents the feeling as knowledge, when the feeling a child has after a morning of assigned cruelty is a fact about the morning and tells him almost nothing reliable about the life he was made to mimic. Where this book uses the word empathy without qualification, it means the second kind, the understanding that changes what a person does, and the argument throughout is that the manufactured feeling does not become that understanding and sometimes blocks the road to it.

This book makes three arguments about that practice, and each runs through every case in it.

The first is a reversal hidden in the title. In 1968, the same year Elliott ran her first exercise, Paulo Freire published *Pedagogy of the Oppressed* and warned that the oppressed cannot be freed with the oppressor's tools, because the tools carry the oppressor's logic into any hands that pick them up. Coercive moral pedagogy is that warning ignored. It assigns a caste by an accident of birth, enforces it through an authority no one can refuse, rewards the cruelty of the favored, and stages the result for spectators. Such is the structure of the thing it claims to oppose. It is a pedagogy of the oppressor, and the cause it serves does not change what it teaches, because a method teaches its own lesson, and the lesson is obedience to cruelty.

This is the claim that gives the book its title and its sharpest edge, because it holds even when the cause is just. A method is not a neutral container that takes the morality of whatever fills it. It carries its own instruction in its form, and the form of this one, caste and command and licensed cruelty and a watching crowd, instructs in submission no matter what slogan rides on top. A good end does not launder a tyrant's means, and a child taught to obey a rigged hierarchy for an afternoon has been taught the lesson the hierarchy exists to teach, whichever side of the armband he wears.

The second argument borrows a word from medicine. An injury caused by the treatment is iatrogenic, and the harsher forms of this pedagogy are iatrogenic, the cure that produces the disease. The failures sort into two kinds. Some programs merely waste the hour, changing nothing they were built to change, and the drug curriculum that

moved no needle is the type. Others actively injure, and the deterrence tour that raised the crime it meant to prevent is the type. Evidence for both is in the chapters ahead, drawn from controlled trials, with the glow of the moving afternoon set aside. The distinction matters because it sets the ceiling on the claim, that this kind of intervention does not reliably deliver its good and in its harshest forms collects a harm on top, which is the charge a doctor would have to answer and a pedagogy does too.

The word matters because it changes the question one is allowed to ask. A doctor whose treatment sickens the patient cannot defend himself by pointing to the purity of his intentions, and neither can a pedagogy. Whether the people running the exercise meant well is never the right question. What matters is what the exercise does to the people inside it, measured honestly and after the glow of the moving afternoon has faded, and the measured answer, again and again, is failure, and in the harshest versions, harm.

The third argument is the one my own classroom handed me. The simulation works only in sealed conditions, with the subject absent or silenced, and it dies the moment it meets the real world. My exercise needed a room with no Black child in it and a teacher standing guard, and it dissolved in the hallway against a hierarchy it could not touch. Raise that to a law and it convicts the whole genre. A ritual that requires the absence of the people it claims to be about teaches nothing about those people, and a lesson that cannot survive its own subject was never a lesson at all, only a performance staged for the comfort of the people running it.

I have come to think of this as the tell, the single feature that exposes the whole genre. Ask of any such exercise where its supposed subject is, and the answer convicts it. The slave is a child in a paper band; the enslaved are nowhere. A privilege walk turns the poor into a station on a course; the poor themselves are not consulted. The ritual requires their absence because their presence would break it, and a lesson that can only run when the people it concerns are kept out of the room is using them under the name of teaching.

I am writing this in the middle of an argument the whole country is having about diversity work, an argument conducted almost entirely in bad faith from both directions. On one side, the goal of equality is under attack from people who use the failures of the method as cover for opposing the aim. On the other, the method is defended by people who will not concede a single one of its failures, and so defend its worst expressions in order to protect a goal that deserves better tools. I am for the goal and against the method, and I see no tension between those two positions. Holding both is the only honest place to stand. The history of how free societies are broken shows a recognizable technique, the conscription of virtue, the demand that people be made to feel the right things by force, and it points toward the largest claim in this book, that a population habituated to coercion for a good reason is tilted toward accepting it for a bad one. I

make that claim as an argued tendency rather than a proven law, and I make it in full only in the chapters ahead, where it belongs. Here I only promise it.

This comes from history, not from rhetoric. I have read how decent countries are talked into indecent things, and the pattern tends to run the same way at the start, a good cause invoked to justify a small surrender of the rule against forcing belief. The surrender is easy to grant when the cause is one you favor and the people forced are people you do not. That is the trap. The machinery cares nothing for the cause and everything for the habit, and a habit of compelled feeling, once built, waits to be turned to whatever end the next hand directs. I am against the method on the left and the right alike, and I am against it for the sake of the goal, not in spite of it, because the freedom the method spends is the freedom without which no goal is safe.

What follows is the case, built case by case and argument by argument, and held to a single standard throughout, that a claim about harm must be shown and not merely felt. I have tried to cite no source in bad faith, to defend the goal on every page even while dismantling the method, and to leave the reader free to disagree. The book asks for nothing it does not earn.

Three distinctions decide how hard this book judges any given case, and naming them now will save argument later. The first is the age of the subject. An adult can consent to be unsettled in a way a captive eight-year-old cannot, and the use of these methods on children, who can neither refuse nor consent, draws the book’s sharpest objection. The second is whether the thing is compelled or chosen. A workshop a person elects to attend is a different object from one his employer or his teacher forces on him, because compulsion is the engine of the backlash the later chapters document. The third is whether the lesson rests on a real encounter or on a staged stand-in. A survivor’s testimony, a memoir, an afternoon spent in actual contact with the people in question, these put the subject in the room as herself, where the simulation puts a proxy there instead, and the proxy is where the method goes wrong. None of this condemns experience as a teacher, only the coerced experience, performed on the unconsenting and built around a substitute for the people it claims to serve.

This book is an indictment of a method and a defense of a goal, and it is built to leave both camps uneasy, because that discomfort is the only proof that the attack was honest. It begins where I began, in a small room with a paper armband and an adult who had mistaken cruelty for instruction. And it ends where I ended, with the smallest and oldest act of freedom there is, a child who looked at the person in charge and declined to play the rigged game, whatever it cost him. That refusal is the thing the entire machine is designed to prevent. It is also the only thing that has ever stopped it.

PART ONE

The Specimen

CHAPTER ONE

The Armbands

The facts in this chapter rest on a single witness, who was eight years old when he gathered them, and the witness is me. Before you weigh anything I build on those facts in the rest of this book, you should know exactly how far the witness can be trusted, because a man who asks you to convict a method on the strength of his testimony owes you the limits of that testimony first.

Here is what I cannot give you. I cannot give you the date, beyond the spring of 1973. I cannot give you my teacher's name, though I sat in her classroom every weekday for a year and was afraid of her for part of it. The conversations I cannot reconstruct word for word the way a novelist would, because I was a child and no one in that room was keeping a transcript. Memory does not store a school day as a recording. It stores a few sharp objects and lets the rest dissolve, and I would rather show you the few sharp objects honestly than fill the dissolved parts with invention and pass the whole thing off as total recall.

Here is what I can give you, and what I would swear to under oath: the armbands; the two letters printed on them; the cup of water; the braying; the jumping jacks; the reversal the next morning; the word the fathers used when they came for the school. These have not faded in half a century, and they have not faded because they were not learned in the ordinary way, the way a multiplication table is learned and then half forgotten. They were burned in. I remember them the way a person remembers a burn, which is to say on the skin, without effort and without wanting to.

I am setting them down at the front of the book deliberately. Every argument I make in the chapters ahead rests on a plain claim, that this kind of thing happened, that it was treated as ordinary and even enlightened, and that the adults who ran it believed they were doing good. I am the proof of that claim. A witness owes precision before he is owed belief, so before I take the machine apart, before I name its parts and follow its descendants into the seminar room and the prison tour and the corporate training center, I want the testimony down plainly, with the gaps marked where the gaps are.

It began with a sorting, and the sorting arrived without a reason, which I have since learned is the opening move in every version of this exercise. We were told to line up. We were sent to one side of the room or the other by a rule withheld from us, the way the terms of a rigged game are always withheld from the people about to lose it. Standing in that line, I did not know what decided which way I went. I learned only later that the boundary ran along the color of our eyes, that the brown-eyed children had been gathered into one group and the blue-eyed and the rest of us into another, and that a grown woman had built an entire school day on a trait none of us had chosen and none of us could alter by lunchtime.

There is a particular dread in being classified without being told the rule, and I felt it before I had a word for it. A child can absorb a punishment he understands. What unsettles him is the machinery that operates on him for reasons kept out of his reach, the sense of a hand sorting him toward an outcome he cannot see coming and cannot argue against. That dread was the first thing the exercise produced in me, before any armband was handed out, and I notice now that the designers of these lessons treat that dread as a feature. The confusion is the method beginning its work, not the method breaking down.

Then came the armbands. Mine was blue paper, and it carried a single printed capital letter, an S. The children on the other side of the room wore brown bands lettered M. No adult explained the letters, not that morning and not afterward, and I have spent a good part of my life arriving at the only reading that fits the facts. M for Master. S for Slave. I am marking that as an inference, because honesty requires it. No one in that room said the words master and slave to me. The letters were never decoded aloud. What I know for certain is narrower and worse. One band gave its wearer the right to command. The other band carried the duty to obey. The letters were printed large enough that any adult in the building could read a child’s assigned rank from across a corridor. And a boy of eight, looking down at the S on his own arm and across at the M on another child’s, understood without being told which of the two he had been made. A child does not need the vocabulary of a system to feel its weight on him. He feels the weight first, and finds the words for it decades later, which is part of what I am doing now.

The rules were simple, and the simplicity was the point. Brown bands could give orders. Blue bands had to take them. A brown-eyed child could send a blue-eyed child to the back of a line, could correct him, could report him, could decide in small ways how his morning went, and behind the whole arrangement stood the teacher as its guarantor, which meant there was no appeal. I want to rest on that word, appeal, because it carries more than it appears to. The cruelty of the day was that an adult had removed the ordinary brakes, not that children can be unkind to one another, which any eight-year-old already knows from the schoolyard. She had told one group of children that its unkindness was now licensed and the other group that its suffering was now required, and she had done it inside the one room where a child is supposed to be safe

from precisely that. A playground is lawless on every side, and a child knows to watch his back there. This was lawlessness with a teacher's signature beneath it, in the place a child is taught to trust, and that signature is what separated it from an ordinary bad day among kids.

I lasted about as long as it took for someone to test the rule on me. A brown-eyed boy, bigger than I was, told me to fetch him a drink of water. I stayed in my seat. He told me a second time, and added that if I would not, he would have to whip me. I remember the word whip with complete clarity, because it was so obviously borrowed from somewhere, the sort of word a child grabs when he is performing a part he has just been handed and is testing how far the part extends. I told him to try. He did not try. He went to the teacher, which tells you exactly where the power in that room sat, and the teacher did not discipline the boy for threatening to whip a classmate. She disciplined me, for refusing to be whipped.

The punishment is the part I would ask you to carry through the rest of this book, because it is the hinge of the whole argument. I was told I would be made to feel the shame of history. At eight I did not know what those words meant, and I am not sure the teacher could have defined them cleanly if pressed, but the phrase has never left me, and its meaning has only grown heavier with age. The shame of history. As though history were a debt that one small boy could be made to settle in a Nebraska classroom by being humiliated in front of his friends. Humiliated is the precise word, not a softer one. While the exercise ran, the blue bands around me were made to bray like donkeys and to grind through forced jumping jacks, a row of seven- and eight-year-olds performing their own degradation on command while the rest of the room looked on, and I was singled out and held up as the one who had earned an extra portion of shame by declining to perform mine.

I have spent years turning that phrase over, the shame of history, because a careless adult handed me, without meaning to, the password to the whole enterprise. I have already said that history is no debt a single child can be made to settle. The deeper trouble lies in the verb the teacher reached for, in the belief that a person can be taught anything worth knowing by being made to feel ashamed. Shame transmits no understanding from one human being to another. What it deposits in the shamed child is a private wound he will spend years learning to dress, and no insight at all into the life it was supposed to explain.

The reflex behind that phrase runs through everything this book examines, so I want to name it plainly while it is still small enough to see whole. The reflex is the belief that a person can be improved by being made to suffer, that virtue is a thing you install in someone by force, and that the installer is doing the sufferer a kindness. It is an old belief, older than any classroom, with a long religious lineage in the idea that the flesh must be mortified to save the soul. The teacher at the front of my room would have been

startled to hear her exercise described as a secular flogging, because she believed she was teaching empathy, and that gap, between what she believed she was doing and what she was doing, is the gap this entire book lives inside. So far as an eight-year-old could judge, she was no monster. She had inherited a method that does its damage regardless of the character of the person running it, which is the most dangerous kind of method there is, because it recruits decent people to do harm and lets them feel righteous while they do it.

There is a darker reading of the phrase, and age has pushed me toward it. To be made to feel the shame of history is to be made a scapegoat, in the oldest sense of that word, the animal driven into the wilderness with the sins of the community on its back. I was eight, and I had refused to fetch a cup of water, and for that refusal I was to carry, on behalf of a roomful of children and a country two centuries deep in its crimes, a guilt that was never mine and could not have been discharged if it were. The exercise needed a scapegoat to complete itself, and on the first day the role fell to me, because I had stepped out of line. The lesson taught no history. It used me to stage a ritual of atonement that cost the adults nothing and showed the children only that stepping out of line is punished. A ritual of atonement that demands no change from the people performing it, and falls hardest on whoever refuses to play, has no honest claim to the word justice. What it manages is guilt, and it manages guilt by moving it onto the powerless, and it works on the comfortable precisely because it asks nothing of them.

Notice who is exempt from this arithmetic. The teacher feels the shame of history at no moment in the lesson; she administers it. Those parents who would later object feel it only when it reaches their own children; until then they are spectators. The cost is carried entirely by children, and among the children most heavily by the one who will not perform, and a moral lesson whose costs fall only on the powerless and never on the powerful has inverted the thing it set out to teach, because the moral content of history is that the powerful owe the powerless, and not the other way around.

The choreography is the evidence, which is why I am dwelling on it. A correction aimed at teaching a child something can be quiet. It can be a word in the hallway, a note carried home in a backpack, a conversation after class. What happened in that room was none of those. It was staged, and a staging needs an audience, and the audience was duly supplied, first by the other children in the class and then, as the day wore on, by the entire school. The lesson did not stay behind the classroom door. We carried our armbands into the cafeteria and out onto the playground, where children from every other room could see at a glance which of us had been marked Master and which Slave, and where the older grades watched a younger class move through its day wearing its assigned ranks like livestock at a fair. A lesson that is honestly about empathy does not require spectators. A spectacle requires them, because the watching is the product. The presence of that audience tells you what kind of event this was, and it was the second

kind.

Now I have to set down the fact that organizes everything else, the fact whose meaning escaped me until I was a grown man, and it is a fact about who was not there. The school was almost entirely white. In the whole building there were, at most, one or two Black children, and in my classroom there were none. A white teacher stood in front of a room of white children and told us, in words I have carried for half a century, that now we knew how it felt to be Black. We did not know. We could not have known. The entire performance depended for its function on the absence of a single human being who would have made it impossible to perform.

Consider what one Black child in that room would have done to the lesson. For every white child present, the armband was a costume. It could be taken off at the final bell. It would be reversed by the next morning. The teacher's standing guarantee made the whole thing temporary, a role assigned at the start of the day and lifted at the end of it. For a Black child, there would have been no final bell and no reversal, because the condition the white children were play-acting for a day was the actual architecture of that child's life on the far side of the door, and no paper band would have been required to deliver the lesson. The pretense would have shattered on contact with the one person it claimed to be teaching us about. So the exercise functioned, to whatever small degree it functioned at all, exactly because there was nobody in the room for whom it was not a pretense. Lifted from my single classroom to a general rule, this becomes one of the book's three central claims: that this entire family of lessons requires the absence, or the enforced silence, of the exact people it claims to serve. A lesson that cannot survive the presence of its own subject teaches nothing about that subject, and stages instead a performance for the comfort and the moral self-regard of the people running it.

I have been describing this absence as though it were a hole in one classroom, a local accident of enrollment. It was nothing so small. The absence was the shape of the city, and the city is part of the evidence, so let me put it on the record as plainly as I can.

Lincoln in 1973 was a white city, overwhelmingly and quietly so. I cannot hand you the exact count of Black residents the census took that year, because I have not been able to verify a figure I would swear to, and I will not invent one to make a point. What I can tell you is the larger and verifiable truth. The communities that would later bring real variety to Lincoln's schools had not yet arrived; the refugee resettlement that reshaped the city, the Vietnamese families and then so many others, began only after the fall of Saigon in 1975, two years past the spring I am describing. Before that, a white child could pass through an entire elementary education in that town and sit beside a Black classmate rarely, or never. I was such a child. My school held one or two Black children in the whole building, by my memory, and my classroom held none, and neither of those facts was unusual for the place and the year. They were ordinary, and that is the part worth sitting with.

The all-white room was the precondition for the exercise, and the room could be all white because the city was. Set those facts beside the teacher's sentence and the structure comes clear. A lesson that claims to teach white children how it feels to be Black, staged in a city arranged so that white children almost never meet a Black child, reaches across nothing. It talks to one side of a divide about the other, in the other's absence, and calls the monologue a conversation. The geography that made the lesson possible is the same geography that guaranteed it would fail, because a community can rehearse the feelings of people it has arranged never to encounter, and the rehearsal teaches it nothing it can use, since no one is on hand to correct the performance or to make it cost a thing. At eight I did not understand that the map of my city was inside the lesson. Only now do I see it for what it was, the lesson's hidden precondition, and I am certain you cannot grasp what was done in that room without first seeing the empty map it was drawn on.

There is a further turn, and it is the one that has stayed with me longest. The condition the exercise set out to teach us about, the experience of being sorted and diminished by an accident of birth, was present in Lincoln all along. It was the daily situation of the few Black families who lived there, in a city that had arranged its neighborhoods and its schools so thoroughly that their children were a rumor to mine. By 1973 the law no longer wrote that separation down. The map wrote it, into where families could afford or were steered to live, into which neighborhoods fed which schools, and a separation enforced by geography needs no signs on the doors to be complete. The lesson treated that separation as something far away, a problem in the South or in the past, to be simulated with paper bands because the real thing was supposedly beyond reach. It sat a few miles off, in the parts of the city my classmates and I were never taken, lived out by children our own age whom we could have met, and learned from, and stood beside, had any adult thought it worth arranging. No adult did. It was easier to hand us armbands. A community that will simulate a hardship it has gone to some trouble to keep at a distance has told you what the simulation is for, and what it is for is the relief of the people performing it, who get to feel they have faced a thing they have in fact spent great effort never facing at all.

None of this excuses my teacher, who was responsible for what she did in her own room. I raise the map of Lincoln because the absence in that room was not her invention and did not end at her door, and because the same absence, in other clothing, sits at the center of nearly every descendant of her exercise this book will examine. The sealed room is never only a room. It works as a smaller copy of a larger arrangement, a community that has organized itself so that the people it claims to be studying are somewhere else, available as a subject and unavailable as a presence.

The hallway proved the same point a second time, from the opposite direction, on the same afternoon. Once we stepped out of the supervised classroom and into the

general population of the school, our carefully constructed hierarchy evaporated. The older children did not care which band a third grader wore. To a sixth grader, a third grader is a small and slightly comic creature whatever the color of his eyes or the letter on his sleeve, and so the fourth and fifth and sixth graders looked at all of us, Master and Slave together, as the little kids we plainly were, and treated us with the ordinary lazy contempt of older children for younger ones. The manufactured pecking order, assembled with such care inside one room, dissolved the moment it touched the real pecking order that already governed the building. The teacher's hierarchy could not survive its first encounter with a hierarchy she had not invented that morning. Whatever the exercise was teaching, it could teach it only inside a sealed box, and the seal failed the instant we walked through a door into the world as it runs.

Those two failures, the empty room and the porous wall, belonged to the setting. The deepest failure belonged to the sentence itself, and it would have held even in a room the exercise could not survive, even with the windows thrown open onto the world. Return to what the teacher told us. Now you know how it feels to be Black. I have shown that no Black child was present to make the claim true, and that the city had been built to keep it that way. Set all of that aside for a moment. Suppose, against the facts, that the equation could have been attempted honestly. It would still have been false, because it makes a plain error about what kind of thing an experience is, and the error rewards a slow taking-apart, since the entire genre survives on no one taking it apart at all.

The claim assumes that a feeling, briefly induced under controlled conditions, is the same kind of thing as a condition lived without relief. It is not, and four differences make it not. The first is consent, or the awareness that stands in for consent. We knew it was an exercise. A child who knows the bad thing will end at the bell is having a different experience from a child who has no bell, because the knowledge that suffering is temporary is itself a comfort, and a large one, and it was a comfort no Black child in America in 1973 was issued along with his condition. The second is exit. We had a door. At the final bell the armband came off and the ordinary day resumed, and an experience you can walk out of is not the experience of a thing you cannot. The third is duration. We endured an afternoon, and the weight of a single afternoon, however sharp, is not the weight of a thousand afternoons stacked one on the next until they harden into the architecture of a life. The cumulative is a different substance from the sample, the way a paper cut is a different thing from a wound that will not close. The fourth is stakes. Our armbands cost us nothing that mattered. No child in that room lost a job, a house, a future, or his safety because of the letter on his sleeve, and the thing the armband was pretending to stand for costs people exactly those things, so a pretense with nothing at risk cannot teach the felt weight of a condition whose whole gravity is in what it puts at risk.

The defender of the exercise wants to treat those four as differences of degree, as though a longer version, with higher stakes and no way out, would shade gradually into the real thing. Run that thought to its end and it refutes itself. A version with no exit, no end, and real stakes stops being an exercise at all; it becomes the condition itself, imposed on a child by a teacher, which is a crime, and the fact that the only honest version of the simulation is a crime tells you the simulation was never going to teach what it promised. Those are not the same lesson. What separates them is a difference of kind, and differences of kind are exactly the ones that more intensity and better facilitation cannot touch, because they are not differences of amount.

I can hear the objection, because it is the reasonable one, and the defenders of these exercises have leaned on it for fifty years. Surely, the objection runs, an imperfect taste of a thing beats no taste at all, and surely a child who has felt even a counterfeit sting will be a little gentler with the people who feel the real one. It is a humane thought, and the evidence does not bear it out, which is a large claim the later chapters of this book are built to carry. For now I will say only what my own classroom showed me, which is that the counterfeit sting made no one gentler. The children who had held the upper hand grew fond of it. The children who had been pushed down grew hungry for their turn to push. And all of us learned that an adult with authority would license cruelty so long as the cruelty arrived wrapped in a lesson, which is not a small thing to learn at eight, and not a kind one.

Let me put the claim at its proper size, because I am building a book on it and it has to bear weight. Empathy can be taught; I have no quarrel with that, and the last chapters of this book describe the ways it is done. My quarrel is with the belief that it can be manufactured in an afternoon, by force, through a costume, because empathy is built the slow way, by knowing particular people and attending to particular lives across time, and any method that promises to deliver it fast and whole is advertising a shortcut to a thing that has no shortcut. The shortcut is the tell. When someone offers to install a virtue quickly and dramatically, the speed and the theater are the warning rather than the attraction, because the virtues worth having are the ones that take the long way in and cannot be hurried without being faked. My teacher believed she had found the fast road to a generous heart. What she had found was the fast road to the feeling of a generous heart, a different destination wearing the same name, and the children she drove there climbed out of the car certain they had arrived, which is the one outcome guaranteed to keep them from ever making the trip.

Here is the part that turns a harmless mistake into a harmful one, and it is the finding I would carry out of this chapter if I could keep only one. The exercise leaves the children worse than it found them. They arrive ignorant and aware of their ignorance; they go home convinced they now understand a thing they have only worn for a day. A child who knows he does not know how it feels to be Black will, if he is decent, keep his

mouth shut and his eyes open, and may one day ask, and listen, and learn. A child who has been told he now knows, by an adult he trusts, in a ceremony built to feel like revelation, walks away holding a false certificate of understanding, and that is worse than an honest blank, because the blank invites the question and the certificate forecloses it. The most lasting thing the exercise gave the children in my room had nothing to do with empathy. What it gave them was the unearned conviction that they had already done the work of feeling for others, which is the exact conviction that makes a person stop doing it. That is the iatrogenic core of the whole enterprise, the treatment that produces the disease. For now it is enough to have watched it work, once, on a single afternoon, in a single room, on me.

The next morning the roles reversed, and the reversal is the cruelest and most revealing feature of the entire design. Brown bands became slaves. Blue bands became masters. Overnight, on the teacher's instruction, the children who had been licensed to command were ordered to submit, and the children who had spent a day obeying were handed the authority, and in that one boy's vocabulary the whip along with it. Look at what the reversal proves. It proves that the cruelty answered nothing any child had done or deserved, that an adult assigned it and reassigned it on a schedule, for effect. The same child who was a master on the first day was a slave on the second, which means the contempt each group had been trained to feel for the other was understood by the teacher herself to be arbitrary and detachable, a costume she could redistribute at will. A lesson that sets out to expose the arbitrariness of cruelty by practicing arbitrary cruelty has enlisted in the thing it meant to indict.

I refused the second day as well, and I want to be exact about what I refused, because it is the act this book takes its shape from. On the first day I had refused to obey. On the second day, promoted now to master, I refused to command. As I remember it, the teacher pressed me on this, because the refusal made no sense inside the logic of her exercise. I was being offered power, the identical power that had been turned against me the day before, and the design assumed I would take it, that the lesson would finally land at the moment the formerly oppressed child discovered how good it felt to oppress in his turn. I declined the discovery. When she pushed, I told her, in the only words an eight-year-old had available, that it was not for me. Not my thing. I lacked the vocabulary to explain that I had examined both roles and found both of them contemptible, that the offer of the whip was the same insult wearing the opposite mask, a promotion in name only. What I knew, without being able to say it, was that I did not want to do to another child what had been done to me, and that the wanting was the precise thing the room was built to install in me, and that I would not let it be installed. That refusal is small, and it belonged to a child, and I have come to believe it is the most important thing in this book.

Then the parents found out, and the school collided with something it could not reverse on a schedule. Every family in that neighborhood was white, and working-class, and conservative, most of the fathers working the line in the city's plants, and not one of those families, so far as I have ever known, had been asked in advance whether their children might be labeled and ranked and walked through a manufactured day of caste. They were furious, and I want to describe the fury accurately, because it is one of the bitterest ironies I have ever lived inside. The exercise had just done to their sons and daughters, in miniature and for a day, the exact thing it claimed to be teaching them to oppose. It had ranked their children at the bottom by an accident of birth and made them feel the weight of it. The parents understood that weight at once, in their bodies, in precisely the way the lesson had intended them to. And their response was not solidarity with the people who carry that weight every day of their lives. Their response was to demand that their own white children be spared it. At least one father, by the account that reached me, made the demand in the exact word the lesson claimed to be combating, the slur itself, deployed in defense of children who had been made to play the people that slur is used to wound. I will not print the word. Its appearance in that moment is the entire point, and putting it on the page would only blunt it.

The protest did not halt the second day. By the time the parents organized, the reversal had already run and the machine had completed its cycle. What the protest accomplished, I have come to believe, was to end the experiment permanently. The school ran that exercise once, in the spring of 1973, and never again, and when I lay out the facts as I hold them, the parents are the only force on the scene with both the standing and the motive to have shut it down. So here is the conclusion I have carried for half a century, and I offer it as the darkest finding in my own small case before the book opens out onto larger ones. The thing that killed a lesson sold as a cure for racism was racism itself. No argument about the method, its cruelty, or its evidence defeated the exercise. It fell because it briefly inconvenienced white children, and a community willing to tolerate the genuine hierarchy without a word of complaint would not tolerate a two-day imitation of it aimed at their own. That is no victory of conscience. The opposite is closer to the truth, and it ought to disturb anyone who believes these exercises accomplish the good their designers promise.

There is one more witness I have to call, and calling her costs me something, because she was my mother. She taught fourth grade in the same building. When word of my refusals reached her, and reached her colleagues, she was embarrassed, a teacher whose own son had made a scene in a sister classroom down the hall, and she punished me for it at home. She took away my television for a week. Her stated reason, and I believe she meant it sincerely, was to teach me what it feels like to lose a privilege, to have a thing you assume is yours simply removed by an authority you cannot argue with. I have turned that punishment over in my mind for half a century, and what stays

with me is not the loss of the television, which I barely missed. What stays with me is that my mother could not see, and so far as I know never saw, that stripping a privilege in order to teach a lesson about privilege was the identical mechanism of the exercise she was, in that moment, defending. The school had taught caste by force and called it empathy. My mother punished my refusal of caste by force and called it discipline. The underlying logic was the same in both rooms, and it was invisible to her, because it is invisible to nearly everyone, and that invisibility is a large part of why this book has to be written. With the television gone, I did the one thing left to a boy alone in a quiet house. I read. I have not run out of that particular freedom since, and I have come to suspect that the punishment handed me the single tool I would need, many years on, to take the whole apparatus apart.

That is the testimony. One classroom, one school, one town, one spring half a century gone, reconstructed as honestly as memory allows and flagged at every seam where memory gives out. For most of my life I believed it had been a local accident, a single strange day produced by a single overreaching teacher, the kind of thing that happens once in a building and is quietly never mentioned again. I was wrong, and the scale of how wrong I was is the subject of everything that follows. The teacher in Lincoln had invented none of it. She was copying, with the assurance of someone working from an approved model, an exercise built five years earlier in another small town by another schoolteacher, an exercise that by 1973 had already been filmed for a national television audience, written up in a book, praised in the press, and recommended to teachers across the country as a humane and daring way to teach children about prejudice. To understand my classroom, you have to walk back into hers. You have to go back to the morning after Martin Luther King Jr. was killed, to a third-grade room in Riceville, Iowa, and to the woman who built the machine the rest of us would inherit.

CHAPTER TWO

April 5, 1968

The machine that reached me in Lincoln in 1973 was built five years earlier, two hundred miles to the east, by a woman who believed she was doing the bravest and most decent thing of her life, and who may have been right about the bravery. Her name was Jane Elliott. She taught third grade in Riceville, Iowa, a small, white, churchgoing farm town in the northeast corner of the state, the kind of place where drawing attention to yourself was treated as a minor sin. On the evening of April 4, 1968, she was at home when the news came that Martin Luther King had been shot dead in Memphis. What stayed with her, and what she repeated for the rest of her life, was a moment in the television coverage when a white reporter turned to a Black man and asked, in so many words, who would keep his people under control now that their leader was gone. The question sickened her. It carried, inside a single sentence, the assumption that Black Americans were a problem to be managed rather than a people to be grieved with, and she resolved that night that her students would not grow up inside that assumption without a fight.

I want to begin with that resolution, and to be fair to it, because the argument of this book depends on holding two things at once that most people prefer to take one at a time. Elliott's goal was right. She looked at a country that had just murdered its most eloquent advocate of nonviolence, and she refused to let a roomful of white children inherit the casual contempt she heard on her television and in her own teachers' lounge. That refusal took real courage in Riceville, a town with no apparent appetite for the subject, and she paid for it across the rest of her career. Nothing in the pages ahead is meant to soften any of that. The argument is narrower and harder: that the method she reached for, in the sincerity of her grief, was the method of the thing she hated, and that the cause did not redeem it. No villain stands at the origin of coercive moral pedagogy, only a good woman building a bad machine for the best reason there is, and that is exactly what makes it the right place to start.

To understand the reach she felt for, it helps to stand for a moment where she stood in the first week of April 1968. King's murder had torn the country open. More than a

hundred cities were burning or about to burn, the National Guard was in the streets of Washington within sight of the Capitol, and the evening news carried images of a grief and a fury that most of white America had spent a decade declining to look at. Riceville sat far from any of it, insulated by distance and by sameness, a town where a child could grow up without ever meeting a Black person and absorb the country's racial assumptions secondhand, from television and from adults who had absorbed them the same way. Elliott understood the insulation as the problem. Her students felt nothing in particular about King's death because they had no one to feel it for, no face to attach to the abstraction of race, and she could not import the missing people into her classroom. So she reached for a substitute. If she could not bring her children into contact with the reality of being despised, she would manufacture a reality of being despised and run them through it, and the decision to manufacture what she could not supply is the decision this entire book is about. It is the moment the substitute is chosen over the thing itself, and once that choice is made, everything that comes after follows from it.

The next morning, Friday, April 5, 1968, the first of her students through the door was a boy who had heard the news, and he put the question to her plainly, the way children put the questions adults have learned to avoid, asking why they had shot that King. The rest of the class filtered in, twenty-eight white children in a town that had almost no others, and Elliott set aside the arithmetic she had planned. She asked them whether they thought they understood how it felt to be a Black child in America, and when some said they did, she pressed until they conceded they could not, and then she made them an offer. Would they like to find out. The children, trusting her the way third graders trust the adult who runs their day, said yes.

That yes deserves a harder look than it usually gets, because the whole exercise rests on it and it cannot bear the weight. A child of eight cannot consent to be a research subject, and consent extracted by the one adult who controls his lunch, his recess, and his standing among his friends amounts to compliance wearing consent's clothing. The children did not know what they were agreeing to, and neither did Elliott, who had built the thing the night before and was about to run it on live subjects with no notion of where it would lead. Strip away the good intentions and look at the structure on its own terms. An adult in a position of total authority proposed to inflict an unknown experience on twenty-eight children for her own pedagogical purposes, secured their agreement by asking a question they were in no position to refuse, and kept no provision for stopping if the experience went wrong. We have built an entire apparatus of ethics review, in the decades since, precisely to prevent adults from doing to subjects what Elliott did to her class that Friday, and we built it because the road to a great deal of harm has been paved by people as sincere as she was. The sincerity is not in question. The sincerity is the hazard, because it is what lets a person skip the questions a colder operator might have been forced to answer.

What she built over the next two days had a structure I can describe precisely, because she described it many times and the cameras later recorded a version of it. She divided the class by the color of their eyes. One group she raised up and the other she pushed down, and the next school day she swapped them, so that every child served a turn as a master and a turn as an inferior. The sources disagree on which color sat on top first, and I will not pretend to settle a detail the record leaves unsettled; what matters is the shape. Whichever group sat on top received the privileges a child counts as wealth: first in line, longer recess, second helpings, the front of the room, the right to drink straight from the fountain. The inferior group wore collars of cloth around their necks so their rank could be read at a glance, sat at the back, lost five minutes of every recess, drank from paper cups, and were forbidden to play with the children above them.

Then Elliott did the thing I find hardest to forgive, harder than the collars and the segregated recess, because it corrupted the one faculty a teacher is supposed to protect. She lied to the children about science. To make the ranking stick, she told them that melanin, the pigment that darkens the eyes, also sharpens the mind, so that the brown-eyed children, carrying more of it, were the natural superiors when their turn came. She wrote the word on the board. She gave a chemical a causal power it does not have, and she did it deliberately, because she understood that children raised on the space program would swallow a hierarchy more readily if it arrived in a lab coat. Consider what that maneuver is. The claim that a heritable biological trait determines intelligence, and sorts humanity into the abler and the lesser, is the foundational claim of scientific racism, the exact pseudo-science that propped up segregation and the eugenics movement and the Nazi racial laws. To teach children to reject that idea, Elliott taught them to believe it, swapping eye color for skin color but keeping the rotten logic intact. She trained them to accept that biology ranks minds, and merely pointed the ranking in a new direction for two days. Tell a child on Friday that melanin makes you smarter, and you have taught him the premise of racism, in a form he trusts because his teacher signed it, and asked him to apply it. That lesson runs in the opposite direction from a defense against the idea. The lie was offered as a tool against bigotry, and it is bigotry’s first principle.

The marker deserves a moment too, because it is the genre’s signature and because it traveled. In Riceville the badge of caste was a cloth collar. By the time the same exercise reached my classroom in Lincoln it had become a paper armband, and in the corporate workshops Elliott would later run it became a lanyard or a colored dot. The material changes; the function never does. The badge exists so that the lower caste can be identified across a room without a word spoken, which is the entire purpose of a badge, and the purpose holds whether the cloth is sewn in Iowa or the paper is cut in Nebraska. It invites a comparison that observers have drawn since the exercise began, to the yellow star the Nazis forced onto Jews, and the comparison is apt, which should have been the

warning rather than the selling point. A method that requires you to brand a child so the others can find and punish him has told you what kind of method it is.

Then came the part Elliott considered her proof, and that I consider the heart of the indictment. The children turned. By her own account it took about fifteen minutes to convert a room of ordinary, friendly third graders into two camps, one swaggering and one shrinking. The children in the favored group, handed a license they had done nothing to earn, used it. They bossed, corrected, and informed on the children below them, and they grew, in her telling, vicious with a speed that frightened her. One child suggested she keep a ruler on hand in case one of the lower group stepped out of line, the same impulse, reached for by an eight-year-old in Iowa, that produced the threatened whipping in my own classroom five years later. The favored children were not bad children. They acted on a permission the one unquestioned authority in the room had handed them, and they believed her, because that is what children do with what a trusted teacher tells them.

Be precise about where the children's cruelty came from, because the usual telling makes it sound spontaneous, as though contempt simply rose out of them once the lid came off. The children did not generate it on their own. Elliott modeled it for them, standing at the front of the room and performing the disdain she wanted the favored group to copy, narrating the failings of the lower caste in the flat, certain voice of an adult reciting facts. The children below were lazy, she told the room. They forgot what they had been taught. Give them something good and they would wreck it. She put leading questions to the favored children about the deficiencies of the others and took their shouted answers as confirmation, building in real time the loop every prejudice runs on, in which the authority asserts the stereotype, the favored repeat it, and the repetition is counted as proof. The cruelty the children produced in fifteen minutes came from a script. She wrote it and they learned it, line by line, because she was the teacher and the teacher was reading it as truth.

That detail carries the whole argument, because it locates the engine of the exercise exactly where this book says it sits, in the authority of the adult running it. Remove Elliott's narration and the eye-color division is nothing, two sets of children who happen to have different colored eyes and no reason on earth to despise one another. Add her narration, delivered with the full weight of the one adult they were required to believe, and the despising arrives inside the hour. The exercise revealed no cruelty lying latent in children and waiting for a trigger. It demonstrated something colder, that an adult with sufficient authority can install cruelty in children who trust her, on a subject of her own choosing, in the time it takes to mark a stack of phonics cards. That is a finding about the power of grown-ups over the young, and a frightening one, and it runs in the opposite direction from the finding Elliott believed she had made. She thought she had shown her students how prejudice feels. What she had shown, to anyone watching

closely, was how prejudice is built, and how few minutes it takes when the builder is trusted and the materials are children.

The lower group, meanwhile, came apart. The collared children went quiet and clumsy. They stumbled over reading drills they had handled easily the day before, and some of them said, in so many words, that they could not think with the collars on. Elliott watched a stack of phonics cards that a child had raced through as a superior take far longer when the same child wore the badge of the inferior, and she drew from it the lesson that would later make her famous, that the bare assignment of low status degrades the performance of the person assigned to it. She was onto something real, something psychologists would eventually study under the name of stereotype threat. Here is the trouble. To demonstrate that branding a child as inferior makes the child falter, she branded children as inferior and watched them falter. The demonstration and the harm were a single act. No method here taught the children about the damage of contempt. She inflicted the damage and pointed at it.

This is the moment to be clear about what the exercise proved, because Elliott and her admirers have read it one way for half a century, and the reading does not hold. The speed of the cruelty was offered as evidence that the lesson worked, that the children now understood prejudice from the inside. The speed of the cruelty proved something colder. It proved how easily an authority can manufacture cruelty in children who trust her, how thin the membrane is between a kind classroom and a vicious one once an adult decides to dissolve it. What the children learned in those fifteen minutes had nothing to do with empathy for the despised. They learned, and learned fast, how to despise on command, how good it can feel to stand on someone, and how an adult with power will permit the standing and even reward it. Those are the lessons of the oppressor, taught with total efficiency, and the noble aim they were taught in the service of did not change what landed in the children. A child who has just discovered the pleasure of the boot has learned only that the boot is available to him. He has tried it on and found that it fits.

On the second school day she reversed the roles, and the reversal is the feature she was proudest of and the one that gives the most away. The group that had suffered was lifted up, and the group that had ruled was brought down and collared. Elliott reported that the formerly oppressed children, now in charge, were gentler with their old tormentors than they themselves had been treated, because they remembered the sting and did not want to inflict it. She took this as the exercise’s redemption, the proof that having suffered teaches mercy. Some of them may well have been gentler; childhood is not uniform, and memory bends toward the version that flatters the teller. Set the question of degree aside, because the reversal proves a harder thing regardless of what the children did with their turn. It proves that the cruelty was never earned and never deserved, that an adult assigned it on a schedule and revoked it on a schedule, a costume handed out on Friday and redistributed on Monday. A hierarchy that the person who

built it can reverse overnight describes nothing about the world. It advertises the power of whoever holds the clipboard, and the children, whatever their manners on the second day, had learned that lesson too, that the line between master and servant is drawn by whoever runs the room and can be redrawn at will.

The claim that the reversal teaches mercy deserves a harder look, because it is the moral heart of Elliott's defense, and it does not survive one. Grant, for the sake of argument, that the formerly collared children were in fact gentler when their turn to rule arrived. What would that show. It would show that a child who has been hurt may hurt others a little less, a thin and unreliable proposition to start with, given how much of the human record consists of the abused growing into abusers. Even granting it, the reversal teaches the wrong lesson. A child who is gentle as a master because he remembers being a slave has still been taught that the world divides into masters and slaves, that he will hold one role or the other, and that his task is to be a kinder sort of master when his turn comes. That leaves the hierarchy standing, merely better mannered, and a kinder boot is still a boot on a neck.

The reversal also smuggles in an assumption this book rejects outright, that suffering is the necessary tuition for the capacity to care. Carry that assumption forward and its cruelty shows plainly. If only those who have suffered can understand, then we are obliged to make people suffer in order to make them understand, and we have just licensed every ordeal ever defended as being for the victim's own good, from the strap to the boot camp to the harshest corners of religious schooling. The decent alternative, that a person can be taught to care for others without first being broken, is both kinder and truer, the only premise under which moral education is anything other than the planned infliction of pain. Elliott took great pride in the reversal, and what it taught her students was that mercy is a thing you earn by being hurt first. That is a lesson the oppressor is always glad to have the oppressed believe, because it tells the person on the bottom that his suffering is educational, that it is doing him good, that he should hold still and learn from it. A method cannot teach its way out of oppression while quietly teaching the central comfort that oppression depends on.

The exercise might have ended there, a strange two days in one Iowa classroom, except that Elliott asked the children to write about it. Their compositions, under the title How Discrimination Feels, ran in the *Riceville Recorder* on the fourth page of its April 18, 1968 edition, and the Associated Press lifted the story off the local page and sent it across the country. A private lesson became a public event in a matter of weeks. The wire story brought an invitation to *The Tonight Show*, where Elliott described the exercise to Johnny Carson's national audience, and the response was immediate and mostly hostile. The switchboard lit with angry calls, and one letter, often repeated since, captured the precise shape of the country's denial. How dare she, the writer demanded, try so cruel an experiment on white children, when Black children were already used to

such treatment. Sit with that sentence, because it does the book's work for me. Nowhere in it is the cruelty itself the problem. The complaint is only that the cruelty had been aimed at the wrong children, that suffering was the natural inheritance of Black children and an outrage solely when it was visited on white ones. The exercise had set out to expose exactly that assumption, and a fair share of the country answered by stating the assumption out loud, in a letter, as a grievance.

From there the thing became an institution. In December of 1970 Elliott staged her exercise for adult educators at a White House Conference on Children and Youth. That same year ABC sent a producer named William Peters to Riceville, and the film he made, *The Eye of the Storm*, put Elliott's classroom in front of a national television audience and turned a rural schoolteacher into a figure of conscience. Peters wrote the exercise up in a book the following year. Fifteen years after that, in 1985, the PBS series *Frontline* aired *A Class Divided*, which brought the grown students of the filmed class back together to testify to what the two days had meant to them. Across roughly two decades, a homemade lesson in a farm-town classroom hardened into a documented, broadcast, published model that any teacher in the country could copy. That is the pipeline that carried it the two hundred miles to Lincoln and set it down, in slightly altered form, on me. The teacher who pinned the armband on my sleeve in 1973 improvised nothing. She was running a program that network television and a major publisher had certified as humane.

Elliott did not run the exercise once. She ran it on the town's children year after year, and then she carried it out of the classroom altogether. By the mid-1980s the demands of the lecture circuit had pulled her away from teaching for good, and she remade the eye-color exercise for adults, for teachers in training, for government workers, and at last for corporations, who paid her to divide their employees by eye color and humiliate the blue-eyed group for an afternoon in the name of teamwork. The lesson she had built in grief became a profession, then a brand, then a product with a price list. I am not interested, here, in litigating whether she should have made a living from it, and the industry that grew up around her is the subject of a later chapter. The point for now is simpler. A two-day improvisation born of a single terrible night became a fixed method, repeated for decades on children who could not refuse it and then sold to adults who could, and at no point along that climb did anyone with the authority to slow it down ask the question this book is asking, which is whether the method teaches what it claims or only what it does.

What the broadcasts never showed, and what it has taken later work to document, is the price. The fullest accounting comes from Stephen Bloom, whose history of Elliott and the exercise calls it, in its subtitle, a cautionary tale, and the caution is earned on nearly every front. Begin with the town. Riceville did not celebrate its famous daughter. When Elliott walked into the teachers' lounge after her television appearance,

colleagues stood and left the room. Downtown, the whispers followed her. The bed-and-breakfast she ran with her husband fell under a quiet boycott once the documentary aired, the slow strangulation a small town can manage without ever holding a meeting about it. Some of her neighbors never forgave her, and decades later she estimated that a fifth of the town remained angry, a number of them still willing to call her the epithet reserved for white people who take the side of Black ones, a word I will not print, though its use in this context proves a great deal. The lesson sold as a cure for racism had not cured the racism of the town that produced it.

The town's ugliness has to be handled carefully here, because it is the kind of detail that tempts a reader to pick a side the book refuses to pick. Riceville behaved shamefully. A community that boycotts a family and slurs a woman for teaching children to question prejudice has indicted itself, and nothing here excuses it. The town's bigotry does not, however, redeem the exercise, and the temptation to let it is precisely the trap this book is built to avoid. Two things are true together. Elliott's method was the oppressor's method, and the town that punished her for using it was steeped in the same prejudice she was trying to fight. A fair account holds both, and refuses the comfort of a story in which the brave teacher's critics are all bigots, so the method must have been sound. The critics in Riceville were largely bigots. Yet the method was unsound anyway, and the soundness of a method is not decided by the character of the people who happen to oppose it.

The heaviest cost fell where such costs usually fall, on children, and not only the children in the classroom. Elliott's own children carried the town's resentment home from school. Her twelve-year-old daughter was surrounded in a school hallway by classmates who taunted her with the claim that her mother would soon be sleeping with Black men. Her son was beaten. A family had been marked, by the same logic of collective punishment the exercise itself ran on, for the offense of the mother's prominence. There is a symmetry in it that Elliott seems never to have named. The town did to her children, out of spite and by accident, a version of what she did to her students on purpose and on principle, singling out the young for a status they had not chosen and could not escape, and making them pay for it in the one currency children cannot do without, the regard of the other children.

Inside the classroom, the children paid a quieter price than Elliott's own, and a longer one, because some of them were filmed. The third graders that Peters recorded in 1970 did not choose to become the faces of a national lesson, and they could not unchoose it afterward. Their worst two days, the hours when they were made cruel or made to cry on camera, were fixed on film and shown to the country, and shown again to their own children, and they carried that exposure into adulthood whether it had helped them or harmed them. When Bloom went looking, decades on, he found the range you would expect and that the celebratory accounts leave out. Some of the grown

students described the exercise the way the reunion footage shows them describing it, as a hinge their lives had turned on for the better. Others remembered it as a wound, a humiliation administered by a trusted adult and then broadcast to strangers, and a few wanted nothing more to do with it or with her. The single experience produced no single result. It produced a scatter, which is what an uncontrolled intervention on twenty-eight separate children was always going to produce, and the scatter is invisible in the documentary, because a documentary is built to show the children who confirm the thesis.

I dwell on the filmed children because their situation makes the consent problem impossible to wave away. An adult cannot ethically take a child's two most vulnerable days, the days she has engineered in order to break him down, and turn them into a permanent public record he will answer for across the rest of his life, however worthy the cause or however moving the footage. We would not accept it now. We should not have accepted it then. The fact that the footage is so powerful is the reason it was easy to accept, and that is the trap exactly. The power of the artifact laundered the ethics of producing it. A nation watched a film that moved it to tears and concluded that something true and good had happened in that classroom, when what the film proved was only that the camera had been pointed at real children in real distress. That laundering, the way a moving result is taken to excuse the means that produced it, is the single move the entire industry would later be built on, and it was already working to perfection in 1970, on children too young to sign a release.

The deepest question is whether the exercise did what it claimed, and here the record turns cold. Elliott has insisted all along that it strengthened the character of the children who went through it, and some of her former students have agreed, returning across the years to tell her that the two days changed their lives for the better. Those testimonies are real, and they deserve respect. They are also exactly the kind of evidence that cannot settle the matter, because the people who seek out a beloved teacher to thank her are a self-selected few, and the children who left the exercise frightened, or humiliated, or quietly convinced that adults could not be trusted, do not as a rule come back for the reunion to say so. When researchers looked past the testimonials for durable effects, they found little. The scholarly verdict on these short, intense interventions, the one Bloom records among them, is that any crash course meant to dissolve racial prejudice in a day yields benefits that fade to almost nothing in the months that follow. The exercise was a powerful experience. A powerful experience and a lasting education are not the same thing, and the assumption that the first must produce the second is the quiet error on which the whole industry that followed would be built.

A question hangs over all of this, the one Bloom's portrait keeps circling. Why did Elliott never doubt. The evidence that her method produced no lasting change was available to her for decades, in the published research, in the critics' arguments, in the

unhealed bigotry of her own town sitting in plain sight. She met every piece of it with the same certainty she had carried out of the classroom on the first afternoon, and she has held it ever since, from stage to stage, sure that she is right and that her opponents are either racists or cowards. That certainty matters, because it reaches past one woman's character. The method manufactures it in whoever runs it.

Consider what the exercise hands its operator. It gives her the moral high ground in the opening minute and never asks for it back. She is the person teaching the lesson against cruelty, which means any objection to her method can be received as an objection to the lesson itself, and any critic can be filed under the heading of the problem she is fighting. The method makes its practitioner feel like the bravest and most righteous person in the room, and a person who feels that way stops asking whether the thing producing the feeling works, because by then the feeling has become the evidence. Here is the danger that runs underneath this whole book, deeper than any single ill-conceived exercise. A technique that confers righteousness on the one who wields it is a technique almost no one will examine honestly, because to examine it is to risk the righteousness, and so it survives every disproof and spreads on the strength of how it makes its users feel rather than on any record of what it does. Elliott is the first and clearest instance. She will not be the last, and the chapters ahead are in large part the story of how a method that fails the people it claims to help keeps winning the loyalty of the people who deliver it.

The borrowed authority of science deepened the trap. Elliott's two days were, from the start, described in the vocabulary of the laboratory. People called it an experiment, and she did not discourage the word, though what happened in her classroom met none of the conditions an experiment requires. There was no control group, no measurement worth the name, no hypothesis that could have failed, and no observer except the woman who had designed the thing and knew in advance what it would show. What she ran was a demonstration, a piece of theater with a foregone conclusion, and a demonstration can only confirm the belief of the person staging it. Dressed in the language of science, though, it took on science's credibility without science's discipline, and a country inclined to trust experts received a teacher's vivid hunch as though it were a finding. That confusion, between a demonstration that moves an audience and an experiment that tests a claim, would shield the whole genre for fifty years. Every later version could point back to Riceville as the place where the thing had supposedly been proved, when Riceville proved nothing, and stood only as the first and most persuasive performance of a belief its author already held.

I do not raise the trauma, the boycott, the harm to her family, or the thinness of the long-term evidence in order to convict Jane Elliott of cruelty, because cruelty was never her aim, and aim is not nothing. I raise these things because the origin already holds, in miniature and in plain sight, every problem the rest of this book will trace through the

larger industry. All of it is there on April 5, 1968: the fabricated science; the badge of caste; the licensed cruelty that arrived in fifteen minutes; the children used as material for a lesson they could not consent to; the reversal that exposed the hierarchy as the teacher's invention; the powerful experience mistaken for a durable cure. The machine left Elliott's workshop already broken, and it reached Lincoln exactly as dangerous as the day she built it. Every fault the industry would later be accused of was present at the source.

There is one element of the Riceville classroom I have saved for last, because it is the element that binds Elliott's room to mine and to all the rooms that came after. Riceville was a white town, nearly without exception, and Elliott's class of twenty-eight held not a single Black child. The lesson that claimed to deliver the experience of being Black was conducted, from first to last, in the total absence of anyone who lived it. No one in that room wore the collar as a fact rather than a costume. There was no one for the children to turn to with the question the exercise pretended to answer. And no one was present whose life would have made the whole performance impossible to stage. The same emptiness I described in my own classroom was the precondition in hers, five years earlier, at the place where the genre began. It is the genre's oldest and most necessary feature, the missing subject, and it is one of the three claims this book exists to prove.

To prove it, and to prove the rest, I have to do to these two classrooms what an engineer does to a machine he means to understand. I have to take them apart. The next chapter sets the two cases side by side, Riceville and Lincoln, the collar and the armband, and pulls from them the common mechanism that every later chapter will test against a fresh example: the caste assigned by birth, the authority that cannot be refused, the cruelty made permissible, the audience, and the reversal that proves the lesson was a performance all along. Elliott built the first machine. It is time to open it up and name its parts.

CHAPTER THREE

The Anatomy of the Machine

Two classrooms in two different states, five years apart, have given us everything we need to build the rest of this book. In Riceville, in 1968, Jane Elliott divided children by the color of their eyes, collared the lower group, and watched cruelty arrive within fifteen minutes. In Lincoln, in 1973, another teacher did a version of the same thing with paper armbands and a boy who would not carry water. I have told both stories as fully as the evidence allows, one from the historical record and one from my own memory, and I have held back, so far, from generalizing, because a witness owes you the particulars before he offers you the pattern. The particulars are in. It is time for the pattern.

The move I am about to make is the move an engineer makes with a device he means to understand. He does not study one unit and stop. He gathers several, sets them side by side, and asks what they share underneath their different casings, because the shared structure is the design, and the design is the thing that can be reproduced, predicted, and judged. Riceville and Lincoln are two units of one device. Strip away the surface differences, the collar and the armband, the melanin lie and the unexplained letters, the Iowa farm town and the Nebraska capital, and the same machine sits under both, built from the same five parts in the same arrangement. This chapter takes the machine apart and names the parts. Then it names the three things the machine does to the people fed into it, the three claims this book exists to prove. By the end you will hold a test, a short list of questions you can put to any lesson that proposes to teach virtue through engineered suffering, and the rest of the book is that test, applied case by case to the whole family of these exercises.

The first part of the machine is a caste assigned by an accident of birth. Every version begins by sorting people into a higher and a lower group using a trait they were born with and cannot change by the end of the day. Elliott used eye color. The Lincoln teacher used eye color too, then translated it into the letters on the armbands. The trait is always chosen for its irrelevance, and the irrelevance is the point, because the lesson means to mimic the arbitrariness of race, and nothing is more arbitrary than being ranked for the color of one’s eyes. Notice what this opening move accomplishes before

anything else happens. It removes the question of desert. A child sorted into the lower caste has done nothing to land there and can do nothing to climb out, which strips away, at the start, the ordinary moral logic a child lives by, the logic in which good behavior earns reward and bad behavior earns punishment. In its place the exercise installs the logic of caste, in which standing is fixed at birth and conduct is beside the point. The designers believe they are teaching children how unjust that logic is. What they are doing, first, is making the children live inside it, and a child made to live inside a logic learns to operate it long before he learns to criticize it.

There is a deeper trouble in this founding move, one that returns in every later case. To sort children by a birth trait and assign worth accordingly is the exact operation of the thing the lesson opposes. Racism is the assignment of worth by a trait of birth. The exercise assigns worth by a trait of birth. It has reproduced the structure of the disease in the first five minutes, before a single cruel word is spoken, and it asks us to believe that reproducing the structure is the way to teach children to reject it. That is the founding contradiction of the whole genre, present at the first step, and everything wrong with what follows grows out of it.

The second part is an authority the subject cannot refuse. A caste means nothing on its own; two groups of children with different eye colors have no power over each other until an adult confers it. The teacher is the engine. She assigns the ranks, writes the rules, enforces them, and above all guarantees them, so that there is no appeal beyond her and no court above her in the room. In Riceville the children swallowed the melanin lie because the teacher told it, and a teacher in 1968 was, to an eight-year-old, an authority close to absolute. In Lincoln the boy who threatened me with a whipping ran to the teacher when I refused, and the teacher disciplined me, not him, because the system she had built required that refusal be punished and compliance be backed. This is the part that separates the exercise from an ordinary schoolyard cruelty, which a child knows how to resist or flee. A playground is lawless on all sides, and a child watches his back there as a matter of course. The classroom is the one place a child is taught to trust, and the exercise turns that trust into the instrument of the harm, using the adult's authority to make cruelty first permissible and then required. Remove the authority and the machine stops. The manufactured ranks dissolve the instant the teacher leaves the room, which is exactly what I watched in the hallway in Lincoln, where the older grades, beyond my teacher's reach, ignored the armbands entirely. A hierarchy that needs a guard to stand over it was never real, only a thing the guard was holding up.

The third part is licensed cruelty. Once the caste is set and the authority is in place, the exercise hands the upper group permission to mistreat the lower, and it hands that permission from the one source a child cannot question. This is the active ingredient, the thing the whole apparatus exists to deliver, and it works with a speed that ought to frighten anyone who has watched it. Elliott said she saw kind children turn vicious in

about fifteen minutes. In Lincoln a boy reached for the word whip inside the first hour, and a child in Riceville suggested the teacher keep a ruler ready for the lower group, the same impulse surfacing in two rooms five years apart, which tells you it belongs to the structure rather than to the children. Hand a child a license to dominate, backed by an adult he trusts, and he will use it, because that is what a license is for, and because the pleasure of standing on someone is among the easiest pleasures there is to discover. The designers read this speed as proof that the lesson is working, that the children are learning how prejudice operates. The speed proves the reverse of what they think. It shows how little it takes to manufacture cruelty in the young when an authority sanctions it, how shallow the reservoir of decency turns out to be once an adult starts draining it on purpose. What the children practice is cruelty from the inside, and practice, in children as in anyone, builds the capacity it rehearses. Recognition of cruelty from the outside, the thing the exercise advertises, is the one lesson it never delivers.

The fourth part is an audience. The cruelty in these exercises is public by design, performed for watchers because the watching is part of the product. A humiliation delivered in private teaches only the person humiliated. A humiliation staged for an audience teaches the audience too, training everyone present in who may be degraded and who may do the degrading. In Lincoln the audience widened by the hour, from the children in the room to the cafeteria to the playground, where every grade could read our ranks off our sleeves. In Riceville the audience began as the class and then, through the essays in the local paper and the wire story that followed, became the readership of the Associated Press, and then, when the cameras came, the national audience of a network documentary. The progression asks to be seen whole, because it reveals what the exercise is for. A lesson honestly concerned with the private moral growth of twenty-eight children would have no use for a camera. The reach toward an ever larger audience tells you that the spectacle is the point, and a thing built to be seen has confessed that the seeing is its true work and the teaching a pretext. The audience does something quieter and worse as well. It converts the children into a demonstration, raw material for a lesson aimed at the watchers, and a child used as material for someone else’s lesson has been turned from a person into a prop, which is its own small cruelty laid over the staged one.

The fifth and final part is the reversal. On the second day, in both classrooms and in nearly every version since, the roles are swapped: the masters become servants and the servants become masters, and the exercise presents this as its masterstroke of fairness, the part where everyone learns the lesson from both sides. The reversal gives the most away of any part of the machine. It proves, past argument, that the cruelty was never deserved and never earned, because the same child is a master on one day and a servant on the next, his worth raised and lowered on a schedule by an adult with a clipboard. A status that the person who assigned it can reverse overnight was never a fact about the

child, only a costume the teacher handed out and took back. Once you see the hierarchy for what it is, you see what the children were in fact being taught. The lesson about injustice never lands. The lesson about power lands completely, because the children have watched, in the clearest possible demonstration, that caste is a game an adult can run, with rules an adult can rewrite, in which their only task is to play whichever part they are handed. Sold as the lesson's conscience, the reversal is in truth the lesson's confession, the moment the machine admits that the whole hierarchy was arbitrary all along, assigned for effect by the one person in the room with the power to assign it.

It helps to see that these five parts form more than a checklist of independent features that happen to co-occur. They are a system, and the order is the design, because each part creates the conditions the next one needs. The caste comes first because the authority needs something to assign; without two groups there is nothing to rank. Authority comes second because the cruelty needs a license, and only an adult the children cannot refuse can issue one that holds. Cruelty comes third because the audience needs something to watch, and a room of children behaving decently makes poor theater. The audience comes fourth because it generalizes the lesson, carrying it from the few children inside the cruelty to the many who witness it and learn the roles. And the reversal comes last because, by the time it arrives, it can do its real work, which is to teach the children that the whole arrangement they have just lived through was a matter of an adult's choosing, rerunnable and reversible at will. Pull any one part out and the machine falters; assembled in order, the parts reinforce one another and produce an effect no single part could produce alone, which is the signature of a machine rather than a mistake. The people who run these exercises did not, for the most part, design it on purpose. They inherited it whole, from Elliott and her imitators, and run it without seeing the engineering, the way a person drives a car without understanding the engine.

Before the parts are assembled into a verdict, one objection has to be met, because a careful reader will already have raised it. Other institutions share some of these parts and draw no complaint. A football team assigns rank, demands obedience, and works its rookies hard in front of a crowd. An army sorts recruits, strips their individuality, and drills them to exhaustion. The flight school drops a pilot into a failing engine, and the surgical residency walks a trainee through a staged crisis built to rattle him. If the five parts were the whole story, every one of these would stand condemned beside Elliott's classroom, and the argument would prove far too much. What separates them is the presence or absence of a real task. In the team, the army, the cockpit, and the operating room, the hardship is the price of a skill the person will carry out of the room and use, the game won, the war survived, the plane landed, the patient saved, and the difficulty is shaped to build a capacity that outlasts the day. The exercises in these pages answer to no such task. They train no skill and leave nothing the person carries forward; the degradation is itself the content, the staged wound offered as the thing to be learned,

when whatever lesson it claims could have been taught by gentler means that build rather than break. That is the boundary of this book’s indictment, and it should be kept in hand through every chapter that follows, because the defenders of the exercise will reach for the boot camp and the football field again and again, and the reply does not change. Those forge a capacity. The exercise only stages a wound.

Two clarifications keep that boundary honest, because it does heavy work and a defender of the exercises will test it. The first is that I am not settling by definition the question of whether these exercises produce the change they promise. A trainer will insist that his workshop serves a real task, a fairer workplace, and that the awareness it instills is a durable skill that shows in what a person later does, and that claim cannot be waved away from an armchair. It has to be checked, and the later chapters check it against the measured outcomes of the exercises themselves, where the lasting change fails to appear and only the passing feeling remains. The boundary drawn here is a claim this book will earn with evidence, not a line imposed by fiat. The second clarification is that none of this is a verdict against learning by doing. Experience teaches, and some things are learned in the doing that no telling can deliver, which Dewey understood and which the remedy this book ends on depends upon, since cooperation among equals and sustained contact are themselves experiences and not lectures. The line that matters falls elsewhere, between an experience a person enters freely and as himself, in the company of a real other who authors his own part and toward a task that pays off in the world, and a degradation assigned to him by category, performed for a crowd, and staged around a proxy standing in for the people it claims to be about. The cure this book reaches is experience of the first kind, the cooperative classroom and the contact among equals, where the other is present and speaks for himself, which is the whole distance between it and the exercise that needs its subject gone. Experience as such sits on the safe side of that line. My target throughout is the coercion and the counterfeit, the assignment of a wound where a person should have been left to build a capacity, and a lesson does not earn my objection merely by being felt.

Those five parts, the caste by birth, the unrefusable authority, the licensed cruelty, the staged audience, and the reversal, are the whole machine. Any lesson that assembles all five, in the bounded sense just drawn, sits at the center of the family this book indicts, whatever it calls itself and whatever banner it marches under, the form Elliott built and the privilege walk and the mock auction inherit. Around that center sits a looser ring, the exercises that run on the same engine without staging the whole theater, the coercion of an authority or a category and the proxy that stands in for a real encounter, while missing one part or another, the fear tour that skips the reversal, the drug lesson that assigns no caste. Both belong to the indictment, because both coerce a feeling and both put a stand-in where the real thing should be, and the difference is one of degree, the core earning the book’s sharpest charge and the ring a milder one. I set the two tiers

down here because the distinction matters later, when a critic objects that a prison tour and an eye-color drill cannot be the same object. He is right, and the taxonomy answers him: they are kin of different degree inside one family, near relations rather than identical twins, which keeps the category from being drawn so wide that a single mismatch breaks it. Naming the parts tells you how the machine is built. It does not yet tell you what the machine does to the people fed through it, and that is the more important question, the one the rest of this book is organized to answer. There are three answers, three claims, and I want to state them now as plainly as I can, because everything ahead is the work of proving them.

The first claim is the inversion, and it is the one the book is built on. The method of these exercises is the method of the oppressor, even when the cause is just. To teach a child against domination, the exercise dominates him. To teach him against caste, it casts him. And to break him of the arbitrary use of power over the powerless, it uses arbitrary power over him. The tools are the oppressor's tools, picked up in a good cause, and the cause does not change what the tools do, because a method carries its own content. You cannot teach gentleness by force, or equality by ranking, or freedom by command, any more than you can teach sobriety by getting a child drunk. The lesson the child absorbs is the lesson the method embodies, and the method here embodies the thing it claims to oppose.

The case was made, with more authority than I can claim, in the same year Elliott built her exercise. In 1968 the Brazilian educator Paulo Freire published *Pedagogy of the Oppressed*, the book whose title this one inverts on purpose. Freire's subject was the education of the oppressed themselves, the peasants and workers he had taught to read, and his central warning runs straight against the exercises in this book. Liberation, he argued, cannot be handed down from above by an authority depositing correct attitudes into passive minds, a thing he called the banking model of education, in which the student is an empty vessel to be filled. And liberation cannot be reached, he insisted, by the oppressed merely trading places with their oppressors, reversing the terms while leaving the structure of domination standing. The goal, in his phrase, is a person who is neither oppressor nor oppressed. Freire saw, more than fifty years ago, that a pedagogy which mirrors the power it claims to fight produces no liberation, only a new turn of the same wheel, and that warning is the theoretical spine of this chapter. Coercive moral pedagogy is the banking model armed, the depositing of correct attitudes by force, and its two-day reversal of master and servant is the exact maneuver Freire said could free no one. The exercise built in 1968 and the book published in 1968 are mirror images. One is the pedagogy of the oppressed. Elliott's exercise, for all its good intentions, is the pedagogy of the oppressor, and I will give Freire and this inversion a full chapter of their own later in the book.

There is a sharper irony still, one that scholars of Freire have noticed and that bears directly on the exercises in this book. Freire wrote for the education of the powerless, the people on the underside of a real hierarchy, and his pedagogy was meant to be practiced by them and with them. The exercises that descend, knowingly or not, from the same era do the reverse. They are aimed mostly at the comfortable, at white schoolchildren and salaried employees, the people nearer the top of the real hierarchy than the bottom. Applied there, a pedagogy built to liberate the oppressed becomes something stranger, a brief and managed taste of oppression administered to the unoppressed, on the theory that a person who has sampled the underside will treat the people who live there with more care. Freire would have recognized the theory and rejected it, because he held that understanding grows from sustained reflection and action in the world, from what he called praxis, and not from a staged ordeal that ends at the bell. The exercises keep the vocabulary of liberation and discard the thing itself, which is how a method can sound like Freire while doing the precise opposite of what he taught.

The second claim is that these exercises are iatrogenic. That word comes from medicine, from the Greek iatros, physician, and genesis, origin, and it names an illness produced by the treatment meant to cure it. The thinker who gave the idea its broader force was Ivan Illich, whose 1975 book *Medical Nemesis* argued that modern institutions, past a certain scale, routinely produce the opposite of what they promise, so that medicine manufactures sickness and schooling manufactures ignorance, and the cure becomes the cause. Illich called the general pattern counterproductivity, and coercive moral pedagogy is among its clearest instances. The exercise promises empathy and produces its opposite. It sets out to reduce cruelty and manufactures cruelty on the spot. It sets out to build understanding and issues, instead, a false certificate of understanding to children who have learned nothing except how the machine of contempt is run. The treatment causes the disease.

A full chapter later in the book lays out the evidence that this is so, because the claim is strong and deserves more than assertion. For now it is enough to see the shape of it in the two cases we have. The children in Riceville left no gentler; some of them left having discovered how good it felt to rule, and having watched a trusted adult reward the discovery. The children in my room left with no sympathy for Black Americans, who were nowhere in the building; they left having learned that an adult with power will license cruelty so long as it comes wrapped in a lesson. Whatever these exercises deposit in a child, it is not the thing printed on the label, and the gap between the promise and the deposit is the defining feature of an iatrogenic treatment, the thing that should stop a careful person cold, rather than a minor inefficiency to be tuned away. A cure that reliably produces the disease is worse than no cure at all. It occupies the ground a real remedy might have stood on, and it sends the patient home believing he has been healed.

Let me be concrete about how the treatment produces the disease, because iatrogenesis is easier to assert than to show. The exercise makes three deposits in a child, and not one of them is empathy. The first is practice in cruelty, the rehearsal of contempt that the favored group performs and that practice, as ever, strengthens. Second is resentment, the residue left in the children who were degraded, who learned that an adult they trusted would turn on them for a trait they could not change, a lesson more apt to breed wariness than fellow feeling. The third deposit is the most insidious, the false certificate, the unearned conviction in every child who went through it that he now knows what prejudice feels like and has therefore done his moral homework. That certificate is worse than ignorance, because the ignorant person knows there is something he has not learned and may yet go and learn it, while the certified person believes the lesson is finished and closes the book. Add the three deposits together, the practiced cruelty, the resentment, and the false certificate, and you have the opposite of the stated goal, manufactured reliably, by a method that congratulates itself the whole time. That is iatrogenesis in full, the cure that does more than fail, since it actively produces the condition it was meant to relieve, and a careful look at the two classrooms shows all three deposits being made in real time.

The third claim is the one this book most makes its own, less the discovery of the absence than its elevation into a general law and a diagnostic test, and it is the simplest of the three to state and the hardest to forget once you have seen it. These exercises require the absence of the people they claim to be about. The lesson that would teach white children how it feels to be Black depends, for its operation, on there being no Black child in the room, because a Black child present would turn the costume back into a life and make the performance impossible to sustain. Riceville was an all-white town and Elliott’s class held not one Black child. My school in Lincoln held one or two in the whole building and my classroom held none. That emptiness made the exercise possible. The subject of the lesson had to be absent for the lesson to proceed, and a lesson that cannot survive the presence of its own subject was never a lesson about that subject at all, only a performance staged in the subject’s absence, for the benefit of the people performing it.

This third claim does a particular kind of work in the book, because it hands you a tell, a single feature you can look for that exposes the whole exercise from the outside. When you find one of these lessons, ask who is missing. Ask whether the people whose experience is being simulated are in the room as participants, present and able to speak, or whether they have been arranged out of it, kept offstage so the simulation can run undisturbed. A privilege walk needs the poor kept at a comfortable distance. The poverty simulation needs no poor people in attendance. The exercise that teaches about a group almost always teaches in that group’s absence, and the absence does real work in the teaching. It is what makes the teaching feel safe to the people who run it, because the

missing subject cannot contradict the performance, cannot complicate it, cannot turn to the room and say that this is not what it is like at all.

The clearest confirmation of the third claim comes from the rare occasions when the subject is not absent, when someone who lives the condition the exercise simulates happens to be in the room. On those occasions the exercise collapses rather than deepens. A poverty simulation run in front of people who have been poor curdles into insult, because the participants who have lived it can feel the distance between the game and the fact, and they tend to say so. A lesson on disability that asks the able-bodied to spend an hour in a wheelchair plays differently when a wheelchair user is present to point out that the hour taught the able-bodied mostly about the awkwardness of a borrowed chair and almost nothing about a life. The presence of the real subject introduces the one thing the exercise cannot survive, which is correction, the live voice that says this is not what it is like. That is why the genre drifts, without anyone planning it, toward rooms the subject has been kept out of, toward the all-white classroom and the all-comfortable workshop. The absence does more than limit these exercises; it is the condition that lets them feel like revelation to the people inside them, and the moment it is lifted, the feeling goes with it.

Before I turn the five parts and the three claims into a test, I have to say plainly what the machine is not, because a test is worthless if it catches things it should not, and the line here matters more than almost anything else in the book. Teaching children about injustice is necessary, and nothing in these pages argues against it. A child should learn what slavery was, and what the Holocaust was, and what segregation did, in detail and without flinching, and that learning will and should disturb him. The disturbance is honest, because it comes from the facts. What separates honest moral education from the machine is that honest education manufactures no victim in the room. It does not sort the class into castes, license some children to abuse others, stage the abuse for an audience, or hand out suffering on a schedule and call it a curriculum. A class that reads a survivor’s account, or hears a survivor speak, or studies the documents of a massacre, is learning from the testimony of the people who lived it, with those people present on the page or at the podium as witnesses rather than absent as an alibi. The direction of the thing is reversed. In honest education the suffering being studied is real, historical, and someone else’s, approached with the humility owed to it. In the machine the suffering is manufactured, present, and the child’s own, inflicted by the teacher to make a point. One asks a child to understand a reality. The other asks him to perform a counterfeit of it. The test I am about to give catches the second and leaves the first alone, and any reading of this book that turns it into an argument against teaching hard history has mistaken the target completely.

With the line drawn, I can deal quickly with the defenses these exercises usually raise, because each of them dissolves against the anatomy. The first defense is that it is

only a simulation, harmless because everyone knows it will end. The anatomy answers that the harm lies elsewhere than in the belief that the caste is permanent. It lies in the three deposits the anatomy named, all of which land whether or not the child knows the clock is running. A rehearsal of contempt is still a rehearsal, and rehearsal builds the thing rehearsed.

The second defense is personal: it worked for me, says the adult who went through it as a child and turned out caring. That testimony is sincere, and it is also the weakest kind of evidence, for the reasons the last chapter set out. The people who seek out the teacher to thank her are self-selected, the harmed stay silent, and a single vivid memory tells you nothing about the average effect across the many children processed alongside the one who remembers it fondly. An anecdote of survival is no measure of safety. Plenty of people survive things that injure most of those subjected to them.

The third defense is the deepest and deserves the most respect: you have to feel it to understand it, and only the felt version teaches what the explained version cannot. This is the claim the whole genre rests on, and the anatomy shows exactly where it fails. What the exercise makes a child feel is the experience of being sorted and bullied for an afternoon under a teacher’s control, with a known end and nothing real at stake. That feeling and the lived condition it claims to convey are different in kind, as I argued from my own classroom, and feeling the first teaches almost nothing reliable about the second. Worse, the felt version carries the iatrogenic freight the explained version does not, the practiced cruelty and the counterfeit understanding that an honest lesson, taught from real testimony, never deposits. The premise that feeling beats understanding turns out to be backward. The understanding is what lasts, and the feeling, in this manufactured form, is what does the damage.

There is a fourth defense, quieter than the others, which is simply that the intentions are good, and that anyone attacking these exercises must be soft on the prejudice they target. The whole architecture of this book denies that move its footing. Equality is defended on every page. The quarrel is with the method, and a method can be condemned with no retreat at all from the goal it claims to serve, because the worth of a method is a separate question from the worth of the aim, and confusing the two is how bad methods survive inside good causes for decades at a stretch.

Put the five parts and the three claims together and you hold a test, the practical gift of this chapter and the tool the rest of the book will use. Confronted with any lesson that proposes to teach virtue through engineered suffering, you can ask a short series of questions and know, quickly, what you are looking at. Does it sort people into a higher and a lower group by something they did not choose. Is the resulting authority one the lower group cannot refuse or appeal. Are the favored licensed to mistreat the others. Is the cruelty performed for an audience. And does it reverse the roles and call the reversal fairness. If the answers run to yes, you are looking at the machine, and you can predict

what it will do, because the machine, fully assembled, does the same three things. It will use the oppressor's method in a good cause. It will tend to produce the harm it promised to cure. And it will require the absence, or the silence, of the people it claims to serve.

The chapters ahead run this test on case after case, and the cases grow larger as the book goes on: the simulation game a California designer built for the classroom; the privilege walk and the poverty tunnel that became fixtures of the corporate workshop; the mock auctions and role-plays that surface as scandals in the schools every few years; the prison tours that promised to scare children straight; the drug-education curriculum a whole generation sat through; the mandatory diversity training that corporations now buy by the millions. Some assemble the full five parts and some belong to the looser ring set out earlier, and the chapters will say which is which as each case arrives. Each looks different on the surface, wears a different banner, claims a different pedigree. Run the test and the same engine appears beneath them, doing the family's work to whatever degree the case completes the design, and the resemblance becomes impossible to miss.

A reader is owed fair warning that this book does not mean to argue from logic and example alone. The cases just named have mostly been measured, with the same tools medicine uses to test a drug, and Part Five sets the measurements out in full; a preview is enough here. The prison tour built to frighten children away from crime raised the crime instead, the drug curriculum a generation sat through moved no use and by some counts nudged it upward, the corporate training meant to lift the standing of women and minorities was followed in the largest study by a decline in it, and the disability simulation built to raise regard for the blind lowered it against people who did nothing, while the alternative the book ends on, sustained contact among equals, carries the strongest evidence in the whole field. Some of those results are active harm and some are only waste, a difference Part Five keeps in view, and it is the convergence of them across unrelated fields, more than any single number, that turns suspicion into proof. The reasoning comes first by design, on the theory that a reader persuaded by the logic meets the numbers ready to receive them. The numbers are real, and they are coming.

One feature of the machine matters more than any other for what follows, and it is the reason this book does not stop at the schoolhouse door. The anatomy is indifferent to scale. The same five parts that ran in a classroom of twenty-eight children run, unchanged, in a workshop of fifty adults, a training center serving a corporation of thousands, or a public culture of millions, because nothing in the design depends on the size of the room. A caste can be assigned to a nation as readily as to a class. An authority that cannot be refused can be a teacher, an employer, or a state. Cruelty can be licensed in a hallway or in a hiring policy. An audience can be a playground, a television network, or a social-media feed. And the reversal, the lesson that the roles belong to an authority to assign and reassign, scales up into something more dangerous than a bad afternoon, into a politics in which whole groups are told that it is their turn now to

occupy the bottom, for the good of their souls. Growing the machine does not improve it. Scaling a method that manufactures cruelty, produces the harm it promised to cure, and requires the absence of its subjects only manufactures more cruelty, more harm, and a larger absence. The later chapters follow the machine up that ladder of scale, from the single classroom to the simulation industry to the training-and-consulting business to the politics of the present, and the argument does not bend as the rooms grow, because the machine does not change. It only takes more people inside it. The reason to take apart a small machine carefully, the reason I have spent three chapters on two classrooms, is that the small machine and the large one are built to the same plan, and a person who can recognize it in a third-grade room of armbands will recognize it anywhere, which is the whole purpose of the test.

The machine is now fully specified, named in its five parts and judged by its three effects, and that specification is the foundation the rest of the book stands on. What began as one boy’s memory of an afternoon in Lincoln, and one historian’s account of a week in Riceville, has become a blueprint, and a blueprint can be checked against any building that claims not to use it. The family of these exercises is large, and its members are respectable, taught in schools of education, sold by consultants, defended by people of obvious good will. None of that respectability survives contact with the blueprint, because the blueprint is in the building whether the builders admit it or not. We begin with the nearest relatives, the simulation games and the privilege walks that took the classroom exercise and carried it into the wider world, and we will find in every one of them the same five parts in the same arrangement, doing the same three things to the people fed through them.

PART TWO

The Family

CHAPTER FOUR

StarPower

Jane Elliott's exercise was no single eccentric invention, the brainstorm of one Iowa teacher working alone. It belonged to a family, born in the same years, raised on the same assumptions, and spread by the same conviction that the way to teach a hard truth was to stage it and make people live through it. The end of the nineteen-sixties produced a small boom in this kind of thing, a movement of educators and designers who believed that simulation, the building of artificial little worlds in which students could act out the dynamics of the real one, was the future of teaching. They had a name for their field, simulation gaming, and journals and conferences and a sense of standing at the frontier of education. The emblem of the movement, the simulation more people have passed through than any other and the one that shows the family's logic in its purest form, is a classroom game called *StarPower*. This chapter is about that game, the man who built it, and the revealing fact at its center, which is that its designer learned to read the cruelty his game reliably produced as proof that it was working.

The movement that produced *StarPower* had a creed, and the creed matters because it explains both the ambition and the blindness. Its premise was that people learn far more from doing than from being told, that an hour spent inside an experience teaches what a year of lectures cannot, and that the classroom of the future would be a place where students lived through scaled-down versions of real situations and drew their own conclusions. There is truth in the premise, which is part of why it spread so fast. A pilot does learn more in a simulator than from a manual, and a medical student learns more from a practice patient than from a diagram. The educational reformers of the late nineteen-sixties took that sound idea and extended it into territory where it stops being sound, the territory of moral and social experience, on the assumption that what works for flying a plane works for understanding injustice. They built simulations of international trade, of legislatures, of war, and they built simulations of prejudice, and the whole enterprise carried the confidence of a field that believed it had found a better way to teach almost anything. What the movement lacked was any equally developed method for telling whether its powerful experiences were teaching the right thing, or

anything durable at all. It had a sophisticated theory of engagement and almost no theory of outcome, and into that gap *StarPower* and its kin walked, vivid and memorable and largely unexamined, their emotional impact taken as a stand-in for their educational value because the field had no better measure on hand.

The man was R. Garry Shirts, and the origin of *StarPower* deserves to be told precisely, because it places the game in exactly the company this book has been keeping. In 1969, the year after Elliott first divided her third-graders by eye color, Shirts was approached by the National Council for the Social Studies and asked to build a simulation game that would help students understand racism, a game he was to unveil at the organization's national convention. Sit with that for a moment. The same year, the same country, the same impulse that drove a teacher to brand her children with collars drove a professional designer, at the request of the national body of social-studies teachers, to manufacture an experience of racial hierarchy for the classroom. Shirts's first attempt built a scenario around a real incident, a fight between a Black student and a white student over a racial slur that had spread into a cafeteria brawl, with the players cast as participants in a meeting to address it. That version did not satisfy him, and what he produced instead abandoned the explicit subject of race for something he judged more powerful, a game about the bare mechanics of advantage, in which the lesson about hierarchy would be delivered without anyone being told that hierarchy was the subject. The result was *StarPower*, and it traveled far past anything Elliott's classroom reached.

The decision to drop race from a game commissioned to teach about race deserves more than a passing glance, because it reveals something about how the dramatic method understands its own subject. Shirts judged the abstract version more powerful, and in a narrow sense he was right, since a game about pure advantage could be run on any group anywhere and would produce its effect regardless of who played. The cost of that power was the subject itself. A lesson explicitly about the hatred of Black Americans became a lesson about colored chips, and the particular history the commission had asked him to teach, with its specific cruelties and its specific victims, dissolved into a general parable about hierarchy that mentioned none of them. This is the absence of the subject arriving by a different road than it took in Elliott's classroom. Where Elliott substituted eye color because no Black child was present, Shirts substituted chips because abstraction made the product more marketable, more universal, more reliably effective across audiences. The two roads reach the same place. In both, the people the lesson was meant to be about vanish from it, replaced by a token the players can pick up and put down, and the designers experienced the vanishing as a gain in power rather than a loss of truth. A method that improves when you remove its subject has told you what it values, and what it values is the effect on the audience, not fidelity to the people the audience is supposedly learning about.

The game is simple to describe and insidious in operation. A group of players, anywhere from a dozen to thirty-five, is given colored chips drawn at random from a bag, with different colors worth different numbers of points, and is told to trade chips with one another to raise their scores, keeping the contents of their hands secret. After a round of trading, the players are ranked by score and sorted into three groups, each marked with a badge they wear, the high scorers becoming Squares, the middle scorers Circles, and the low scorers Triangles. The players are told that the sorting reflects achievement, that they have earned their places by skill at the game. That telling is a lie, and the lie is the engine. The initial random draw, and no skill, determines most of what follows, and the game is rigged from the second round forward to make the advantage compound, because the Squares now draw their chips from a bag loaded with high-value pieces while the Triangles draw from a bag of low-value ones. Those at the top are made richer by the structure and told they are rich by merit. A thin pretense of mobility is preserved, an occasional Triangle handed enough chips to rise, so the meritocratic story stays plausible, but the overwhelming reality is that Squares stay Squares and Triangles stay Triangles while the distance between them grows. Anyone who has followed the third claim of this book will already see the shape, a caste assigned by accident and then rebranded as desert.

The way the game feels to play is part of its design, and the feeling follows a curve worth tracing. Its opening rounds are pleasant, even festive, a room of people bartering chips and joking, with nothing yet at stake beyond a score that means little. The badges change that, slowly. Once a player is wearing the symbol of his rank, the symbol begins to do work on him, and the work is the point. The Square learns to carry his badge as a mark of worth, the Triangle to feel his as a mark of failure, and because the symbols are visible to everyone, they organize the whole room into a hierarchy each person can see and feel his place within. This is the same machinery that makes a uniform or a yellow star do its work, the conversion of an assigned category into a felt identity through a marker worn on the body, and the designers understood its force well enough to build the game around it. The anthropologists who later adapted *StarPower* used exactly this feature to teach how visible markers of identity function in stratified societies, how a badge can come to stand in for a person and to license his treatment. That lesson is sound. The trouble is the means, because to teach a child how a degrading marker feels, the game makes a child wear one and be degraded, and the wearing is real even though the badge is plastic. By the time the Squares are sent off to write the rules, the room has been sorted into the comfortable and the humiliated as thoroughly as any real hierarchy sorts people, and the sorting was accomplished, as it always is in this genre, by an authority no one in the room was free to refuse.

Then comes the move that makes *StarPower* what it is. After several rounds, the facilitator announces that because the Squares have done so well, they have earned the

right to make the rules for the rest of the game. The Squares withdraw to confer, and the moment they do, the temperature of the room changes. Up to that point the game has been social and loose, people laughing and bargaining and enjoying themselves. The instant a group of children or adults is handed unchecked power over everyone else and sent off to write the rules, the laughter stops and the game turns earnest and grim. What the Squares do with their power is the most reliable finding in the whole history of the game. They use it to entrench themselves. They reshuffle the bags in their own favor, award themselves free points, write rules that make it nearly impossible for anyone below to rise, and they do this, in the observed pattern, with relish, consulting the lower groups rarely and heeding them never. Below them the Circles spend their energy trying to climb into the Square group, and the few who manage it begin at once to behave like Squares. At the bottom, the Triangles move through a sequence the game produces over and over, from anger to self-blame to a flat despair, sitting out a contest they cannot win, coming alive only when they sense a chance to cheat or to overturn the board. The game stages a society stratifying itself, licenses its winners to be cruel, and watches the predictable result unfold.

Here is the detail that opens the whole enterprise to the light. Shirts has described the early testing of the game, when he was still unsure it would work, still worried that the brightly colored plastic shapes the players wore around their necks would strike them as too silly to take seriously. His worry vanished, he reports, at a particular sight. He watched one of the Squares, the top group, shake his plastic square in the face of a Triangle and threaten him, telling the lower-status player he would do as he was told, in Shirts's recollection, "because I'm a Square that's why." That was the moment the designer knew his game was a success. Read that again with the book's argument in mind. The proof that the game worked, the sight that allayed its creator's fears, was a child using a manufactured status to threaten and dominate another child. The cruelty was no side effect the designer noticed with regret. It was the evidence of efficacy he had been hoping to see. And in that reaction, multiplied across the whole simulation-gaming movement, lies the error this chapter exists to name. The designers of these games built machines that reliably produce the spectacle of cruelty, and because the spectacle was reliable and vivid and emotionally powerful, they concluded that the machines reliably produced understanding. They mistook a dependable production of cruelty for a dependable production of insight, and the two are not the same thing.

The reliability is real, and *StarPower*'s defenders are right to insist on it. The game does produce the same results almost everywhere, with players of every age, nation, and politics, and that consistency is interesting as a fact about human beings. Put people in the position of the Square and they will, with few exceptions, behave like Squares; put them in the position of the Triangle and they will, with few exceptions, sink into the apathy of the Triangle. One of the most telling moments anyone has recorded came

when a Square argued to her fellow Squares that the unfairness was making the game tedious and ought to be corrected, and the other Squares responded by stripping her of her points and demoting her to Triangle. The person who told the story drew the right conclusion, that the Squares who punished her were not unusually cruel people, only people occupying the position of Square. That observation is correct, and it is the heart of *StarPower*'s intended lesson, that structures shape behavior, that the system a person is placed in does more to determine his conduct than his private character. The irony, which the game's designers and champions seem never to have absorbed, is that the lesson applies to *StarPower* itself. If a structure can turn ordinary people into petty tyrants and broken subordinates, then the exercise is such a structure, manufacturing the tyranny and the brokenness it claims to be merely revealing, and its reliability in doing so is no sign that it teaches the lesson well, only a sign that it commits the act the lesson condemns.

The fate of the Square group deserves particular attention, because it is where the game's iatrogenic character shows most plainly. The intended lesson is that power corrupts and that systems shape behavior, a lesson aimed in large part at the powerful, meant to make them examine their own advantage. What the record consistently shows is that the Squares learn no such thing. They leave the game, by the repeated observation of those who run it, largely unable to see the oppression they committed, reluctant to admit they abused their power, and in many cases unwilling to concede that they enjoyed it. The reason is the one the earlier chapters established. An hour spent practicing domination, with permission and with success, trains a person to perform domination fluently and to find it natural rather than to question it, and it rewards him for doing so with points and status and the deference of the room. The game sets out to inoculate the powerful against the temptations of power and instead gives them a supervised hour of indulging those temptations, which is closer to training than to warning. For the group the lesson most needed to reach, the game reliably produces the opposite of its aim, sending the Squares away more comfortable with power than they arrived, having learned in their bodies that being on top feels good and that the people below can be managed. The cure manufactures the disease in exactly the population the cure was for, and the debrief at the end, a conversation, is asked to undo an hour of lived and pleasurable dominance, which is a great deal to ask of a conversation.

If the Squares learn the wrong lesson, the Triangles pay the wrong price. The accounts of the game describe what happens to the players sorted to the bottom in language that ought to give its champions pause, a progression through anger and resentment into self-blame, self-criticism, and finally a dull, withdrawn despair. These are children, in the school versions, being walked deliberately into the feeling of being a failure in front of their classmates, and the feeling is not pretended. A child who has been ranked last, made to wear the mark of last, and then watched the rules rewritten to

keep him last has been given a real experience of exclusion, with real distress attached, and the fact that the exclusion is staged does not reach down and soften what he feels while it is happening. The defenders call this a controlled dose, a manageable taste of what real exclusion is like, and treat it as the productive core of the exercise. That framing understates the harm and overstates the benefit. The harm is real distress inflicted on a child for the instruction of others, and the benefit, the durable understanding meant to justify the distress, is the thing the evidence chapters will show to be uncertain. A method whose lower half consists of deliberately producing despair in children, on the promise of a lasting insight that may not survive the week, carries a cost the genre's enthusiasm tends to hide, and the children sorted to the bottom are the ones who pay while the lesson, if it lands at all, lands somewhere else.

Everything in the defense of these games eventually rests on the debrief, the discussion after the play in which the facilitator helps the participants name what happened and draw the intended lessons from it, and the weight placed on that conversation is more than it can bear. The debrief is asked to perform a conversion, to take the raw experience of an hour and turn it into durable understanding, to make the Square see his cruelty and the Triangle see the system rather than himself. Sometimes it helps. The deeper problem is an asymmetry the debrief cannot overcome, between the kind of learning the experience installs and the kind the talk attempts. Experience writes its lesson in the body, in the felt pleasure of dominance and the felt sting of exclusion, where learning is deep and durable and hard to argue away. The debrief offers a verbal correction, a set of propositions about power and systems, delivered to people whose bodies have just been taught something else. When the visceral lesson and the verbal lesson agree, the debrief reinforces what the experience taught. When they conflict, as they do for the Square who enjoyed his power and is now told he should have questioned it, the words tend to lose to the feeling, and the participant leaves having been told one thing and having learned another. A conversation at the end is a thin instrument for undoing what an hour of lived experience has written, and the genre's faith in it is the faith of people who would rather not look too hard at what their exercises deposit.

StarPower did not stay in one classroom. By its creator's estimate the number of people who have played it runs into the millions, and it has been administered in schools and universities, in corporations and government agencies, in military training and management seminars, across decades and around the world. It became a fixture, a thing a teacher or trainer could order, with a director's manual and a set of materials, and run on a Tuesday afternoon to deliver a guaranteed emotional experience. Its designer went on to build other simulations in the same spirit, including a widely used cross-cultural game, and the broader movement he belonged to produced dozens of these artificial worlds for the classroom, on poverty, on trade, on war, on prejudice. A version of

StarPower even found its way into the teaching materials of a major public-education project on race mounted by professional anthropologists. The dramatic simulation, in other words, did not remain the eccentric practice of a few enthusiasts. It professionalized, acquired institutional backing, and embedded itself in the standard repertoire of American education and corporate training, where much of it remains. The family that produced Elliott's exercise grew into an industry, and a later part of this book follows the money. Here the point is the spread itself, the movement of a single dubious idea from the margins into the mainstream on the strength of how powerful it felt to administer.

StarPower repays examination because it shows the book's three claims operating in a case quite different on its surface from the eye-color exercise. The five-part anatomy is all present. A caste is assigned by something other than merit, the random draw dressed as achievement. The authority is unrefusable, the facilitator who runs the game and grants the Squares their power. The cruelty is licensed, handed explicitly to the top group as the right to oppress the others. An audience watches, the whole room observing the badges and the huddle and the slow humiliation of the bottom. And the whole structure is that of a lesson depending on people behaving badly on cue. The inversion is present too, in its sharpest form yet. *StarPower* is a lesson against the abuse of power that operates by licensing the abuse of power, a teaching about oppression conducted by oppressing a third of the room, and the contradiction is constitutive of the method rather than incidental to it, since without the real oppression of the real Triangles there would be nothing to observe. The iatrogenic pattern appears in the fate of the Squares, who, the record shows, typically leave the game without seeing the oppression they took part in, having practiced domination and enjoyed it and learned to find it natural. And the absence of the subject is there in the origin, a game built to teach about racism that contains no racial subject at all, only chips and badges, the hated group of the original commission replaced by a proxy so abstract that the lesson can be run on anyone because it is about no one.

One feature of the game's authority structure repays drawing out, because it sharpens the inversion to a fine point. *StarPower* teaches, among its official lessons, that power without checks will be abused and that every system needs limits on those at the top. The game makes that case by handing the Squares unchecked power and letting the room watch them abuse it. Notice, though, where the unchecked power in the room sits, which is with the facilitator. He rigs the bags, assigns the ranks, decides when the Squares may rule, and stops the game when he pleases, and his authority over the whole arrangement answers to no one inside it. The game stages a lesson about the danger of unchecked power inside a structure governed by an unchecked power that no one questions, the adult running it, and the irony goes unremarked because the facilitator's control is the one thing the participants are trained to accept as natural. The Squares are

the visible tyrants, the designated villains whose abuses illustrate the lesson, while the real sovereign of the room stands outside the frame, having arranged the entire spectacle of domination and cast himself as its neutral narrator. A lesson about checking power that exempts its own author from any check has hardly escaped the problem it describes. It has reproduced that problem and hidden the reproduction behind the costume of instruction.

It is fair to ask what the honest alternative would be, since the lesson *StarPower* aims at, that systems and structures shape behavior, is true and worth teaching. The alternative is the one the later chapters develop, and it does not require manufacturing a rigged society in a classroom. A teacher who wants students to understand how stratification works can show them the real thing, the actual data on who holds wealth and power and how it stays concentrated, the documented histories of how hierarchies were built and maintained, the testimony of people who have lived at the bottom of real ones. This material is more accurate than any simulation, because it is the thing itself rather than a chip-game model of it, and it carries none of the simulation's costs, since no child has to be made to feel like a failure and no child has to practice being a tyrant to learn that tyranny exists. The choice the dramatic tradition never quite faces is between teaching about hierarchy and building one. *StarPower* builds one, on the theory that building it teaches about it, and the theory fails for the reasons this chapter has traced, that the people inside the built hierarchy mostly learn to inhabit their positions rather than to question the structure. Teaching about hierarchy, through the real record and the real voices, asks more patience and offers less drama, and it has the advantage of being both true and harmless, which the simulation is neither.

Set the fair case for *StarPower* at its strongest. Its defenders argue that the game delivers an insight no lecture could, the felt knowledge that one's own behavior is hostage to one's position, and that a person who has watched himself become a Square or a Triangle can never again be quite so confident that he would have behaved well in someone else's circumstances. They argue that the rage of the Triangles is instructive, a controlled dose of what exclusion feels like, and that the debrief at the end converts the raw experience into durable understanding. This case is not empty, and the effect it describes is real for some participants. The method is effective because it does generate a genuine and sometimes lasting realization about the power of systems, delivered with an emotional force that abstract instruction cannot match, and because the structural insight at the center of it is true and worth having. The method is not effective because it purchases that insight at the cost of humiliating real people, because the Squares it creates tend to learn the opposite of the intended lesson, enjoying their power rather than questioning it, and because, as the chapters on evidence will show, the durability of the realization past the door of the classroom is far less certain than its champions claim. Faced with a tool that reliably produces a vivid experience and only sometimes produces

a lasting lesson, while reliably inflicting real distress on the people sorted to the bottom, the honest recommendation is to set it down. The insight it offers can be reached by routes that do not require manufacturing an hour of cruelty, and a later part of this book takes up what those routes are.

The claim that the game delivers felt knowledge deserves a closer look, because the phrase hides a question its users rarely ask, which is felt knowledge of what. A participant in *StarPower* does acquire a vivid, embodied experience, and the experience does leave a strong impression. That impression, though, is of a rigged chip game, with its plastic badges and its arbitrary point values and its known ending an hour away, and the question is whether the felt knowledge of that experience transfers to the real hierarchies the game means to depict. An earlier chapter on the absent subject gives the reason for doubt. The Triangle's distress is real, but it is the distress of a person who knows the badge comes off at the end of the period, a bounded discomfort with a visible horizon, and the suffering of actual exclusion has no such horizon, which is most of what makes it suffering. The Square's enjoyment is no different in kind, the pleasure of winning a parlor game rather than the entrenched advantage of a real elite with real stakes. What the participant comes to know in his body is the feel of the simulation, and he is then encouraged to believe he has come to know the feel of the thing simulated, which is a larger and shakier claim. Felt knowledge is powerful precisely because it feels like direct acquaintance, and that is also its danger, because it can deliver a confident sense of having understood an experience the game only gestured at through a proxy. The participant leaves certain he knows what it is like, and the certainty outruns the knowledge.

Shirts's reassurance at the sight of a threatening child was no personal quirk. It was a small instance of the way the whole dramatic tradition judges its own success, and the way earns a name because it explains the tradition's longevity in the face of its failures. These methods are evaluated by their vividness. A facilitator runs an exercise, watches it produce strong emotion, tears or anger or a stunned silence, and reads the strength of the emotion as the measure of the lesson. Elliott pointed to her weeping participants and her grateful letters; Shirts pointed to his threatening Square; the privilege-walk facilitator points to the hush that falls over the room. In every case the evidence offered is the intensity of the experience, and intensity is taken to certify depth, as though a lesson that moved people must have taught them. The inference is natural and wrong. Strong emotion is easy to produce and proves nothing about what was learned, since the same techniques that generate it, the sorting, the licensing, the public exposure, would generate it whether the lesson landed or not. A field that measures itself by how powerfully its exercises feel has equipped itself to detect engagement and disabled itself from detecting failure, because the one signal it trusts, emotional force, is present in the failures and the successes alike. This is why a method can backfire for fifty years and

still command the confidence of intelligent practitioners. They are watching the wrong gauge, and the wrong gauge reads full every time.

There is a distinction that cuts to the bottom of the matter, between a demonstration and a lesson, and *StarPower* sits on the wrong side of it while pretending to sit on the right. A demonstration shows something true to an observer. *StarPower* is a fine demonstration; run it and you will see, reliably, that ordinary people placed in a rigged hierarchy will rationalize their advantage or sink into their disadvantage, and that is a real and interesting fact about human beings, well worth a social scientist's attention. A lesson changes the person who undergoes it, leaving him wiser or better than before. These are different achievements, and a procedure can manage the first while failing the second, which is exactly *StarPower*'s situation. As a demonstration of how hierarchy shapes behavior, observed from outside, the game is informative. As a lesson delivered to the people inside it, the game tends to teach the Squares to enjoy power and the Triangles to feel like failures, which is close to the reverse of what it intends. The designers, watching the demonstration succeed so reliably, concluded that the lesson must be succeeding too, and conflated the value of the spectacle to an observer with its value to a participant. A psychologist could learn a great deal by watching *StarPower* played a hundred times. That tells us nothing about whether the players learned anything worth keeping, and the evidence suggests that many of them learned the wrong thing while their teachers, dazzled by the reliability of the show, recorded a success.

A defender might offer one more distinction in the game's favor, that *StarPower*, unlike Elliott's exercise, is honest, since the rigging is revealed in the debrief and the players learn that the whole hierarchy was an arranged fiction. The honesty is real and it is to the game's credit, yet it arrives too late to do the work asked of it. The deception that matters is the one that operates during play, when the players are told their ranks reflect merit, and that deception is fully live for the hour the cruelty and the despair are being produced. A Square who spends an hour believing, or half-believing, that he earned his place and acting on that belief has practiced the entitlement of the advantaged in real time, and learning at the end that the game was rigged does not reach back and erase the hour of practiced entitlement. A Triangle who has spent an hour feeling like a failure does not have that hour refunded when the facilitator explains the trick. The revelation corrects the players' beliefs about the game while leaving untouched what the game has already deposited in their bodies and habits, laid down while the fiction was still operating. Honesty after the fact beats dishonesty, and it does not convert an hour of manufactured hierarchy into a harmless one. The lived experience happened, and it happened under the lie, and a later confession of the lie changes the record without changing what the experience did.

The most sobering fact about *StarPower* is that it is still for sale. More than half a century after its creation, the game can be ordered with its director's manual and its

chips and its plastic badges, marketed now to corporations and government agencies and schools as a tool for teaching about power, leadership, and diversity, and run in conference rooms and classrooms every week. Its creator built a successor that puts participants in charge of a dysfunctional system and invites them to reform it, a gentler design, and the original endures alongside it, a product with a track record measured in decades and in millions of players. Nothing in that track record includes the kind of outcome evidence that would tell a buyer whether the game leaves participants wiser, less prejudiced, or better at sharing power a month or a year later. What it offers instead is the thing the dramatic tradition has always offered, a reliable and powerful experience, a guaranteed hour of strong feeling that everyone will remember. That the experience reliably produces cruelty in its winners and despair in its losers, and that its educational payoff remains unproven, has not slowed its sale, because the market for these tools runs on the same currency the designers minted, the vivid experience mistaken for the lasting lesson. *StarPower* is the family's patriarch, and its long, profitable life is the clearest sign of how completely the confusion at its origin took hold.

Underneath the specific failures sits the conceptual mistake the whole movement made, and *StarPower* displays it with unusual clarity. The designers confused two kinds of reliability. A simulation can reliably produce an effect, the curdle, the rage, the huddle of the Squares, the despair of the Triangles, and *StarPower* does this as dependably as a chemical reaction. A simulation can reliably produce an education, a change in understanding that survives the room and improves the person, and this is a different and much harder thing, which no amount of dramatic reliability guarantees. Watching a game produce the same powerful scene every time it ran, the designers inferred that it must be teaching the same powerful lesson every time, and the inference does not hold. The reliability of the effect tells you only that the structure dependably moves people through certain emotions and behaviors. Whether anything worth keeping remains a week later is a separate question the spectacle cannot answer, and the temptation to read the one as proof of the other is the central seduction of the dramatic method. A reliable reaction is no proof of a reliable lesson, and an experience that always moves people need not ever improve them.

StarPower is the eye-color exercise grown into a portable product, the same machine in different materials, sorting people by an arbitrary mark, licensing the favored to crush the rest, and reading the resulting cruelty as a job well done. Its designer saw a child threaten another child and felt reassured. That reaction, and the movement that shared it, took a reliable engine of humiliation and sold it to the schools as a reliable engine of insight, and the two have been confused ever since. The confusion did not stay in the classroom. It moved into the privilege walks and the poverty simulations and the oppression tunnels that colonized the campuses and the human-resources departments in

the decades that followed, the descendants of Elliott and Shirts, each one staging a hierarchy and calling the performance education. To those descendants, and to what they ask of the people made to walk through them, the next chapter turns.

CHAPTER FIVE

The Privilege Walk and the Tunnel

The dramatic method did not stay in the third-grade classroom or the social-studies convention where it was born. It grew up, left home, and found work in the two places most eager to hire it, the American campus and the corporate human-resources department, where over the past several decades it has installed itself as a fixture of orientation weeks, diversity trainings, and residence-hall programming. The descendants of Elliott's exercise and Shirts's game are everywhere in those settings, and this chapter examines three of the most common, the privilege walk, the poverty simulation, and the tunnel of oppression. They differ in their props and their topics, and they share a deep structure with their ancestors and with each other, the same staged hierarchy, the same assigned roles, the same authority running the room. What is new in the descendants, and what this chapter is about, is a particular move they all make. Each takes a structural matter, a question about how systems distribute advantage and harm, and converts it into personal theater, a performance about how the individual in the room feels, and in making that conversion each tends to humiliate the people it processes without teaching them much of anything.

Why these two settings, the campus and the corporation, took so eagerly to the dramatic method is worth a moment, because the reasons have little to do with whether the method works. A university orientation office or a corporate diversity department faces a standing demand to address inequality and to be seen addressing it, and the dramatic exercise meets that demand with rare convenience. It is cheap, requiring a room and a facilitator and an afternoon. The exercise is portable, sold as a kit or carried in a binder, run by staff with no special training in the subject it dramatizes. It is memorable, leaving participants with a strong experience they will mention afterward, which reads to an administrator as impact. And it is defensible, a documented program the institution can point to when asked what it is doing about privilege or poverty or bias. Every one of those virtues is a virtue of the exercise as an institutional artifact, a thing that fills a requirement and generates evidence of effort, and none of them is a virtue of the exercise as education. The campus and the corporation adopted these rituals

for the same reason they adopt most things, because they solved an institutional problem, the problem of visibly doing something, and the question of whether the something taught anyone anything was, as it usually is, secondary to the question of whether it could be scheduled, documented, and survived without lawsuit.

Begin with the privilege walk, the most portable of the three and the one a reader is likeliest to have been put through. The format is simple. A group lines up shoulder to shoulder across the middle of a room, and a facilitator reads a series of statements aloud, each tied to an instruction to step forward or step back: if English was your first language, step forward; if a store detective ever followed you because of how you looked, step back; if your parents held a college degree, step forward; if you ever skipped a meal because there was no food at home, step back. Statement by statement, the line that began as a single rank disperses across the room, some participants advancing toward an imagined finish line, others falling behind, until the floor displays a visible spread of accumulated advantage and disadvantage. Then the facilitator stops the exercise and asks the participants to look at where they are standing relative to one another, and the discussion that follows turns on a question put to those left behind, how it felt to watch others move ahead. The exercise descends from the idea of the invisible knapsack of unearned advantage, a real and useful concept from the scholarship on privilege, and its purpose is to make that concept visible by turning it into steps across a floor.

The concept the walk borrows needs restoring to its original form, because the distance between the essay and the exercise is the distance this chapter is measuring. It was given its influential statement in an essay that listed the small daily conditions a white American could count on without noticing, the bandages that matched her skin, the assurance that her race would not be held against her in a routine transaction, the freedom to speak without being taken as a representative of her group. The essay's power lay in its specificity and its restraint. It catalogued ordinary, verifiable facts and let them accumulate into a structural picture, and it asked the reader to notice a pattern in her own life that pointed outward, to a system, rather than inward, to her conscience. The walk takes that essay and stands it up on a gymnasium floor, and in the translation the structural picture collapses back into the personal reckoning the essay had worked to transcend. What was an analysis of a system becomes a tableau of individuals, what was an invitation to notice becomes a command to step forward and be seen, and what was restrained becomes theatrical. The exercise keeps the essay's vocabulary and loses its mind, retaining the word privilege while discarding the careful structural reasoning that made the word mean something other than personal guilt.

Notice what the conversion does to the idea it is meant to teach. The concept the walk derives from is structural, a claim about how a society distributes goods and burdens across groups, about patterns that hold at the scale of millions and across

generations. The walk takes that claim about systems and relocates it inside the individual body, asking each person to embody his own coordinates in the structure by physically stepping forward or back. That relocation feels like clarity and is mostly a loss. A structural fact is a fact about distributions, about how being born into one group rather than another shifts the odds across a whole population, and it is precisely not a fact about any single person’s worth or guilt. The walk, by staging the structure as a row of individuals advancing and retreating, converts the impersonal into the personal and invites exactly the misreading the structural concept was designed to prevent, the reading of advantage as a personal possession to be proud or ashamed of rather than a pattern to be understood. Whoever ends at the front is positioned to feel guilt, the person at the back to feel exposed, and both feelings are about the self, where the structural analysis was about the system. The walk trades an understanding for an emotion, and the emotion it most reliably produces, personal discomfort about one’s place in a visible ranking, is the one least likely to deepen anyone’s grasp of how the ranking came to exist.

This conversion is the signature move of the whole campus-and-corporate genre, a general error that recurs in every exercise this chapter examines. The error is to treat a structural fact as if it were a personal one. Structural facts are claims about populations and systems, about how an institution or an economy or a body of law distributes outcomes across groups over time, and they are established and corrected the way other empirical claims are, by evidence about aggregates. Personal facts are claims about an individual, his character, his choices, his feelings. The two live at different levels, and a great deal of careful social thought consists precisely in keeping them apart, in seeing that a pattern across millions is not the sum of millions of personal virtues and vices. The dramatic exercise collapses the levels on purpose. It takes the structural claim and renders it as a personal experience, a step on a floor, a failed budget, a walk through a haunted corridor, on the theory that the personal rendering will make the structural claim vivid. What the rendering does instead is convert a question of systems into a question of selves, and the conversion has a political cost its designers rarely count. A person who has been taught to experience structural inequality as personal advantage or personal guilt has been equipped to feel about himself and disequipped to think about the system, and feeling about oneself, however intense, builds neither the understanding nor the will required to change a structure. The genre produces guilt where it promised analysis, and guilt is famously poor fuel for the patient, collective, unglamorous work that altering a real structure demands.

The humiliation the walk inflicts comes in two forms, depending on who is in the room. In the common campus and corporate version, the participants are a relatively comfortable group, and the disadvantages the statements name belong mostly to people not present, which means the walk is performed on behalf of absent others, sometimes

quite literally, since some versions hand participants fictional identities to walk as. That version is personal theater with no subject in it, a room of the advantaged performing a choreography of disadvantage that none of them live, and its humiliation is the mild and self-regarding discomfort of the guilty conscience. The second form is worse. When the room is mixed, when some of the people in the line do live the disadvantages the statements name, the walk exposes them, publicly and by design, sorting them to the back of the room in front of everyone and then asking them, in the debrief, to describe how it felt to be there. A person who lives that disadvantage, who is poor, who is marked, who is watched in stores, is made to enact it as a performance for the education of the comfortable, and is then invited to narrate his exposure for the group. To ask someone who lives at the back of a real hierarchy how it felt to be sent to the back of a pretend one is a cruelty dressed as a lesson, and it is the kind of cruelty this method specializes in, the conversion of a real person's real pain into a teaching aid for an audience he did not ask to instruct.

Two features of the walk's choreography sharpen the point. The first is the ritual assurance, offered by almost every facilitator before the exercise begins, that the room is a safe space and that no one will be made to share more than he wishes. The assurance sits oddly atop an exercise whose entire mechanism is public exposure, whose lesson is delivered precisely by making people's positions visible to the room, and the contradiction is rarely noticed by the people running it. An exercise cannot both protect a person from exposure and teach by exposing him, and the safe-space framing functions less as a real protection than as a permission slip, a phrase that lets the facilitator proceed with the exposure while feeling that he has cared for the exposed. The second feature is the prize. In a common version, the participants are told that a reward waits at the front of the room and that everyone is competing to reach it, a detail meant to dramatize that the advantaged win the race. What it dramatizes instead is the exercise's own logic, that the point is to sort people into winners and losers and to let the losers watch the winners collect. A walk that ends with the front of the room claiming a prize the back of the room was never positioned to reach has staged, in miniature, the exact humiliation it claims to be teaching against, and has done so to a group of people who came to a workshop and found themselves ranked.

The poverty simulation moves the same logic from a line on a floor to a role-playing game with a script. The most widely used version, developed by a state community-action network and sold to schools, agencies, and companies across the country, runs like this. Participants are assigned to fictional families, each handed a card describing its situation, a single mother with three children and no car, an elderly couple choosing between rent and medication, a family one paycheck from eviction. The room is set up with tables around the edges representing the institutions of poor people's lives, the employer, the bank, the pawnshop, the welfare office, the school, the police, staffed

by volunteers. Players then attempt to survive a simulated month, compressed into four fifteen-minute weeks, by shuttling between the tables to pay bills, cash checks, keep jobs, and feed their pretend children, while the structure of the game makes survival nearly impossible. The designers are emphatic that the exercise is not a game, and the insistence is telling, since the thing they are denying is exactly what it resembles, a board game of poverty in which the comfortable roll the dice for an hour. Its stated aim is admirable and aligned with this book's own emphasis, to move participants away from blaming poverty on the poor and toward seeing it as a structural condition, a failure of systems rather than of character.

The poverty simulation is the gentlest member of the family, and it earns a more careful verdict than the others. It does not license cruelty the way *StarPower* does, and it does not sicken people the way the tunnel does, and the evidence shows that it can shift attitudes in the short term, softening the reflex to blame the poor for their poverty. That is a real result and worth acknowledging. The trouble is twofold, and both troubles are by now familiar. The first is the switch-off. A participant who spends an hour failing to make a pretend budget work knows the whole time that the hour will end, that the pretend children are paper, that he will drive his real car home to his real solvency, and the felt knowledge he carries away is the knowledge of an afternoon's frustrating game, not of the grinding, horizonless condition the game stands in for. The second is the re-centering. The exercise teaches poverty by making the comfortable person the protagonist of a poverty story, casting him as the one who struggles and feels and learns, with actual poor people present only as the absent originals his fictional family is modeled on. Its emotional center is the comfortable participant's experience, his stress, his frustration, his eventual relief, and a method that teaches about the poor by making a well-off person the hero of the afternoon has quietly kept the poor in the position they occupy everywhere else, offstage, spoken for, the raw material of someone else's growth. Even at its gentlest, the family converts a structural subject into a personal experience and measures success by how the comfortable felt.

The makers' insistence that the simulation is not a game rewards a closer look, because the denial reveals the discomfort at the heart of the form. They are right to sense that calling it a game would cheapen it, and they are wrong about why, because the structure they have built is a game in every formal respect, with roles, rules, resources, obstacles, and a clock, and the players experience it as one, competing against the system and against the ticking weeks to keep their pretend families afloat. The denial is an attempt to hold off an obvious objection, that turning poverty into a board game for the comfortable is a strange way to honor the poor, and the attempt fails because the objection is correct. There is something off about learning the texture of a desperate life by playing at it for an hour and then going home, and the makers feel the wrongness keenly enough to deny the word that names it. As for the results, the honest picture is of

a tool that moves attitudes in the short term and loses its grip soon after. Studies find that participants leave more inclined to see poverty as structural and less inclined to blame the poor, and other studies find that those gains fade, that the effect measured a few weeks or months later has largely washed out. A method that produces a real but evaporating change in attitude, at the cost of casting the comfortable as the heroes of a poverty story and the poor as the absent originals, has bought little at a price worth questioning, and the genre's habit of measuring success at the moment of maximum feeling, before the evaporation, is the same habit that let *StarPower*'s designers mistake a reliable effect for a lasting lesson.

The tunnel of oppression is the most theatrical of the three and the most disturbing by intent. It began in the residence halls of a midwestern university in the early nineteen-nineties, modeled on the Museum of Tolerance in Los Angeles, and it spread from there across American campuses, where it is now an annual fixture run by housing offices, multicultural centers, and student-life staff. The format is a walk-through. Visitors move in small groups through a series of rooms, a dozen or more, each staging a brief scene of some form of oppression, racism in one, sexual violence in another, homophobia in a third, then poverty, disability, religious persecution, and on down a list that organizers say could fill a new tunnel every week. The scenes are built and performed by student volunteers and are designed to be visceral, to assault the senses and provoke strong emotion through graphic enactment, recorded sound, and confrontation. In one common scene, actors scream a slur into the face of a student playing its target; in another, participants act out a true account of an assault. Groups are led through the corridor, scene after scene, and then gathered at the end for a facilitated discussion meant to convert the shock into reflection and the reflection into a commitment to social justice. On many campuses the event is co-sponsored by the counseling center, an arrangement whose logic deserves more attention than it usually gets.

That a tunnel needs its counseling center on standby tells you most of what its critique requires. The exercise is built to upset people, and it succeeds, by the testimony of its own participants and guides, who describe leaving scenes sickened and frightened, and in some cases unable to continue, having to skip rooms because the actors had gotten to them. Its intended theory is that the upset is productive, that being made to feel the horror of oppression viscerally will move the comfortable to understand and to act. The theory mistakes a strong reaction for a strong lesson, the same error the whole tradition makes. Being sickened by a staged scene of violence teaches a person that the violence is sickening, which he knew, and floods him with an emotion that research on fear and shock suggests is more likely to produce avoidance and shutdown than understanding and engagement. And the absence of the subject is total. The tunnel is built explicitly for the sheltered, for the students in the campus bubble who, organizers

say, have never encountered these realities, and the realities themselves are supplied by student actors performing the suffering of people who are not in the room. The comfortable watch a dramatization of an absent group's pain, feel appropriately terrible, process the feeling in a debrief, and leave having undergone an emotional event that stood in for an education. For any participant who has lived what the tunnel stages, who has been assaulted or slurred or hunted, the experience offers something worse than nothing, a sudden ambush by his own worst memory, staged for the instruction of his sheltered classmates.

The tunnel's lineage bears tracing, because what it borrowed and what it discarded are both revealing. Its model was a memorial institution, a museum built to confront visitors with the machinery of historical atrocity, and a serious memorial does difficult and honorable work, presenting documented history, the testimony of survivors, the record of what was done, to an audience that comes to learn. The campus tunnel kept the memorial's intensity and dropped its substance. Where the museum presents the real record and the survivor's own voice, the tunnel presents student actors performing invented scenes, and where the museum trusts documented horror to do its work on a willing visitor, the tunnel manufactures horror to do its work on a captive tour group. The memorial is anchored to a real history and a real subject; the tunnel floats free of both, a sequence of affecting tableaux assembled to produce an effect. Some campuses have gone further still and built the tunnel into a for-credit course, so that students earn academic standing for staging the suffering of others, a development that completes the form's drift from education toward production. And the harm runs in two directions at once. The sheltered visitor the tunnel is built for receives a shock mistaken for an insight, while any visitor who carries the real version of a staged scene, the survivor of the assault the actors are performing, the person who has lived the slur being screamed, is led without warning into an ambush, his memory rebuilt as a teaching prop for a classmate's afternoon. A method that injures the people it depicts in order to move the people it addresses has its priorities exactly backward, and the standby counseling center is the institution's quiet admission that it knows.

There is a word for what these exercises make of suffering, and the word is spectacle. To witness another person's pain is one thing, an act that can deepen understanding and oblige a response. To consume it as a staged performance arranged for one's own benefit is another, and the line between the two runs through every exercise in this family. The privilege walk arranges the disadvantaged into a tableau for the room to view. A poverty simulation stages a desperate family's month as an afternoon's diversion. The tunnel builds a corridor of enacted atrocities for groups to tour. In each, the suffering of an absent or silenced group is converted into a show, with the comfortable as its audience and their feelings as its measure, and the conversion does something subtle and corrosive to those feelings. It teaches the spectator that other

people's pain is a thing he is entitled to be moved by, a content provided for his growth, an experience he can enter and exit at the door. The emotion this produces wears the clothing of compassion and works more like consumption, a sentiment aroused on schedule by a performance designed to arouse it, and a sentiment of that kind is closer to the manufactured feeling of an audience watching a sad film than to the obligation a person incurs by encountering a real other in distress. The genre flatters the spectator that he has felt for the suffering when he has mostly felt himself feeling, and the difference matters, because the manufactured sentiment fades with the performance that produced it and leaves no debt behind, while the real encounter with a real other tends to leave a person changed in a way that asks something of him afterward.

Each of these exercises leans, in the end, on its debrief to redeem it, the closing discussion in which a facilitator gathers the shaken participants and helps them turn what they felt into what they should think. The debrief carries the same impossible burden here that it carried in the simulation game. It is a verbal event asked to correct a visceral one, a conversation about systems offered to people whose bodies have just been taught a lesson about themselves, and when the two disagree, the body usually wins. The walk has just taught a participant to feel personally exposed or personally guilty, and the debrief asks him to convert that feeling into a structural understanding it did nothing to build. A tunnel has just flooded its participants with shock, and the debrief asks them to metabolize the shock into reasoned commitment in the few minutes before they leave. The conversion the debrief attempts is the reverse of the one the exercise performed, an effort to turn the personal back into the structural after the exercise spent its whole length turning the structural into the personal, and a short talk at the end is no match for the experience that preceded it. The debrief is where the genre keeps its conscience, the place it points to when challenged, and its faith in that closing conversation is the faith of a tradition that needs to believe its powerful experiences are also good lessons and lacks the evidence to know.

A defender backed into a corner will fall to the last line, that whatever their flaws these exercises at least raise awareness and start conversations, and that something is better than nothing. The claim deserves an answer, because it is the one most people reach for, and the answer is that the something on offer is often worse than nothing on the two measures that matter. Awareness is the weaker half of the claim. The exercises raise feeling, which is not the same as awareness, and the feeling they raise, personal guilt or secondhand shock, can crowd out the structural understanding that genuine awareness would require, leaving a participant more emotional and no better informed about how the system he is now upset about works. The conversation half fares no better. A conversation that opens from a place of manufactured guilt or defensiveness, with the participants sorted into the exposed and the ashamed, is a charged and guarded exchange rather than the open one the defender imagines, the kind that hardens positions

rather than softening them, and the research on the backlash these methods can provoke gives the worry teeth. Then there is the cost the better-than-nothing claim ignores entirely. Nothing does not humiliate the disadvantaged person in the room, does not ambush the survivor with a staged version of his trauma, does not consume an afternoon and a budget that might have funded the patient teaching that works. The choice was never between these exercises and nothing. Set against the better thing the time and money could have bought, the dramatic ritual reveals itself as a costly substitute for the education it crowds out rather than a modest good.

The alternative this book keeps pointing toward deserves to be stated plainly here, since the campus and the corporation, having reached for the dramatic exercise, tend to assume there is nothing else to reach for. There is, and it is older and better attested than any of these rituals. The way to reduce prejudice and to teach people about a group they fear or misunderstand, the most reliable finding in the entire research literature, is sustained, cooperative contact with members of that group under conditions of equal standing, the slow work of knowing actual people rather than the fast work of performing their conditions. The way to teach about a structure is to teach the structure, through the data and the history and the testimony that show how it operates and what it does, with the people who live inside it present as the authorities on their own experience. Neither alternative is dramatic, and that is part of why the institutions pass them over, since neither produces the strong afternoon that reads as impact. Both carry the homely advantage of working, and of working without humiliating anyone, exposing anyone, or ambushing anyone, which is more than the dramatic exercises can claim. The whole family, from Elliott’s classroom to the campus tunnel, rests on a single bad bet, that the shortcut of manufactured experience can substitute for the long road of real knowledge and real encounter, and the bet keeps failing in the same way, producing a powerful feeling that masquerades as a lasting lesson. Strip the drama away and what remains is the unglamorous truth the genre was invented to avoid, that understanding other people takes time, takes their actual presence, and cannot be installed in a single staged hour no matter how intensely that hour makes anyone feel.

Consider what a participant carries out of one of these rooms, because the takeaway is revealing and it is rarely the intended one. He does not carry a structural analysis, which the exercise did not supply, or a changed pattern of behavior, which a single afternoon cannot install. What he carries is a story, the memorable tale of the day he did the privilege walk or walked the tunnel or went broke in the poverty game, an anecdote he can produce on cue as evidence that he has confronted these matters and grown. The story is the product, and it functions as a credential. Having undergone the experience, the participant can now count himself among the people who take inequality seriously, who have done the work, who have felt the discomfort and earned the standing it confers, and the credential is durable in a way the understanding is not, surviving long

after whatever was learned has faded. This is the same counterfeit described earlier in the book, issued now by a different office. The exercise hands the comfortable a certificate of conscience, a documented experience they can cite to themselves and to others, and the certificate does its own quiet harm, because a person who believes he has already confronted his privilege or understood poverty or witnessed oppression is a person who has closed the question, and a closed question is the opposite of the open and ongoing attention that living rightly alongside other people requires. The genre's most lasting effect may be the complacency it licenses, the sense of a debt discharged by an afternoon, when the debt it gestured at can be discharged by no afternoon at all.

There is a deeper irony in a ritual that claims to decenter the self and works by placing the self at the center of everything. The whole point of structural analysis is to look away from one's own conscience and toward the system, to understand that the important facts are not about how any individual feels. The dramatic exercise reverses that gaze and trains it back on the self, asking the participant to attend, for an entire afternoon, to his own position, his own advantage, his own guilt, his own discomfort, his own growth. A practice that fills a room with people examining their own feelings about inequality has turned their attention inward rather than outward, onto the most absorbing subject any person has, which is himself, and away from the people who suffer the inequality. The marginalized, once again, are the occasion for this self-examination rather than its beneficiaries, present as the prompt that lets the comfortable contemplate their own souls. A lesson in looking outward that operates by looking inward has misunderstood its own assignment, and the misunderstanding is no accident, because looking inward is what the dramatic method does, and feeling about oneself is the only product it reliably delivers. The participant leaves having performed a great deal of feeling about himself and having learned little about anyone else, which is the genre's recurring result wearing its most flattering disguise, the disguise of conscience. And the people the exercise was supposedly about leave with nothing, because they were never the point, only the material, and a method that turns the suffering of others into an occasion for the comfortable to feel things about themselves has built its self-absorption into its design. The decentered self the rituals promise is a self that has spent an afternoon thinking about nothing except itself.

Step back from the three and the family resemblance is plain, and so is the inheritance from Elliott and Shirts. The anatomy runs through all of them, in full in some and in part in others. A role or position is assigned by the facilitator, the step on the floor, the family card, the place in the tour group. The authority is unrefusable, vested in the person running the exercise, who controls the script, the pace, and the meaning. Discomfort is licensed and administered, the exposure of the walk, the engineered failure of the simulation, the assault of the tunnel. An audience is always present, the room watching the line, the families watching each other, the group moving

through the scenes together. And each ends with the debrief, the conversion ritual in which the manufactured experience is supposed to become a durable lesson. The inversion is present in each, since each teaches against a structural injustice by staging a small version of it and making people undergo it. The iatrogenic risk is present, since each can produce the opposite of its aim, guilt that curdles into defensiveness, tourism that confirms distance, shock that prompts avoidance. And the absence of the subject runs through all three, the disadvantaged and the oppressed appearing as roles, cards, and performed scenes, almost never as themselves with a voice of their own. The descendants kept the ancestors' engine and changed the décor.

The fair case for these exercises is the case for experience over abstraction, and it deserves a hearing. Their defenders argue that a lecture on privilege or poverty or oppression slides off the comfortable, that the structural facts remain inert until a person feels something, and that the walk and the simulation and the tunnel supply the feeling that makes the facts stick. There is a partial truth here, the same partial truth that powered the whole simulation movement, that emotion aids memory and that a felt experience is harder to dismiss than a statistic. The method is effective because it does generate strong feeling and does, for some participants in the short term, soften a hostile attitude or open a closed mind, and because the structural facts it points at are real and worth teaching. The method is not effective because the feeling it generates is feeling about the self rather than understanding of the system, because the experience it supplies is of a proxy that switches off rather than of the condition it represents, because it tends to humiliate the people it claims to serve when they are present and to perform on their behalf when they are absent, and because, as the evidence chapter will show, the attitude changes it produces are mostly shallow and short-lived. Offered a set of tools that reliably manufacture an emotional event and only unreliably produce a lasting understanding, while risking real harm to the vulnerable, the honest recommendation is to teach the structural facts as what they are, and to stop staging them as personal theater.

The structural facts these exercises gesture at can be taught directly, and taught better, by the means that respect their nature. Privilege, poverty, and oppression are patterns in the world, and the way to understand a pattern is to study it, through the data that show how it works, the history that shows how it was built, and the testimony of the people who live inside it speaking in their own voices. That kind of teaching keeps the subject structural, which is what the subject is, and keeps the real people present as authorities on their own lives rather than as absent originals for a comfortable person's performance. It asks more of a teacher than ordering a tunnel kit or reading a list of step-forward statements, and it offers less drama, and it has the advantage of teaching the thing it claims to teach without humiliating anyone to do it. The descendants of Elliott and Shirts took a set of structural truths, converted them into personal theater, and

produced rooms full of guilty, frightened, or tearful people who had felt a great deal and learned little. There is, however, a version of this theater in which the conversion fails completely, in which the staged suffering meets a real subject who will not play along, and the whole machine breaks down in public. That is the auction, and it is where the next chapter goes.

CHAPTER SIX

The Auction

This book has argued that the dramatic exercise depends on an absence. The lesson works, in the sense of running smoothly and producing its glow of significance, only when the people it claims to be about are not in the room, present as a proxy, a fictional family, a performed scene, a collar worn by a child of the wrong group standing in for a child of the right one. The claim invites a test. If the absence of the subject is what lets the ritual proceed, then the presence of the subject should break it. This chapter runs that test, using the cases in which teachers did the unthinkable thing the logic of the method always pointed toward, and assigned the real subject the despised role, casting actual Black children as slaves to be sold and actual Jewish children as victims to be gassed. What looks like a more intense version of the lesson is its collapse. The presence of the real subject detonates the pretense, a sudden break into real harm, real trauma, and real scandal, and the rupture is the most revealing thing the genre produces, because it shows in the open what the gentler exercises keep hidden, that the method was always a machine for manufacturing degradation, tolerable only as long as the degradation fell on a stand-in.

The cases are real, they are documented, and they recur with a regularity that should disturb anyone who believes them rare. In April of 2011, a fourth-grade teacher at an elementary school in Norfolk, Virginia divided her class by race as part of a lesson on the Civil War, sent the Black and mixed-race children to one side of the room, and let the white children take turns buying them. The school learned of it only when parents complained, and the principal wrote to families calling the activity inappropriate while allowing that the teacher had meant well. The same spring, near Columbus, Ohio, another fourth-grade class was divided into slaves and masters, with a ten-year-old Black child assigned to be sold. In 2017, a mock auction in a fifth-grade class in Maplewood, New Jersey produced the same outrage. In March of 2019, at a private school in Bronxville, New York, a fifth-grade teacher told the African American children in two classes to raise their hands, leave the room, and return with imaginary chains placed on their necks, wrists, and ankles, whereupon she conducted a mock

auction of them before their classmates. Weeks later, at an elementary school in Watertown, New York, a fourth-grade teacher had white students bid on their Black classmates, instructed the children sold to address the buyer as master, and told them they would take the buyer's surname. The dates change and the script does not.

That list could be extended, and its persistence is part of the indictment. The same years produced milder variants of the same impulse, fifth-graders asked to dress as Civil War figures for an experiential day, students walked through reenactments of the Underground Railroad in which the children playing the escaping enslaved were chased and hunted. As recently as this decade, parents in Chicago and its suburbs have reported slavery reenactments in their children's classrooms, which means the practice has survived every warning issued against it. That survival is the point. These exercises do not persist because teachers are cruel or ignorant of the warnings, since the warnings are everywhere and the teachers are, by all evidence, decent people trying to teach a hard subject. They persist because the dramatic method is the default the culture hands a teacher who wants a lesson to land, the first tool that comes to hand, reached for precisely by the conscientious teacher who feels that a lecture on slavery is not enough and that the children must somehow be made to feel it. The reflex to dramatize is the disease, and it is so deeply seated that it survives decades of documented harm, producing the same scene in a new town every year or two, each teacher apparently unaware that she is the latest in a long line, each community reacting as though the thing had never happened before. A practice that keeps recurring despite repeated condemnation has become a structural feature of how the culture teaches, the recurrence itself the symptom that gives the diagnosis away.

What stays the same across these cases is the argument itself. The caste is assigned by birth, and this time the caste is real, the actual race of an actual child rather than the color of his eyes or the draw of a chip. The authority is unrefusable, a teacher running a required lesson, and no fourth-grader sent to the auction block had any standing to decline. Cruelty is licensed by the form, the auction demanding that some children be sold and others buy. An audience watches, the class observing its own members sorted and priced. Every structural element this book identified in the eye-color exercise is present, intensified by the substitution of a real despised identity for a pretend one, and that single substitution is what turns the lesson into an injury. A child playing a blue-eyed inferior for a morning is enacting a fiction he will shed at recess. A Black child made to stand on a pretend auction block while white classmates bid on him is enacting the specific, documented degradation of his own ancestors, in front of the people he must sit beside tomorrow, on the instruction of an adult he cannot refuse, and there is no recess at which he sheds the identity that placed him there. The as-if collapses. What remains is a Black child who has been sold, in a classroom, in front of his peers, and no debrief reaches the place where that lands.

Dwell on what the exercise asks of the particular child, because the abstraction of the word degradation can hide the concrete thing. The Black child at the front of the room meets slavery as something done to his body rather than as a topic, made to perform his own subjugation, to stand while peers who are mostly white assign him a price, to answer to the name of an owner, to feel in his body the posture of the sold, and to do all of this on the authority of the one adult in the room he has been taught to trust and obey. The betrayal runs in several directions at once. His teacher, the figure charged with his protection, is the author of his humiliation. His classmates, the children he must face the next morning, have watched him priced and bought. The history being taught is his own, the actual fate of his actual ancestors, which means the lesson hands him, in front of an audience, the message that the room is rehearsing the proper order of things. Whatever the teacher intends, the child cannot help receiving some version of that message, because the exercise stages it directly, his body in the position of the slave and his classmates in the position of buyers, and a child has no way to hold the framing at the distance an adult might. The morning ends and the topic closes, and the child carries out of the room an experience that has no clean closing, because he cannot return the identity that placed him on the block. This is the harm the gentler exercises distribute in smaller doses to stand-ins, delivered here at full strength to the one child least able to refuse it and most marked by it afterward.

The aftermaths tell the rest. In the Bronxville case the state attorney general opened an investigation and concluded that the exercise had a deeply harmful effect on all the students present and on the African American children above all, the school fired the teacher, agreed to hire a diversity officer and broaden its faculty, and called in a counseling center to address what it described as the need for healing. The pattern of the response is the thing to mark. The institution reaches for counseling after the fact, the same reflex the oppression tunnel revealed, the standby therapist who is the clearest admission that the exercise was expected to wound. That teacher was later hired by a nearby public school and, when parents recognized the name, sent home. None of these teachers, by the available record, set out to harm a child. Each meant to teach the evil of slavery and reached for the dramatic shortcut the culture had taught them to reach for, and each produced, instead of understanding, a Black child in tears and a community in revolt. The gap between what they intended and what they did is the gap this whole book has been measuring, and in these rooms it is at its widest.

The same logic governs the other great atrocity the schools insist on staging. In December of 2021, at an elementary school in Washington, D.C., a staff member leading a library class directed a group of third-graders to act out scenes from the Holocaust. The children were made to dig imaginary mass graves, to simulate shooting one another, to pretend to be packed onto trains, and to mime suffocating to death in a gas chamber. The instructor cast one boy as Adolf Hitler and told him to pretend to kill

himself at the end, as Hitler did; the boy was Jewish. She is reported to have told the children that the Germans had acted because the Jews ruined Christmas, an antisemitic claim offered to eight- and nine-year-olds as history, and to have instructed them to keep the activity secret, which they did until they told their regular teacher. The children came home sobbing. The district placed the staff member on leave, called the lesson unapproved, apologized to the families, and routed the children to its equity and counseling team, and the major Jewish civil-rights organizations met with the school system in the aftermath. A boy of eight, Jewish, made to play Hitler and mime his suicide, is the purest illustration this book can offer of what happens when the subject is present.

Two features of the Washington case deserve to be drawn out, because each turns a bad lesson into something darker. The first is the instruction to keep it secret. That staff member told the children not to tell anyone what had happened in the library, and a teacher directing children to conceal an activity from their parents has crossed from poor pedagogy into the grammar of abuse, since secrecy imposed on a child by an adult is the signature of harm meant to escape adult scrutiny. The children, to their credit, broke the secret, carrying the story to their regular teacher, which is the only reason it came to light. The second feature is the antisemitic claim folded into the lesson, the assertion that the Germans had murdered the Jews because the Jews had ruined Christmas. A lesson nominally against antisemitism delivered, as part of itself, an antisemitic libel to a room of eight-year-olds, and the libel came as a line inside the exercise rather than a slip outside it, offered as explanation. Hold those two features together with the casting of a Jewish child as Hitler, and the case stops looking like a lesson that went wrong and starts looking like the thing the lesson was supposed to oppose, reproduced in miniature inside a school, with real Jewish children as its targets. The exercise meant to teach children to recognize and reject antisemitism instead subjected Jewish children to it, under instruction to stay quiet, which is as exact an inversion of an educational aim as this book will find.

The DC case was florid and unusually cruel, and the quieter versions make the same point with less noise. In 2011 a sixth-grade teacher at a school in South Carolina built a multi-day simulation of nineteen-thirties Germany, dividing the class into Nazis and Jews. The children cast as Nazis were given the run of a privileged class, seated in front, flattered as bright, treated as the teacher’s favorites. Those cast as Jews ate in silence in the hallway, sat on the floor at the back of the room, and wore stars pinned to their clothing, and a sign on the cafeteria door announced that Jews and dogs were not admitted. The design is Jane Elliott’s, transposed from eye color to ethnicity, the same partition into a flattered group and a degraded one, the same badges, the same authority arranging it all. This version was reported to run smoothly and to draw little complaint, which tells its own story, since a simulation that assigns a real Jewish child to the starred

and excluded group, and runs smoothly, has simply found a child willing or compelled to absorb the degradation quietly. Smoothness shows only that the hurt stayed silent.

The institution with the most authority on this subject has reached a clear verdict, and the verdict is the one this book has been arguing toward. The United States Holocaust Memorial Museum, in its guidance for teachers, states plainly that simulating the experience of the Holocaust is pedagogically unsound, that even a careful simulation leaves students with the false impression that they now know what it was to suffer the event or to take part in it, and that teachers should rely instead on primary sources and the testimony of survivors. Read that reason again, because it is the switch-off this book described, named by the foremost authority on the subject. The student who has spent an hour pretending to die in a gas chamber walks away believing he knows something he does not, the feel of a horror he sampled and survived, and the false knowledge is worse than ignorance because it is confident. The most carefully designed version on record, studied by a scholar of Holocaust education, had the teacher play the Gestapo so that no child was cast as a Nazi, and ran for seven weeks, and even produced measurable gains in what the students knew. That careful version still asked every child in the class, including any actual Jewish child, to live for seven weeks as a hunted Jew under a teacher playing the secret police, and it still left its students, by the museum's reasoning, mistaking a simulation for the thing itself. The best version of the exercise fails for the same reason the worst version does, and only the volume of the failure differs.

The museum's recommended alternative repays attention, because it marks the exact line between a method that teaches and one that traffics in feeling. A survivor's testimony and a simulation of the Holocaust both put a student in contact with the event, and they do so in opposite ways. The simulation removes the survivor and installs the student in his place, asking the student to imagine from the inside an experience the student has not had and cannot have, and rewarding him with the sensation of having had it. Testimony does the reverse. It keeps the survivor present and the student in his proper place as a listener, a witness to a witness, learning the event from the one person entitled to describe it from within. The difference is the difference between consuming an experience and receiving a testimony, and it maps precisely onto the absence and presence of the subject this book has tracked from the first page. A simulation works by the subject's absence, since a real survivor in the room performing his own gassing for an audience is an obscenity no one would stage, which is exactly why the simulation replaces him with a child. Testimony works by the subject's presence, by handing the authority over the experience back to the person who lived it and asking everyone else to be quiet and attend. The student who hears a survivor describe the selection at the ramp learns something true and is changed by it, and he leaves knowing that he has heard a witness rather than become one. The student who mimes death in a pretend gas chamber leaves believing he has touched the thing itself, and the belief is false, and the museum,

which has spent decades on the question, says plainly that the second student has been taught less than the first and harmed more.

These cases give the book's first principle its plainest demonstration, the principle the next part will argue in full, that the dramatic method is the oppressor's own. In the gentler exercises the inversion is a structural resemblance, the facilitator borrowing the oppressor's tools of sorting and licensing and exposure. In the auction and the Holocaust reenactment the inversion is no longer a resemblance, because the teacher performs the oppressor's acts directly rather than borrowing his method. A teacher who conducts an auction of Black children is doing, in a classroom, the specific thing the slaveholder did, pricing and selling human beings by race. A teacher who plays the Gestapo, or who sends a Jewish child to wear a star and sit on the floor, is doing the specific thing the Nazi did, marking and degrading a people by descent. The method has stopped pointing at the historical oppression and started reproducing it, on the actual descendants of its actual victims, in a room arranged for the purpose. This is what it means to say the method is the oppressor's. The slaveholder and the Nazi also sorted people by birth, also wielded an authority their victims could not refuse, also licensed cruelty and staged it for an audience, and the teacher who builds a lesson on those same four moves has built, whatever the banner over the door, a small working model of the thing the banner condemns. These cases are the model running at full power, with the historical roles filled by the historical groups, and they show that the resemblance the book has been tracing was exact all along, and that only the absence of the real subject kept it from showing.

The comparison between these rooms and the gentler ones amounts to something close to an experiment, and the conclusion in those terms is hard to escape. Hold everything constant except one variable. The structure is the same across all of them, the sorting by category, the unrefusable authority, the licensed cruelty, the watching audience, the closing reversal. Intent is the same, a teacher meaning to teach against injustice by making it felt. The framing is the same, the language of empathy and experiential learning and lessons that stick. The one thing that changes between a privilege walk and a mock auction is whether the despised role is filled by a stand-in or by a real member of the despised group, and that single change is what separates a room that produces a glow of significance from a room that produces a traumatized child and a fired teacher. When one variable accounts for the entire difference in outcome, that variable is the cause, and the variable here is the presence of the real subject. The gentler exercises relocate the harm the auction inflicts rather than avoiding it, moving it onto someone positioned to absorb it quietly, a volunteer, a stand-in, a child of the wrong group performing the wrong group's pain. The auction removes that buffer and lets the harm fall where the structure always aimed it, on the real subject, and the explosion that follows is the structure's true output made visible. Everything this book has claimed

about the gentler rituals is confirmed by the louder ones, because the louder ones are the same machine with the muffler removed.

The second of the book's principles, that these exercises produce the disease they mean to cure, also reaches its clearest form here. In the gentler cases the iatrogenic harm is partly hidden, a guilt that hardens into defensiveness, an empathy that curdles into distance, effects that take time to show and are easy to deny. In the auction and the reenactment the cure and the disease are the same event, occurring at the same moment, impossible to separate. A lesson against racism that consists of auctioning a Black child is a racist act in its execution, whatever it is in its intention. A lesson against antisemitism that consists of marking Jewish children and feeding them an antisemitic libel is an antisemitic act in its execution, whatever the banner says. There is no lag here between the well-meant cause and the harmful effect, no window in which the lesson does its good work before the side effect appears, because the harmful effect is the lesson, performed in real time on real children. The exercise produces the thing it opposes as its content rather than merely risking it, with the opposition printed on the label. This is iatrogenesis with no delay and no disguise, the disease administered as the treatment, and it is the reason the cases feel so much worse than a simple failure. A lesson that merely failed would leave the children no worse than before. These lessons leave Black children freshly degraded and Jewish children freshly attacked, in the name of teaching them that degradation and attack are wrong, which is the cruelest joke the method plays.

The strongest defense of these exercises must be stated, even here, because the book owes every method its best case before judging it. That defense is that some horrors are so far outside ordinary experience that no lecture and no document can convey them, and that only by living a version, however partial, can a sheltered student begin to grasp what slavery or the Holocaust was. The scholar who studied the seven-week Gestapo simulation lends it some support, since her students did come away knowing more than they had. By this account the method is effective, because it can produce engagement and even knowledge that drier instruction sometimes fails to produce, and because the felt weight of an experience can fix a historical horror in a young mind that a paragraph would not. The method is not effective, and the case against it is decisive, because the engagement it produces is purchased by harming the children the lesson claims to honor, because the knowledge gained can be gained by testimony and documents without that harm, because the felt experience is a false one that leaves students believing they know a horror they have only sampled, and because, in the form these cases take, the lesson against an evil is itself an instance of that evil. Set the modest knowledge gains of the best-designed version against the trauma of the typical one, and against the existence of a method that teaches the same history without selling or gassing a single child, and the recommendation writes itself. The honest course is to abandon the simulation entirely

and teach these histories the way their own scholars and their own museums beg teachers to teach them, through the records and the voices of the people who endured them.

A word is owed to the children, who tend to disappear from these stories once the news cycle moves on. The institutional response follows a fixed and revealing sequence. A school expresses shock, places the teacher on leave or fires her, issues an apology, brings in counselors, announces new training, and within a week or two the matter is closed from its side, a crisis managed. The child's side of it has no such tidy ending. A boy who was sold at auction in front of his class, or made to mime his own death in a gas chamber, does not have the experience annulled by the teacher's dismissal, and the apology addressed to the families does not reach into the memory and remove what was laid down there. The asymmetry between the institution's quick recovery and the child's slow one is itself part of the harm, because it repeats in miniature the larger pattern these lessons claim to oppose, the comfortable moving on while the marked carry the weight. Firing the teacher, the standard remedy, addresses the symptom and misses the disease, since the next teacher in the next town is reaching for the same method with the same good intentions, and the structure that produced the harm remains the default. The children are not helped by the ritual of accountability that follows their exposure, a ritual aimed mostly at restoring the institution's standing, and they are owed something the apology never provides, which is the assurance that the method that harmed them has been recognized for what it is and set down for good. That assurance the institutions have never given, because they have never diagnosed the method, only deplored its worst instances while leaving the engine in place.

A careful reader might press a question at this point. The book has insisted that the subject should be present, and these cases are catastrophes of the subject present, so has the argument contradicted itself? The answer turns on the difference between two ways a subject can be present, and the difference is the whole matter. A subject can be present as a witness, an authority on his own experience, speaking in his own voice to an audience that listens, which is the survivor giving testimony and the descendant telling his family's history. A subject can also be present as a performer of his own degradation, cast by an authority into the despised role and made to enact his subjugation for the instruction of others, which is the Black child on the auction block and the Jewish child in the star. The first presence honors the subject and teaches the audience. The second presence uses the subject as a prop and harms him. When this book calls for the subject's presence, it means the first, the witness with standing and voice, and when it condemns these exercises, it condemns the second, the victim conscripted into a reenactment of his own suffering. The distinction is not subtle once stated. One puts the microphone in the subject's hand; the other puts the chains on his wrists. The whole reform the book argues for can be compressed into the choice

between those two, between asking the subject to testify and forcing him to perform, and every exercise in this chapter chose to force the performance, which is why each one blew up and why the alternative, the testimony the museums recommend, does not.

The child cast as victim is the most visibly harmed, and he is not the only one the exercise injures. Around him the other children are assigned the complementary roles, and those roles do their own quiet damage. White children told to bid on their Black classmates are made to rehearse the posture of the buyer, to raise a hand and name a price for a human being, to practice, in play, the act of treating a person as property. The children cast as Nazis are flattered into the role of the favored and licensed to look down on the starred, made to feel, in their bodies, the pleasures of being the in-group in a system of contempt. Whatever the teacher hopes they will conclude, the immediate thing these children do is perform domination, and the book has argued throughout that performing domination teaches it more reliably than it teaches its wrongness. The bystander children, neither bought nor buying, learn a third lesson, that an authority they trust will arrange this scene and call it school, that the degradation of a classmate is a thing adults stage and require, which is a corrosive thing for any child to absorb. A lesson that means to build solidarity across a classroom instead assigns its children to buyer, bought, and bystander, and rehearses each in his part, so that the exercise distributes the injustice across the room rather than uniting the room against it, handing every child a role in the structure the lesson was meant to teach them to refuse.

The five-part anatomy this book set out in its third chapter runs through every one of these rooms, and it is the fifth part that breaks most instructively. There is the caste fixed by birth, now a real race or a real religion. An unrefusable authority runs it, the teacher. The cruelty is licensed by the form, the block and the gas chamber. The class looks on. And there is the reversal, the moment when the badge comes off and the lesson announces that the hierarchy was arranged, that the inferiority was never real, the moment the gentler exercises rely on to convert the experience into instruction. In these rooms the reversal cannot do its work, because the child sold was a Black child and the child gassed a Jewish one, and no announcement at the end can hand them back an identity they never took off. The comfortable participant in a privilege walk steps out of his assigned disadvantage when the workshop ends. A Black child sold at auction goes home Black, to a country where the auction he reenacted is his family’s actual history, and the Jewish boy made to play Hitler goes home Jewish, carrying whatever a child carries after that. The reversal that is supposed to prove the lesson safe proves instead that the lesson was never safe for the one person in the room it was nominally honoring, and the proof is the blowup.

A defender will say, correctly, that these teachers were not bigots, that they meant to convey the horror of slavery and the Holocaust, that the outcomes were the opposite of what they intended. All of that is true and none of it helps, because the argument of this

book has never been that the practitioners are wicked. The argument is that the method is wrong independent of the practitioner, that a machine for manufacturing degradation produces degradation whoever feeds it and however pure his motives, and these cases are the proof at full strength. A well-meaning teacher who reenacts an auction with a real Black child has produced a real act of racial humiliation, and her good intentions are present in the room and change nothing about what the child experienced. The persistence of intent alongside catastrophe is exactly what the book predicts, because intent was never the variable that mattered. The structure was the variable, and the structure delivers its result regardless of the heart of the person running it, which is why the same scene, with the same good intentions, keeps producing the same wreckage in Norfolk and Bronxville and Watertown and Washington.

None of this is an argument against teaching slavery or the Holocaust, and the distinction could not be more important. The history must be taught, taught early and taught hard, and the failure of the dramatic method is an argument for teaching it better rather than for teaching it less. A truer curriculum, not a softer one, is the alternative, and the institutions closest to these histories say so directly. The Holocaust museum asks teachers to bring survivors' testimony and primary documents into the classroom, the record of what was done and the voices of those it was done to, which keeps the subject present as a witness rather than absent behind a proxy. Educators who work on the teaching of slavery say the same, that reenactment is inappropriate at any grade and that the path runs through documents, narratives, and the words of the enslaved themselves. The right response to a child traumatized by a mock auction is to teach the auction block better, through the testimony of the people who stood on it, rather than to stop teaching it, so that the subject speaks and the child learns without anyone being sold. Anyone who reads these cases as a reason to retreat from the hard history has misread them. They are a reason to teach the hard history with the rigor and the restraint the dramatic shortcut refuses.

It is fair to ask what the better teaching looks like in practice, since a critique that only forbids is half a critique. The materials exist in abundance, and they begin with the words of the people who lived the history. For slavery, the record holds thousands of first-person narratives, the accounts dictated by formerly enslaved people in the nineteen-thirties, the published memoirs of those who escaped, the advertisements and bills of sale and ledgers that priced human beings in the language of commerce, all of which a teacher can read with a class and unpack without asking any child to stand on a block. A bill of sale read aloud, with its cold arithmetic of bodies and dollars, teaches the obscenity of the auction more truthfully than a staged auction does, because it is the actual document and it implicates no child in the room. For the Holocaust, the survivor testimony the museum recommends runs to many thousands of recorded hours, the filmed accounts of the people who stood at the ramp and survived the camps, available

to any classroom that wants them, and a survivor's recorded voice describing the selection carries a weight no eight-year-old miming death can approach. The diaries, the photographs, the transport lists, the recovered letters, the museum itself with its mountain of victims' shoes, all teach the event through its evidence and its witnesses rather than through a reenactment that puts a child in the victim's place. The pattern in every case is the same. Keep the subject present as the authority on his own suffering, through his narrative, his testimony, his recorded voice, and keep the documents present as the cold record of what was done, and let the student do the difficult work of reading, listening, and understanding, which is harder than performing and teaches incomparably more. None of this is a softening of the history. A class that reads the narratives of the enslaved and hears the testimony of survivors confronts the full horror, more of it and more accurately than any simulation delivers, and confronts it without manufacturing a new victim in the third row. The better teaching is more demanding than the dramatic shortcut, not less, and that added demand is the cost of teaching the truth without reenacting the crime.

The auction is the eye-color exercise with its mask removed. Everything the gentler rituals smuggle past the audience, the licensed cruelty, the manufactured hierarchy, the real harm dressed as a lesson, stands exposed the moment the subject is real, because a real subject cannot be made to absorb the degradation quietly the way a proxy can. The detonation that follows every one of these cases, the tears, the complaints, the investigations, the firings, is no malfunction at all, only the method working exactly as it always works, with the one difference that the harm has landed on someone whose pain the institution is finally forced to see. The blue-eyed child, the chip-holding Triangle, the comfortable walker, the touring student, all absorbed a smaller version of the same harm and were told it was education, and the only reason their rooms did not blow up is that the person harmed was a stand-in or a volunteer, present by choice and released on schedule. Put the real subject in the despised role and the truth comes out, that the genre has always trafficked in degradation and depended, for its respectability, on aiming that degradation at someone who could be told it did not count. Having followed the family from the gentlest of its rituals to the most catastrophic, the book turns now to the laws that govern all of them, the principles that explain why a method built to teach against cruelty should reproduce it at every scale, beginning with the oldest and deepest of those laws, that the method is the oppressor's own.

PART THREE

The Law

CHAPTER SEVEN

The Pedagogy of the Oppressor

This chapter's argument fits in four words, pedagogy of the oppressor, and this is where those four words are paid for. Everything before now has been preparation. The two classrooms established the specimen, the anatomy chapter dissected it into its five working parts, and the three claims were named and set aside to be proven one at a time. The first of those claims is the hardest of the three, and it is time to earn it in full. That claim holds that the method of these exercises is the method of the oppressor, and that the goodness of the cause does not change what the method is or what it does to the people who pass through it. To make that claim stand, I have to bring forward the man whose title I have turned inside out, because the case against coercive moral pedagogy was made, in its essentials, more than fifty years ago, by a writer who would have been appalled to see his name attached to the exercises that now borrow his vocabulary.

Paulo Freire was a Brazilian educator, born in 1921 in Recife, in the poor northeast of the country, and marked early by hunger after his own family slid into poverty during the Depression. He spent his working life teaching the poor to read, first in Brazil and then, after a military coup drove him into exile in 1964, across Latin America and beyond. His literacy work was practical before it was theoretical. He sat with peasants and laborers, the people at the bottom of one of the most unequal societies on earth, and taught them the written word, and out of that work came a conviction that reading and freedom were the same project, that a person learning to name the world was a person learning he could change it. In 1968, living in exile, he wrote the book that carried these convictions outward, *Pedagogy of the Oppressed*, first published in Spanish that year and in English two years later. It became one of the most widely read books on education ever written, and its central ideas have shaped how a great many teachers understand their work, including, by a bitter irony this chapter will trace, the teachers who run the exercises this book opposes.

Freire's most famous idea is his critique of what he called the banking model of education. In the banking model, the teacher is the one who knows and the student is the one who does not, and education consists of the teacher depositing knowledge into the

student the way a person deposits money into a bank, filling an empty vessel with the correct contents. The student's task is to receive, store, and repeat. He is treated as an object to be acted upon, a container to be filled, never as a person with his own mind and his own power to question. Freire's objection to this model was moral before it was practical. The banking model, he argued, does more than teach badly. It trains the student in passivity, in the acceptance of a world handed to him by an authority he cannot question, and that training serves domination, because a person schooled to receive without questioning is a person prepared to accept whatever the powerful choose to deposit in him. Against the banking model Freire set what he called problem-posing or dialogical education, in which teacher and student investigate the world together, in which the student is a co-creator of knowledge rather than a receptacle for it, and in which the act of learning is itself an act of becoming free. The point of education, for Freire, was the awakening of critical consciousness, conscientização in his Portuguese, the capacity to see the structures that shape a life and to act on them. Education was meant to produce people who think for themselves. Anything that produced obedient receivers was, to him, education in the service of oppression, however benign its intentions.

Freire had a name for the kind of mind the banking model produces, and set against the kind these exercises produce, the two turn out to be the same. The banking model, he held, yields what he called naive or submerged consciousness, a mind that takes the given world as natural and fixed, that cannot stand outside the arrangement it lives in to see that the arrangement was made by people and can be unmade by them. Critical consciousness, the thing real education exists to grow, is the opposite, the capacity to see the structure, name it, and recognize that it is no one's destiny. Now ask what the exercise grows. It hands the child a structure, a caste system invented that morning, and trains him to inhabit it, to feel its weight, to operate its cruelties, and to accept his place in it on the authority of the adult who built it. At no point does it equip him to step outside the structure and see it for the human contrivance it is. The reversal might look like that stepping-outside, the moment the child learns the hierarchy can be flipped, except that the flip is performed by the same authority and teaches only that the structure is the authority's to rearrange, which leaves the child more submerged than before, more convinced that the shape of his world is something handed down from above. The exercise produces the exact consciousness Freire spent his life trying to dissolve, a mind fluent in its cage and grateful to the keeper. It calls that fluency understanding and sends the child home with it.

That much of Freire is widely known. The part that matters most for this book is quoted less often, and it cuts directly against the exercises that fly his flag. Freire understood, with a clarity few have matched, that the great danger facing the oppressed is that they will free themselves by becoming oppressors in turn. He saw that the person

crushed by domination hates it while also, in a part of himself he cannot easily reach, admiring it and longing to wield it. The oppressed, Freire wrote, carry the image of the oppressor inside them, having absorbed his values and his guidelines, so that they are at once themselves and the oppressor whose consciousness they have taken on. This is why, he observed, the oppressed so often turn their rage sideways, onto each other, in what he called horizontal violence, striking at their own comrades for small offenses because the real source of their suffering sits beyond their reach. And it is why liberation is so hard and so easily betrayed, because the oppressed, handed a sliver of power, will tend to use it as they saw it used on them. The peasant who gains a little authority, Freire observed, tends to wield it as hard as the owner did, or harder.

From this insight came the warning that is the spine of my argument. Liberation, Freire insisted, cannot be achieved by reversing the terms of oppression, by the oppressed simply trading places with the oppressors and running the same machine with the roles swapped. A revolution that only changes who holds the whip leaves oppression in place; it has merely moved the whip to a new hand. The goal, in the phrase that became the heart of the book, is a person who is neither oppressor nor oppressed, a human being freed from the whole relation, on either end of it. To get there, the method itself has to be free of domination. You cannot deposit freedom into people by force, because the depositing is the thing that has to be unlearned. You cannot command people into critical consciousness, because the commanding reproduces the same passivity you are trying to break. The method has to embody the freedom it intends, or it will teach, underneath whatever it says, the lesson that the powerful set the terms and the rest comply. Freire built his whole pedagogy around this principle, that the means and the end cannot be split apart, that a pedagogy of liberation has to be liberating in its form and not only in its content.

There is one more piece of Freire I need before turning to the exercises, because it explains why the false solution is so tempting. Freire wrote, drawing on the psychologist Erich Fromm, who had described the same thing a generation earlier in *Escape from Freedom*, about what he called the fear of freedom. Freedom is a burden, he saw, because it demands that a person take responsibility for himself, eject the internalized oppressor, and stand on his own judgment, and many people, oppressed and comfortable alike, would rather not. They prefer the security of a world whose rules someone else has set. They would sooner be told what is right than work it out and answer for it. The oppressed fear freedom because it asks them to become authors of their own lives after a lifetime of being told they are nothing; the powerful fear it because it threatens the order that serves them. Either way, the fear of freedom is the quiet engine of submission, the reason people accept the deposits made in them and the roles assigned to them, and any honest education has to contend with it. The exercises this book examines never contend with the fear of freedom. They feed it.

Hold Freire's book in one hand and Jane Elliott's exercise in the other, and look at the year stamped on both. 1968. The book and the exercise were born in the same twelve months, on the northern and southern halves of the same hemisphere, addressed to the same problem, the problem of how human beings come to dominate one another and how that domination might end. They are mirror images, and the mirror reverses everything. Freire's pedagogy treats the learner as a subject and refuses to deposit anything in him by force. Elliott's exercise treats the learner as an object, a thing to be sorted and labeled and acted upon, and deposits in him, by the full force of a teacher's authority, an assigned identity and a set of feelings chosen for him in advance. Freire's pedagogy aims at a person who is neither oppressor nor oppressed. Elliott's exercise makes every child both, an oppressor on one day and a victim on the next, and calls the alternation a lesson. Freire warned that reversing the roles of master and servant changes nothing essential. Elliott built the reversal into the design and presented it as the design's genius. The exercise is, point for point, the thing Freire wrote his book to oppose. It is the banking model armed, the depositing of correct attitudes through manufactured suffering rather than mere repetition, and its proudest feature, the overnight swap of high and low, is the exact maneuver Freire identified as the false road that leads nowhere. One of these two things from 1968 is a pedagogy of the oppressed. The other, wearing the borrowed clothes of the first, is a pedagogy of the oppressor.

I have used the phrase a method carries its own content several times, and it is the load-bearing idea of this chapter, so I will set it down plainly and defend it. The claim is that how you teach is itself something you teach, that the form of an action delivers a lesson independent of, and often louder than, the lesson the action is supposed to deliver. A teacher who screams at children to be calm has taught them that the powerful person in the room screams. A parent who hits a child for hitting his sister has taught that hitting is what you do to those smaller than you when you are angry and in charge. The content of the message, be calm, do not hit, is contradicted by the form of its delivery, and children, who learn far more from what adults do than from what adults say, absorb the form. The lesson that sticks is the lesson the method embodies. Nothing about this is soft or sentimental; it ranks among the most reliable findings about how the young learn anything, that example outweighs instruction, that the medium teaches.

Apply this to the exercise and the result is damning. The stated content is a lesson against prejudice and cruelty. The form is an act of prejudice and cruelty, administered by an authority, justified by a trait of birth, and performed for an audience. By the principle that the method carries its own content, the lesson that lands is the lesson the form embodies, which is that an authority may sort people by a trait they cannot change, may license the favored to abuse the rest, and may stage the whole thing for an audience, so long as the authority says it is for a good reason. That is the lesson the child practices with his own hands and his own voice across two days, and practice, here as

everywhere, is how a capacity is built. Far from coming away inoculated against the structure of oppression, the child comes away fluent in it, having operated it with his own hands and his own voice, and having learned that it is a tool an adult in charge can pick up and assign. You cannot teach a child to reject a machine by making him run it. You can teach him only how it runs.

This is the inversion in its full form, and it holds regardless of the cause the exercise serves, because the principle is about method, not motive. A good motive attached to the oppressor's method does not convert the method into something else. It only disguises it, and the disguise is what makes coercive moral pedagogy so durable, because the good motive sits in front of the method like a screen, and everyone looking at the screen sees the kindness of the goal and misses the cruelty of the means. Freire saw past the screen. He insisted, against allies and enemies alike, that you cannot reach a humane end by inhumane means, that the means are the end in the process of becoming, and that a liberation pursued through domination arrives, if it arrives at all, as a new domination wearing liberation's name.

Freire drew one more distinction that lands squarely on these exercises, between false generosity and the real thing. False generosity, in his account, is the charity that eases the giver's conscience while leaving in place the conditions that made the charity necessary. The almsgiver who hands a coin to the beggar and feels cleansed, while the order that produced the beggar stands untouched, has performed false generosity, because the gesture is aimed at the giver's feeling rather than the sufferer's freedom. True generosity works to end the need for itself, to dismantle the arrangement that keeps some people dependent on the goodwill of others. Hold the exercise up to that distinction and it shows its nature. The exercise is false generosity rendered as pedagogy. It is a gesture that makes the people performing it, the teacher and, later, the consultant and the institution, feel that they have done something about injustice, while the structures that produce injustice go wholly unaddressed, untouched by an afternoon of armbands in a room those structures never entered. The child cries, the teacher is moved, the watchers are moved, and at the end of the day everyone has felt a great deal and changed nothing, which is the precise signature of false generosity, the swapping of feeling for the harder work that would cost the comfortable something. Worse than the almsgiver's coin, which at least feeds the beggar for an hour, the exercise gives the sufferer in the room nothing at all, because there is no sufferer in the room except the one the teacher has manufactured, and that one is sent home cured of nothing. The whole performance is generosity turned inward, a kindness the giver does to himself and calls a kindness to the world.

Honesty requires me to turn the same blade on Freire that I have turned on Elliott, because Freire sits no higher than the critique this book makes, and a chapter that used him as an unquestioned authority would be doing to him what these exercises do to their

subjects, treating a complicated human being as a convenient object. Freire has been criticized, from the right and from the left, for tensions in his own thought, and the criticisms bear on my argument. The sharpest charge is vanguardism, the worry that Freire's educator, for all his talk of dialogue, still arrives knowing what critical consciousness looks like and still aims to lead the student toward it, which is a form of the same authority Freire claims to reject. If the teacher already knows the destination and the student does not, then the dialogue has a thumb on the scale, and the problem-posing classroom can become the banking classroom in subtler dress, depositing the correct politics in place of the correct facts. This is a real tension in Freire, and serious readers have pressed it. I press it too, because it sets the standard I am holding the exercises to. The test of any pedagogy is whether it leaves the learner freer, more able to think for himself, less dependent on an authority to tell him what to feel. By that test Freire's own method is open to question at its edges. By that same test the exercises fail completely, and they fail at the center rather than the edges, because they do not even aspire to leave the child thinking for himself. They aim to make him feel a specific thing on command, and they count the command a success when he feels it.

There is a further irony, noticed by scholars of Freire, that sharpens the point. Freire wrote for the education of the oppressed, the people at the bottom, and his methods were meant to be practiced by them and with them, in the work of their own liberation. The exercises that now invoke his spirit are aimed, almost always, at the comfortable, at white schoolchildren in functioning schools and salaried employees in corporate training rooms, people who sit nearer the top of the real hierarchy than the bottom. A philosopher writing on this paradox called it, pointedly, a pedagogy of the privileged, and the phrase names something true. A method built to help the powerless find their voice has been repurposed to give the powerful a brief, managed taste of powerlessness, on the theory that the taste will make them kinder. Freire would have rejected the theory root and branch, because for him understanding grew from sustained reflection and action in the world, from praxis, and never from a staged ordeal that ends when the bell rings and everyone goes back to their lives unchanged except for the certificate they now carry. The exercises keep Freire's vocabulary of liberation and empty it of everything that made it liberating. They are how a method can quote Freire while doing the precise reverse of what he taught.

I come now to the part of the argument that reaches beyond the classroom, because the inversion is more than a pedagogical error. It is a political danger, and naming the danger is the real work of this chapter. The exercise has a stated lesson and a hidden one. Its stated lesson is about prejudice. The hidden lesson, the one carried by the form, is about authority, and it runs like this: that when a trusted authority tells you to accept a hierarchy, to occupy your assigned place in it, and to treat the people below you accordingly, you should comply, and your compliance is virtuous because the authority

says the purpose is good. That hidden lesson is the oldest and most dangerous lesson in the political world, because it is the lesson every authoritarian system needs its subjects to have learned. The exercise teaches submission to authority and dresses the submission as moral education. It trains the child in the thing that, more than any other, makes oppression possible, which is to obey a command against the protest of his own conscience because the person giving the command has assured him it is right.

Consider what the exercise asks of the child in the lower caste. It asks him to accept, on a teacher's say-so, that he is inferior today, to wear the mark of it, to absorb the mistreatment without resisting, and to trust that the adult in charge knows what she is doing. The child who complies, who bows his head and plays his assigned part and waits for the bell, is the child the exercise rewards. The child who refuses, who says this is wrong and will not carry water for it, is the child the exercise punishes, as I was punished in Lincoln for refusing the command. Read that outcome against the politics of the last century and it should chill you. The disposition the exercise rewards, obedience to authority in the face of one's own moral objection, is the disposition that staffs every machinery of cruelty ever built. The disposition it punishes, the refusal to comply with an unjust order whatever the cost, is the disposition that resists them. In the name of teaching children to oppose oppression, then, the exercise trains them in the precise habit oppression depends on and punishes the precise habit that defeats it.

This is no abstract worry, and the thinkers who studied how atrocity happens have left no room to treat it as one. Hannah Arendt, watching Adolf Eichmann on trial in Jerusalem in 1961, found in him no monster, only a functionary, a man who had organized the transport of human beings to their deaths and who defended himself by saying he had followed orders and done his duty. The horror of Eichmann, in Arendt's account, was his ordinariness, the absence in him of any independent moral judgment that might have stood against the orders he received. She gave the phenomenon a name that has never lost its force, the banality of evil, and the banality lay exactly in the surrender of conscience to authority, the substitution of obedience for thought. Arendt's portrait of Eichmann himself has not gone unchallenged, and a later chapter will report the historians who, working from his own words, judge him a convinced ideologue rather than a hollow clerk. That correction matters for the man and changes nothing in the mechanism. Arendt's lasting insight reached past the one defendant to a capacity that lives in ordinary people, the readiness to hand one's conscience to an authority and let obedience stand in for judgment, and that capacity is real whether or not Eichmann was its purest specimen. The reflex the exercise rehearses runs the same whether the man in the glass booth believed in the killing or merely processed it, because the danger Arendt named is the surrender of the inner check, and a great many people have made that surrender under far less pressure than he faced. The lesson Arendt drew, and that the century screams, is that the great crimes are committed less by sadists than by ordinary

people who have been trained, or have trained themselves, not to question what they are told to do. A society that wishes to prevent such crimes has to cultivate the opposite capacity, the habit of independent judgment, the willingness to refuse. Coercive moral pedagogy cultivates the wrong one. It rehearses obedience to authority under the banner of teaching resistance to it.

The historian Timothy Snyder, writing for our own anxious moment about how free societies fall, opens his guide to resisting tyranny with a single instruction that doubles as the deepest criticism of these exercises: do not obey in advance. Snyder draws on the example of Austria in 1938, where ordinary people, before they were compelled to do anything, anticipated what the new Nazi authorities wanted and began doing it on their own, humiliating their Jewish neighbors without being ordered to, because they had read the direction of power and adapted to it on instinct. Anticipatory obedience, Snyder argues, is how authoritarianism installs itself, because the volunteered compliance of the population teaches those in power what is possible and hands them their cruelty for free. Now set the exercise beside that warning. The exercise is a school for obedience in advance. It takes a room of children and trains them, in the space of an afternoon, to read the authority's cues, accept the hierarchy on offer, and perform the roles assigned, all before any real pressure has been applied, all on the strength of a teacher's word that this is good for them. It teaches the young the exact reflex Snyder spent a book trying to break, the reflex of falling into line with power because power has told a story about why they should.

The same finding came, in a different form, from the laboratory. In 1961, the year of the Eichmann trial, Stanley Milgram brought ordinary volunteers to deliver what they believed were escalating electric shocks to a stranger, past screams and pleas, so long as a calm authority in a lab coat told them the experiment required it. A later chapter takes apart how those demonstrations were staged and how their famous numbers were chosen, so I lean on none of that here. The narrow point survives the demolition: when an authority a person trusts directs him to harm another and assures him the harm serves a purpose, a great many people will obey. That is the reflex the exercise rehearses, and the one a moral education worth the name should be dismantling.

I should guard the argument against a misreading here, because the names I have just invoked could be turned into a weapon I do not intend. Nothing in this chapter says that children should be kept from learning about Eichmann, or Milgram, or the machinery of obedience. The opposite is true. A child should learn, in detail and at the right age, how ordinary people were led to commit atrocities, how anticipatory obedience works, how an authority can talk a decent person into cruelty, because that knowledge is among the few real defenses a person has against doing it himself. The distinction this book keeps drawing holds here with full force. Teaching a child about obedience is the work of a free mind examining the trap from the outside. Running an obedience exercise on a child

is the work of springing the trap on him and calling the spring a lesson. You can read Milgram's findings to a class and discuss what they mean for how each student should act when an authority pressures him, and you will have strengthened the same capacity for refusal that Milgram found so rare. You can, instead, do to the class what Milgram did to his volunteers, place them under an authority and watch them comply, and you will have demonstrated the finding by adding the children to the data, training the obedience rather than warning against it. One approach builds the defense. The other rehearses the vulnerability. The exercises in this book are firmly of the second kind, which is why invoking Arendt and Snyder against them is no contradiction. Those thinkers wrote to wake the conscience. The exercises put it to sleep and call the sleep instruction.

Put Arendt, Snyder, and Milgram together with Freire and the shape of the danger is complete. A method that trains people to accept coercion for a reason they are told is good has trained them to accept coercion, and the training outlasts the particular reason. The form is what endures. A child who learns that it is right to obey an authority who sorts people by birth and licenses cruelty, because this authority promised the purpose was kind, has learned a lesson that reaches far past the topic of racism. He has learned a general permission, the permission to suspend his own conscience when an authority supplies a sufficiently noble reason, and that permission is exactly what the next authority will ask of him, the one whose reasons are not kind at all. A population taught to accept coercion for a good reason is a population prepared to accept it for a bad one, because what it has been taught is to accept coercion, and the reason is a detail the next regime will supply for itself. The exercises rehearse the submission and trust the cause to justify it. Causes change hands, though, and the submission, once learned, serves whoever holds the authority next.

The Czech dissident Vaclav Havel, writing in 1978 from inside a regime that ruled less by terror than by ritual, gave this mechanism its most exact description, and it completes the case. Havel asked how a tired, cynical dictatorship that almost no one believed in could nonetheless endure, and his answer was that it endured because people performed it. His famous example was a greengrocer who places in his shop window, among the onions and carrots, a sign bearing an official slogan he does not believe. The grocer faces no direct threat over it. He posts the sign because everyone posts the sign, because posting it is how a person signals that he will go along, and his small daily act of going along, multiplied across millions of grocers and clerks and citizens, is the thing that holds the whole edifice up. Havel called this living within the lie, and he saw that the regime's true power lay less in its police than in the willingness of ordinary people to enact, in public, a thing they did not hold to be true, because enacting it was easier and safer than refusing. The opposite, the act that threatened the regime more than any pamphlet, was simply to stop, to live within the truth, to take the sign out of the window

and bear what followed.

Now look at what the exercise trains. It trains the child to post the sign. It places him in a ritual whose premise he need not believe, the premise that eye color marks the worth of a person, and it rewards him for enacting the ritual anyway, for wearing the mark and performing the role and treating his classmates as the script demands, because the authority has asked him to and going along is easier than refusing. The child who plays his part well, who throws himself into the assigned cruelty or absorbs the assigned humiliation without protest, has learned exactly the disposition Havel identified as the lifeblood of soft tyranny, the readiness to live within a lie when an authority sets the lie in front of him and the cost of refusal is real. And the child who will not post the sign, who says this is wrong and takes the mark off his sleeve, has done the one thing Havel said could bring such a system down. The exercise, which imagines itself teaching children to resist, has built a small machine for teaching them to comply, and it has punished, in advance, the only act that resistance is ever made of.

A fair reader will want to push back here, and the push-back deserves an answer rather than a dodge. Am I not making too much of a classroom exercise, treating a third-grade afternoon as though it produced storm troopers, reaching for Eichmann and Havel to crush a lesson plan. The objection has force, and meeting it honestly sharpens the claim rather than weakening it. I am not arguing that the eye-color exercise manufactures fascists, or that a child who wore an armband in 1973 is one refusal away from a uniform. The effect of a single exercise on a single child is small, and I will not pretend otherwise. The argument operates at the level of disposition and direction, not of immediate cause. What the exercise does is rehearse a reflex and attach a moral reward to it, and reflexes rehearsed in childhood under the approval of trusted adults are exactly the kind that settle in and last. A child is not made authoritarian by one afternoon. Children are shaped across thousands of small lessons about when to obey and when to refuse, and this exercise is one of those lessons, pointed in the wrong direction and stamped with the school’s approval, teaching that going along with an authority’s cruelty is the mature and caring thing to do.

The scale argument from the anatomy chapter applies with full weight here. The machine is indifferent to size, and so is the disposition it trains. A reflex rehearsed in the small room is the same one the large rooms require, which is why it matters where the training points even when any single dose is mild. A culture does not arrive at the acceptance of coercion in one step. It arrives through a long accumulation of moments in which coercion was presented as reasonable, as necessary, as kind, and in which refusal was treated as the problem. Each such moment is small. The accumulation is not. This exercise earns that much attention precisely because it is small and respectable and repeated, taught to millions of children and adults as a model of moral seriousness, which makes it a steady contributor to the accumulation rather than a dramatic one. The

danger of a thing like this lies in its modesty, in how reasonable it looks, in how easily a decent person nods along. That is how the wrong reflex gets trained at scale, less by monsters demanding it than by good people recommending it, one harmless-looking afternoon at a time.

Here the inversion rejoins the personal story this book began with. The single act in either of my two classrooms that embodied resistance to oppression was the act the exercise was built to crush. When I refused, on the second day in Lincoln, to put on the armband and play my part, when I said the words that have become this book's refrain, that this was not my thing and I would not do it, I was doing the one act in either room that Arendt and Snyder and Freire would all have recognized as the real lesson, the only lesson, in the whole rigged afternoon. The refusal declined an unjust order from an authority I was supposed to trust. It rejected an assigned place in a hierarchy I had never consented to. In the smallest way available to an eight-year-old, it did what the people who resist tyranny do, which is to say no to power when power is wrong and to bear the cost of the no. And the exercise punished me for it, because the exercise, whatever it believed about itself, stood with the compliant and against the refusers. A lesson that punishes the child who refuses an unjust command and rewards the children who carry it out has declared, in the clearest terms available, which side of the human story it is teaching. Far from teaching children to resist the next Eichmann, it rehearses in them the disposition of the ordinary people who made him possible, and it calls that education in justice.

Those four words are now paid for. The method of these exercises is the method of the oppressor, in any hands, for any cause, because the method is built from caste, coercion, and spectacle, and those are the oppressor's tools no matter whose grip is on them. Freire saw it in 1968 and wrote the book that names the alternative. The exercises took his vocabulary, inverted his method, and have spent the decades since teaching, under his borrowed banner, the submission he gave his life to ending. A method carries its own content. Teach a child obedience to cruelty in a good cause, and you have taught him obedience to cruelty, and the good cause will not be standing there to protect him on the day a worse one comes asking for the same obedience he was trained to give. There is a test buried in everything this chapter has argued, and it is the one the whole book turns on. The measure of a moral education is whether it strengthens the conscience or sedates it, whether it leaves a child more able to refuse an unjust order or more practiced at obeying one. By that single measure the exercises fail, and they fail in the direction that matters, training the obedience and punishing the refusal, building the disposition every tyranny has needed and starving the one disposition that has ever stopped a tyranny. Freire, Arendt, Snyder, Havel, and a boy with an armband he would not wear all point at the same conclusion from their different vantages. The free conscience, the willingness to say no to power when power is wrong, is the thing worth teaching, and it

is the precise thing these exercises are built to break.

That is the first claim, earned in full. The second goes deeper into the damage, into the question of what these exercises deposit in the people who endure them, and for that I turn to a doctor’s word for a sickness caused by the cure.

CHAPTER EIGHT

Iatrogenic Empathy

The first claim established that the method of these exercises belongs to the oppressor. The second claim, the one this chapter exists to prove, says something about results. It says that the exercises do more than use a tainted method; they produce the opposite of the good they promise. The cure manufactures the disease. A program built to grow empathy grows its absence, a program built to reduce cruelty rehearses cruelty, a program built to open minds shuts them, and the failure is reliable enough to count as a property of the design rather than an accident of a bad day. To name this kind of failure I am borrowing a word from medicine, because medicine has studied it longer and more honestly than education ever has, and because the man who turned the word into an indictment of modern institutions left a framework that fits the empathy industry as if it had been built for it.

The word is iatrogenic, and the man is Ivan Illich. Illich was an Austrian-born thinker, a former Catholic priest who spent the second half of his life as one of the sharpest critics of the institutions of industrial society, and in 1975 he published the book that made the term famous, *Medical Nemesis*, later reissued as *Limits to Medicine*. The book opens with a sentence that cost him a great many friends in the medical profession: the medical establishment has become a major threat to health. His argument was that modern medicine, past a certain point of growth, had begun to produce sickness rather than cure it, and that the sickness it produced was of a particular and self-concealing kind. Iatrogenesis, from the Greek iatros for physician and genesis for origin, names illness caused by the treatment meant to heal. Illich did not invent the word; doctors had long used it for the narrow case of a patient harmed by a procedure. What he did was widen it into a social theory, and the widening is what matters here.

Illich built the charge on the profession’s own data, the infections, the adverse drug reactions, the unnecessary surgeries, and many of his warnings about overtreatment and the medicalizing of ordinary life have since moved from the fringe to the mainstream of medicine’s self-criticism. What gave the book its lasting force was the idea underneath the numbers. Illich argued that health, properly understood, is a capacity a person has

rather than a commodity a profession dispenses, the capacity to meet pain, sickness, and death on one's own terms. Modern medicine, by claiming that capacity for itself, had committed what he called the expropriation of health, taking from people the thing it promised to provide, so that the patient who learned to see every ache as a problem for the expert lost the competence to care for himself and became a permanent consumer of care. That move, the expropriation of a human capacity by the institution built to serve it, is the engine of everything Illich wrote, and it is the move these exercises make with the capacity for moral feeling. Keep the word expropriation close; it returns when the damage these exercises do comes into full view.

Illich saw iatrogenesis working at three levels, and the three map onto the empathy exercises so cleanly that I will introduce them in their medical form first and turn them on the exercises afterward. The first level is clinical iatrogenesis, the direct injury a treatment does to the body it was meant to help, the infection caught in the hospital, the drug whose side effects are worse than the complaint, the surgery that leaves the patient sicker than it found him. The second level is social iatrogenesis, subtler and larger, the way a medical system, by claiming ever more of life as its territory, trains people out of the capacity to care for themselves, turns ordinary conditions of living into diagnoses requiring professional management, and makes a population dependent on the institution that is harming it. Cultural iatrogenesis, the deepest of the three, sets in when the medical enterprise erodes a culture's inherited capacity to face suffering, pain, and death with meaning, replacing the hard arts of endurance and acceptance with the expectation that every kind of suffering is a technical problem for an expert to solve. At the first level medicine injures the body. The second level disables the person. The third hollows out the culture's ability to bear the unavoidable. Each level feeds the next, and together they form what Illich called a self-reinforcing loop, a system that generates the demand for more of itself by producing the helplessness it claims to relieve.

Illich gave the whole argument a political edge in the book's final section, on what he called the politics of health, where he held that the medical system served power by absorbing discontent, converting what might have been anger at the conditions of industrial life into demands for more therapy, so the sick blamed their own bodies rather than the arrangements that had sickened them. That move is the one to hold onto, because the empathy industry performs the same service for the institutions that buy it, absorbing the discontent that injustice produces, converting it into a managed ritual, and sending everyone home feeling that something has been done while the arrangements that produced the injustice carry on undisturbed. A later chapter develops that political function in full. Here it is enough to see that Illich's medicine and these exercises share more than the shape of the harm; they share the use to which the harm is put.

Behind the three levels sat a more general idea, the one that makes Illich useful far beyond medicine. He called it counterproductivity, and the claim was that any

institution, once its growth passes a certain threshold, begins to undermine the purpose it was created to serve. Transport past that threshold produces congestion rather than mobility, the cars so numerous that no one can move. Schooling past that threshold produces credential-chasing and the conviction that knowledge is something only institutions can dispense, a thing Illich had attacked four years earlier in *Deschooling Society*, where he argued that schools, by monopolizing education, had taught people they could not educate themselves. Medicine past that threshold produces iatrogenesis. The pattern holds across every case: a good thing scaled past its proper limit flips into its opposite, and the institution that delivers it acquires what Illich called a radical monopoly, a grip so complete that it crowds out every other way of meeting the need and then defines the need in its own terms, so that people forget there was ever another way. The doctor becomes the only legitimate source of health, the school the only legitimate source of knowledge, and the profession, in Illich's harsh word, becomes disabling, leaving the people it serves less able to do for themselves than they were before it arrived. Hold that whole apparatus in mind, the three levels and the counterproductivity and the disabling monopoly, and turn it now on the exercises this book examines, because every part of it has a match there.

Coercive moral pedagogy is a helping institution. It presents itself as treatment for a sickness, the sickness of prejudice, and it promises a cure, the growth of empathy and the reduction of cruelty. By Illich's logic the right question is whether this treatment, like the medical one, produces the disease it claims to fight, and the answer, level by level, is that it does. Call the result iatrogenic empathy, empathy-injury caused by the empathy treatment, and trace it through Illich's three levels.

At the clinical level, the direct injury, the exercise does exactly what the earlier chapters showed. It deposits the three harms the earlier chapters named straight into the person it was meant to improve. This is the hospital infection of moral education, the treatment that leaves the patient carrying something he did not walk in with.

The injury repays a closer look, because the medical analogy can make it sound abstract when it is concrete. A child who spent an afternoon licensed to mock his classmates has, in plain fact, practiced mockery, has felt the specific pleasure of looking down on someone with permission, and has stored that experience where experiences are stored, in the body of habit. A child who spent the afternoon on the bottom has, in plain fact, been mocked by people he trusted, under the eye of an adult who allowed it, and has learned something about how fast the people around him will turn. Neither of these is a lesson about distant injustice. Each is an event that happened to the child himself, in his own classroom, with his own name on it, and the residue is his to carry. When a treatment's first and most certain effect is to inflict on the patient a version of the harm it was prescribed against, the burden of proof shifts hard onto anyone who would keep prescribing it, and that burden, as the rest of this book shows, has never

been met.

At the social level, the exercise de-skills the capacity it claims to build. Real moral understanding grows from the slow, unmanaged work of encountering other people, listening to them, reading their lives, and revising oneself over time, the ordinary human labor by which a person becomes less cruel. The exercise replaces that labor with a managed product, a packaged ordeal administered by a facilitator on a schedule, and in doing so it teaches the child, and later the adult, that empathy is a thing you receive from an expert in a session rather than a thing you build yourself across a life. It creates dependency on the ritual and the professional who runs it, exactly as Illich's medicine created dependency on the doctor, and it prepares the ground for an entire industry of facilitators and consultants whose existence depends on the belief that ordinary people cannot grow decent on their own. That industry is the subject of a later chapter. Here the point is structural: the exercise, by professionalizing empathy, disables the unmanaged human capacity for it.

The de-skilling runs deeper than dependence on a facilitator. A person taught that empathy is delivered in sessions comes to believe that the moral life has checkpoints, that one can attend the workshop, complete the module, receive the certificate, and be, for a time, finished. Moral attention, which in a healthy life is a thing without end, a daily practice of noticing other people and adjusting oneself, gets repackaged as a discrete event with a beginning and a conclusion. The institution benefits twice over from this, first by selling the event and then by selling the next one when the effect of the last has worn off, exactly as Illich's medicine sold the appointment and then the follow-up. What the person loses in the exchange is the sense that the work is his own and never done, the sense that no expert can do it for him and no certificate can close it. He has traded a capacity for a service, and the service, like all the disabling services Illich described, leaves him needing it more rather than less.

Illich's idea of the radical monopoly sharpens this further, and it explains a feature of the present moment that puzzles many observers. A radical monopoly, in his sense, does more than dominate a market; it crowds out every rival way of meeting a need and then redefines the need so that only its own product counts as meeting it. Medicine became the radical monopoly over health, until coping with one's own body without a professional came to seem reckless. The dramatic exercise, and the industry grown up around it, has reached for the same kind of monopoly over the work of becoming less prejudiced. It has helped install a definition of that work under which the real activity is the workshop, the training, the staged confrontation, the certified program, and under which the older and quieter forms, the friendship across a line, the patient argument, the sustained private effort, no longer register as doing the work at all. A person who has built real relationships across difference but skipped the training can now be made to feel he has not done his part, while a person who has done nothing but attend the session

holds the certificate that says he has. The monopoly inverts the accounting, crediting the ritual and discrediting the substance, and once it holds, it becomes self-protecting, because to question the ritual is made to look like refusing the goal. That confusion, the ritual wrapping itself in the goal so an attack on the method reads as an attack on justice, is the monopoly defending its territory, and it is the reason a method this counterproductive can present itself as the one serious option and brand its critics as the unserious.

At the cultural level, the deepest, the exercise does to moral life what Illich said medicine did to suffering. It erodes the inherited frameworks through which people become better, the encounter, the friendship across a line, the patient study of history, the religious and ethical traditions that train conscience over years, and substitutes a manufactured shortcut that feels like moral seriousness while delivering none of it. Illich said cultural iatrogenesis sets in when medicine saps a people's will to bear their own reality. Cultural iatrogenic empathy sets in when the dramatic exercise saps a people's capacity for the real moral confrontation that would change them, by handing them a counterfeit confrontation that closes the question. The child who has done the exercise believes he has faced the hardest thing and come through. He has faced a staged thing and learned to mistake the staging for the substance, and that mistake, repeated across a culture, leaves a population convinced it has done its moral work when it has only performed it.

The cultural level is where Illich's word expropriation comes due, because the deepest thing these exercises take is the one hardest to get back. A person's conscience is his own instrument for telling right from wrong and for knowing when he has fallen short, and like Illich's health it is a capacity rather than a possession, something exercised across a life and never delivered whole by anyone else. The exercise expropriates it. By staging a moral ordeal and certifying its completion, the exercise tells the person that his conscience has been seen to, that the hard internal work of becoming less cruel has been done on his behalf in a managed session, and it hands him, in place of the ongoing labor, the feeling of having arrived. That feeling is the false certificate, and it is the most damaging deposit of all, worse than the practiced cruelty and the resentment, because those at least announce themselves as injuries while the certificate disguises itself as a cure. A person who knows he is still cruel may yet work on it. A person issued a certificate of empathy believes the account is settled and stops working, and the stopping is the injury. Expropriation of health left people unable to care for their bodies and dependent on the system that had taken the job. Expropriation of conscience leaves people convinced they have done their moral work when they have handed it to a ritual, and a society full of such people is one that has performed its decency rather than practiced it, complete with paperwork, and grown no kinder for the show.

Naming the failure is one thing, and a careful reader is owed the mechanism, the answer to why the dramatic intervention backfires so reliably. Two features of the design account for most of it, and the first is coercion. The exercise does not invite a child to consider a question. It compels him to undergo an experience, assigns him feelings, and rewards him for producing them, and compulsion in matters of belief has a well-documented habit of producing the opposite of what it demands. The phenomenon has a name in psychology, reactance, described by Jack Brehm in the nineteen-sixties: when people feel their freedom to think and choose being squeezed, they push back, reasserting the threatened freedom by moving against the position being forced on them. Tell a person he must believe something, leave him no room to arrive at it himself, and a part of him digs in against it precisely because his autonomy is under threat.

The clearest test of this in the domain that matters here was run by the psychologist Lisa Legault and her colleagues, in a study published in 2011 with a title that states the finding plainly, the ironic effects of antiprejudice messages. They gave people one of two kinds of material aimed at reducing prejudice. One kind was controlling, the message that society demands you not be prejudiced, that you must comply, with the consequences of failure stressed. The other was autonomy-supportive, the message that being free of prejudice is personally valuable and worth choosing for one's own reasons. A third group received nothing. The result should be engraved over the door of every diversity office in the country. The people who got the controlling message came out more prejudiced than the people who got nothing at all, on both explicit and implicit measures, while the people who got the autonomy-supportive message came out less prejudiced than the control group. Pressure did more than fail. It reversed the outcome, producing more of the prejudice it set out to reduce, while the same goal pursued through invitation succeeded. Coercive moral pedagogy sits at the far extreme of the controlling end of that scale, applying a controlling ordeal rather than a controlling brochure, and Legault's finding predicts exactly what the exercises produce, the reinforcement of the attitude they meant to dissolve. Legault enters here as illustration rather than proof. A single study, however clean its result, carries less than its vividness suggests, and the caution this book brings to the contested findings it cites elsewhere belongs on this one too. The weight on this point rests on the older and wider record of reactance, observed across many settings, and on the field evidence gathered later in the book, so the argument would hold if this lone study were withdrawn tomorrow.

The reason coercion fails so reliably in this domain has to do with what a belief is. A fact can be drilled into a person against his will; he can be made to recite that a war ended in a certain year whether he cares or not. A value cannot be installed the same way, because a value held only because it was forced is held not at all, only complied with, and compliance evaporates the moment the force is removed. To believe that other people deserve respect is to have made that judgment one's own, and a judgment made

under compulsion belongs to the compeller, not the person, who experiences it as an imposition to resist rather than a conviction to keep. This is why moral matters are exactly the ones where pressure backfires worst. The harder an authority pushes a value, the more the value comes to feel like the authority's property rather than the person's own, and the more a free mind reasserts itself by holding the value at arm's length. The same logic explains the long record of failure for interventions built on fear and threat, the campaigns that tried to frighten people into virtue and found them more attached to the vice. A later chapter takes up the clearest case, the prison-visit programs that set out to scare children straight and produced more crime, but the principle is general and it is the principle at work here. You can frighten a person, you can shame him, you can compel him, and you will move his behavior in the moment while pushing his convictions in the wrong direction over time, because the part of him that decides whether he becomes a decent adult is the part that decides for itself, and that part hardens against everything done to it by force.

The second feature is shame, and to see why it matters one has to draw a distinction the exercises ignore and that decades of research have made unmistakable, the distinction between shame and guilt. The two feel similar from the inside and do opposite work. Guilt is about a behavior, the sense that I did a bad thing, and it points toward repair, toward apology and changed conduct, because the self remains intact and capable of doing better. Shame is about the self, the sense that I am a bad thing, a global verdict on the whole person, and it points away from repair and toward escape. The research on this is large and consistent, much of it the work of June Tangney and her colleagues, and it finds that shame, far from motivating better behavior, tends to produce defensiveness, withdrawal, denial of responsibility, and the externalizing of blame onto others, and that when shame grows unbearable it converts into anger and aggression. Shame also, the research finds, interferes with empathy, turning the person inward onto his own distress rather than outward toward the one he has wronged. Guilt builds the bridge to the other; shame burns it.

The exercises trade in shame rather than guilt. They do not say to a child, you did a thoughtless thing and here is how to do better. They place him in a condition designed to deliver a verdict on what he is, a member of the group that oppresses or the group that is oppressed, and they make him feel the verdict in his body. For the favored child the lesson is that he is the kind of person who dominates; for the degraded child it is that he is the kind of person who gets crushed. Both are verdicts on the self, and both summon the defensive machinery that shame always summons. The strongest evidence for how badly this backfires comes from a setting far from the classroom. Tangney and her colleagues followed hundreds of jail inmates after release and measured which moral emotion predicted staying out of trouble. Guilt-proneness predicted less reoffending. Shame-proneness predicted more, and it did so through exactly the mechanism the

laboratory had identified, the externalizing of blame. The people most prone to global shame about themselves were the people most likely to offend again, because shame, rather than turning them toward responsibility, turned them toward defense and toward blaming everyone else. An intervention that shames is an intervention betting on the emotion the evidence says makes people worse.

Two further consequences of the shame design deserve note, one for each kind of child the exercise creates. For the favored child, the verdict is global in a way that summons the defense the research predicts. Told in effect that he belongs to the category of oppressors, he is handed a charge against his whole self rather than against any act he chose, and a self under global indictment does what selves under global indictment do, it protects itself, by denying, by minimizing, by deciding the accuser is unfair, by concluding that if this is what being good requires then he would rather not bother. There is also a subtler trap, the one researchers call moral licensing, in which a person who has performed a credential of virtue feels entitled to relax afterward, having banked the proof of his goodness. The child who has done the exercise, cried, and been praised for it has earned, in his own ledger, a license, and a person who believes he has proven he is not prejudiced is in some ways more dangerous than one still worried that he might be, because the worry was doing useful work. For the degraded child the harm is different and rarely counted at all. He has been taught, by a trusted adult and for a day, that he is the kind of person the world steps on, and while the lesson is framed as temporary the experience lands as permanent, because an eight-year-old files no feeling under provisional. He has rehearsed being the victim, practiced the posture of the one who absorbs and does not resist, and that rehearsal, like every rehearsal in this book, builds the thing it practices. The exercise meant to grow his sympathy for the oppressed has instead trained him to occupy the role of the crushed, which is no gift to anyone.

A third feature compounds the first two, and it follows from the staged audience the anatomy chapter identified. The exercise is a performance, watched, and a virtue performed for watchers tends to become a performance of virtue rather than the thing itself. When a child knows he is being observed for evidence of the right feelings, the incentive that forms in him is to display those feelings convincingly, to be seen having the breakthrough, to produce tears at the right moment and say the expected words, and the skill he builds is the skill of seeming, which is a different skill from becoming and often an enemy of it. This is the moral version of a problem I have written about elsewhere under the heading of manufactured sentiment, the way an audience converts real feeling into a product staged for approval, so that what looks like a room full of moved people is a room full of people performing being moved, each watching the others and matching the expected display. The exercise builds exactly this reflex. It rewards the visible signs of empathy and has no way to measure the invisible substance, and so it selects, across thousands of repetitions, for the children and adults who learn to

produce the signs. A culture trained this way fills with people fluent in the gestures of conscience, the right words, the right nods, the public display of having been transformed, and the fluency masks the absence of the transformation, because the gestures have come loose from any inner change and can be produced on demand by anyone who has learned the script. Performed virtue does worse than fall short of the real thing. It is a substitute that actively crowds the real thing out, because once a person can buy the social rewards of goodness with a convincing display, the harder and unrewarded work of becoming good loses its last external incentive.

Put these features together and the backfire is overdetermined. The exercise coerces, which triggers reactance, it shames, which triggers defense, and it performs, which trains display in place of change, and the three reinforce one another, so the child leaves armored rather than persuaded, having met an assault on his autonomy and his self with the predictable human response of resistance. None of this means the goal is hopeless or that prejudice cannot be reduced. Legault's autonomy group got better, after all, and the approaches that work are built on voluntary engagement, on contact between equals, on guilt that points to repair rather than shame that points to escape. The failure traced here is specific. It is the failure of the dramatic, coercive, shaming intervention, the failure of the method this book is about, and it is a failure the better alternatives do not share. A person can hold, at the same time and without contradiction, that reducing prejudice is a worthy and reachable aim and that the eye-color exercise is among the worst available ways to pursue it.

A serious objection deserves a hearing before this chapter closes. Moral growth, someone will say, is not supposed to be comfortable. Confronting one's own capacity for cruelty is hard and ought to be, and a pedagogy that spares the learner all discomfort spares him the thing that changes him. The objection is right about discomfort and wrong about which kind. Real moral growth often does hurt, and nothing here argues for a frictionless conscience. The distinction that matters, again, is the one between guilt and shame, and between a confrontation a person undergoes freely and an ordeal imposed on him. Reading a survivor's testimony and feeling the floor drop away is painful, and it works, because the pain is guilt and recognition, fastened to a real account of a real wrong, arrived at by a mind that agreed to look. Being sorted into a caste by a teacher and made to feel inferior on command is also painful, and it backfires, because the pain is shame and coercion, fastened to a manufactured wrong, imposed on a person given no choice. The two share the surface feature of discomfort and share nothing beneath it. One is the productive pain of a free mind meeting a hard truth. The other is the counterproductive pain of a captive self under assault. To defend the exercise by appealing to the necessity of discomfort is to hide the difference that matters behind the feature that does not, and the difference that matters is the whole question. Whether moral growth should be hard was never in dispute. What is in dispute is whether the

hardness is the kind that builds a conscience or the kind that arms it against the lesson, and the research answers that against the exercise.

If coercion, shame, and spectacle are what make these exercises iatrogenic, then the remedies point the other way, and they deserve a sketch now, because a diagnosis that names no cure invites the despairing conclusion that nothing helps. Something helps. The oldest and best-supported finding in this whole field is the contact hypothesis, the discovery that prejudice falls when people from different groups meet as equals in the pursuit of a shared goal, over time, with the backing of the institution around them. What changes a person is a real relationship rather than a staged ordeal, the colleague who becomes a friend, the neighbor whose life turns out to be a life, the slow accumulation of contact that turns the abstraction called the other into a particular person with a name. Notice how completely that remedy inverts the exercise. Contact is voluntary where the exercise is compelled, sustained where the exercise is brief, equal where the exercise imposes a hierarchy, and grounded in real people where the exercise depends on their absence. The same inversion holds for the emotional register, since the approaches that work appeal to guilt that points toward repair and to a person’s own freely held values, the autonomy-supportive path that Legault’s better condition traced. And where the exercise performs for an audience, the real work is mostly quiet, unwitnessed, and uncertifiable, which is one reason institutions prefer the exercise, since the exercise can be scheduled, photographed, and filed. The remedies share a humbling feature. They are undramatic. They cannot be finished in an afternoon, cannot be sold as a product, and produce no certificate, and that is among the deepest reasons the iatrogenic version persists while the effective version languishes.

The argument so far has been one of mechanism, an account of why this kind of intervention should be expected to backfire, drawn from the theory of iatrogenesis and the psychology of reactance and shame. Mechanism invites a fair challenge: show that it happens, not only that it should. The chapter on evidence meets that challenge, assembling the outcome research across the whole family of dramatic interventions, and the pattern it documents is the iatrogenic signature repeated across domain after domain: across crime and drugs and the workplace, the dramatic intervention that shames and coerces produces the reverse of its aim, while the quieter measures that persuade and include do better. The concept built in this chapter predicts that pattern, and the evidence confirms it.

One question hangs over all of this and deserves an answer before the chapter ends. If these exercises backfire as reliably as the theory and the evidence say, how have they survived for more than fifty years, defended by intelligent people of obvious good will. The answer lies in the same two features that do the damage, the certificate and the spectacle, which together make the failure almost impossible to see from inside. A program that issued no certificate would invite the question of whether it had worked.

This one closes that question by design, sending every participant away feeling changed, and a feeling of having changed is the thing an honest evaluation would have to look past, because it is produced whether or not any change occurred. The facilitator sees a room of moved people, tears and confessions and visible breakthroughs, and reads the display as proof, when the display is the product the spectacle was built to manufacture. The institution sees a completed program and a satisfied report and files the matter under solved. Nobody in the loop is positioned to notice that the prejudice went up rather than down, because nobody measured it, and the one signal everyone can see, the emotional intensity of the session, points the wrong way, since intensity reads as efficacy to people who have not checked. This is Illich's self-reinforcing loop in its purest form. The treatment produces a feeling of cure, the feeling of cure is mistaken for cure, the mistake licenses more treatment, and the only thing that could break the loop, a hard look at outcomes, is the one thing the loop is built to make seem unnecessary. The exercise survives less despite its failure than because of the way that failure conceals itself, and a harm that hides inside the appearance of success is the hardest kind to stop.

Return for a moment to the specimen, because the iatrogenic frame settles a question the early chapters left open, the question of the testimonials. Elliott and her defenders have always pointed to the grown participants who seek her out to thank her, the adults who say the exercise marked them for life and made them kinder, and that testimony has done more than any study to keep the method alive. The frame built in this chapter explains the testimony without crediting it. A person who underwent an intense, shaming, memorable ordeal in childhood and was told, then and afterward, that it had made him good, is a person carrying the certificate decades later, and his sincere conviction that it worked is the certificate speaking, the cultural iatrogenesis in its mature form. He feels changed because he was told he was changed at a moment of high emotion, and the feeling has hardened into a memory he trusts. What the testimony cannot tell us is the thing that matters, whether he is in fact less prejudiced than he would have been without the ordeal, and that is exactly the question his certainty forecloses. The grateful adult offers no evidence that the exercise works. He is a specimen of how it fails, a person so thoroughly issued his certificate that he has spent a lifetime defending the thing that handed it to him. The warmth of the testimony, which feels like the strongest argument for the method, is on this reading the clearest symptom of the disease the method spreads, the conviction of having been cured standing in for the cure.

The second claim is now established as concept and mechanism, with its empirical proof set aside for Part Five. These exercises are iatrogenic. They are the treatment that produces the disease, and they produce it by the routes the research marks out, coercion that provokes reactance, shame that provokes defense, and spectacle that trains the display of virtue in place of the substance, layered on top of the clinical, social, and

cultural injuries that Illich's framework exposes. Hold the whole indictment in view, because its force is cumulative. A single one of these flaws might be survived by a strong intervention. This intervention carries all of them at once. It coerces a belief that can only be chosen, it shames a self that answers shame by defending itself, it performs for an audience that rewards the look of change over the fact of it, and it certifies a completion that forecloses the work still to be done, and each flaw deepens the others. An intervention designed from scratch to fail at growing empathy, while convincing everyone involved that it had succeeded, would look much like the one a great many schools and workplaces still run by choice.

The cure manufactures the sickness, and then, in the cruelest turn of all, it issues a certificate of health, sending the patient away convinced he has been healed and closing off the search for a remedy that might have worked. That closing-off, the way the counterfeit drives out the real, is the heart of the third claim, and it rests on a condition I have mentioned at every stage and now have to confront directly, the strange fact that these exercises can only run in the absence of the people they are about. To that absence, and what it reveals, I turn next.

CHAPTER NINE

The Absence of the Subject

The first two claims of this book were borrowed from men who earned them. Freire supplied the inversion, the proof that the method belongs to the oppressor in any hands, and Illich supplied the iatrogenic frame, the account of a cure that manufactures the disease. The third claim is the one this book most makes its own, drawn from the two classrooms it began in, and it concerns a feature of these exercises so strange that once it is seen it cannot be unseen, though almost no one has named it whole. Others have caught pieces of it before. The disability movement reached the core of it in its refusal of the blindfold and its demand for nothing about us without us, and a longer tradition of standpoint argument circled the same ground, that a group's experience is not the property of outsiders to perform. What this book adds is the rule that runs the absence through the entire family of these exercises and the single question that exposes any member of it. That feature is an absence. The exercises that claim to teach children what it is to be hated for who they are require, as a condition of running at all, that the people hated for who they are be nowhere in the room. Their subject is missing from the lesson. In place of the real, the exercise installs a proxy, a stand-in trait that can be assigned and then removed, and the whole performance depends on that substitution. This chapter raises the observation from my own classroom to a general law of the genre and shows what the law reveals, which is that a ritual unable to survive the presence of its own subject was never about that subject at all.

Begin where the evidence is plainest, in the rooms themselves. Jane Elliott built her exercise in Riceville, Iowa, an overwhelmingly white town where her students, by her own account, had no contact with Black people, and she ran it on a third-grade class of twenty-eight children, every one of them white. The lesson was about the treatment of Black Americans, staged the morning after the murder of Martin Luther King, and there was not a single Black child present to be its subject, to recognize it, to correct it, or to say whether it bore any resemblance to a life lived under contempt. Elliott understood this perfectly well; it was the premise of her design. She reached for eye color precisely because her students had no contact with the thing the lesson was about and she needed

a trait already present in the room that she could charge with meaning, sort by, and reverse. The blue eyes and the brown eyes were no random pick. They were chosen because the real marker of the real hatred was absent, and a substitute was required to make the absent thing performable. My own classroom in Lincoln a few years later worked the same way, the armband doing the job the collar did in Iowa, a proxy pinned to children who would never wear the real thing because the real thing was not theirs to wear.

Stay in that Lincoln room a moment longer, because the absence had a texture I felt without yet having words for it. We were a class of white children being taught, with great moral seriousness, what it felt like to be on the receiving end of racial hatred, and the only thing in the room standing in for race was a strip of cloth the teacher handed out and collected. No one present had lived the thing the lesson was about. No one present could have raised a hand and said the teacher had it wrong, or right, because no one present held the experience against which wrong and right could be measured. The teacher's account of the suffering was the only account available, delivered with total authority to an audience in no position to check it, and that, I came to see much later, is the quiet engine of the whole enterprise. An exercise about a group's experience, conducted entirely without that group, cannot be corrected by anyone in the room, which means it cannot be shown wrong in any way the room could detect. Whatever the teacher said the experience was, the experience became, because the only witnesses who could have contradicted her were never invited. The armband opened no window onto a real life. It was a closed loop, a lesson about strangers performed for people who had never met them, graded by the one person who had built it.

Notice that the absence is no accident of geography, no unlucky fact that these particular towns happened to be white. The absence is structural, required by the design, and the exercise could not do what it does without it. Consider what a real member of the hated group would bring into the room. He would bring a life, a real history of the real thing, against which the staged version could be measured and found either false or grotesque. He would bring the standing to say that is not what it is like, or worse, that is nothing like what it is like, and his standing would outrank the teacher's, because he would be speaking from the inside of an experience the teacher was approximating from the outside. And he would bring the power to refuse the role assigned to his people's suffering, to decline to have it simulated by children for an afternoon and packed away at the bell. The exercise cannot tolerate that presence, because the exercise depends on the teacher's total authorship of the meaning, and a real subject is a second author who cannot be controlled. So the design requires his absence and substitutes a proxy he has no part in, and that substitution is the original sin from which the rest of the trouble flows.

The genre did not stumble into this. Absence was its founding purpose. The dramatic empathy exercise was invented, in Elliott's hands and in the hands of those who followed, to give people who had never suffered a particular hatred a manufactured taste of it, and the intended audience was always the group that does the hating, or is presumed to, rather than the group that endures it. The whole point was to take comfortable people and make them feel, for an afternoon, what the exercise asserted the uncomfortable feel all the time. A lesson built for that purpose has the absence of its subject written into its mission, because if the subject were present the lesson would have no one left to instruct; the people who live the hatred need no afternoon's simulation of it, and the exercise has nothing to teach them. The genre is addressed, by design, to everyone except the people it is about. That is a strange thing for a lesson in empathy to be, and the strangeness is no accident. A practice that can only be aimed at non-members of the group whose experience it teaches has already conceded that the group itself is beside the point, present in the lesson as a subject to be depicted and absent from it as an audience to be served.

Absence of the subject is one of three conditions the genre needs, and the three together describe the sealed environment without which none of these rituals can run. The first is the controlled room, the requirement that the exercise take place inside a space the facilitator wholly commands, a classroom, a training room, a workshop, with the doors closed and the ordinary world shut out. The second is the controlled audience, the requirement that the people present be a known and managed set, sorted into their roles by the facilitator, watched by the facilitator, and permitted to speak only on the facilitator's terms and at the facilitator's chosen moments. Third comes the absent or silenced subject, the requirement we have been tracing, that the people the lesson is about be either missing entirely, as in Riceville, or, when present, stripped of the authority to define their own experience and folded into the same managed script as everyone else. Seal those three and the ritual works. Break any one of them and it falters, because all three exist to protect the single thing the exercise cannot do without, the facilitator's unchallenged control over what the experience means.

The deepest reason the sealed room matters is that it makes the lesson unfalsifiable, and a lesson that cannot be proven wrong cannot be trusted to be right. In any honest inquiry, a claim is worth something only if some possible observation could count against it, some test it could fail. The empathy exercise removes that test by removing the only witness who could administer it. The teacher claims to be showing children what a hated life feels like; the one person who could say whether the showing is accurate, a person who has lived that life, is absent by design, so the claim floats free of any check. This is how the exercise can persist for half a century while teaching falsehoods about the experience it depicts. There is no one in the room with the standing to falsify it, and the design ensures there never will be. Set this against the way

understanding is built honestly, through contact with the real, where the other person is present precisely as the authority on his own life, free to say you have misread me, free to complicate the picture, free to be a particular human being instead of a category. That presence is what makes real understanding correctable and therefore trustworthy, and its removal is what makes the simulation a sealed and self-confirming thing, accurate by fiat because inaccuracy has been ruled out of the room. The exercise cannot survive scrutiny from its subject, and so it is built to ensure its subject is never there to scrutinize it.

The middle requirement, the controlled audience, repays a closer look, because the composition of the audience is part of the deception. Recall from the anatomy of the machine that these exercises always play to spectators, and that the real lesson lands on the watchers as much as on the sorted. Now add the fact this chapter has established, that the watchers are drawn from outside the group the lesson is about. The audience for the eye-color exercise is white children and the adults who later watch the footage; the audience for the disability simulation is the non-disabled; the audience for the privilege walk is the room of the comfortable. In every case the spectators are outsiders to the experience being depicted, which means the depiction answers to no one who could judge its accuracy and flatters everyone whose comfort it confirms. An audience of insiders would change the event entirely, would murmur and object and correct, would turn the performance into an argument. The genre cannot allow that, so it selects its audience as carefully as it selects its proxy, filling the seats with people disposed to be moved and emptied of people equipped to push back. The controlled audience and the absent subject are the same requirement seen from two angles, the demand that everyone in the room, performer and spectator alike, be someone the experience happened to in imagination only, never in fact.

That first requirement, the sealed room, works the same protective magic on the setting that the others work on the cast. These exercises cannot be performed in the open world, in a public square, on a city street, inside the actual hierarchy of a real society, because the open world is full of the people and the conditions that would break the spell. The exercise needs a bounded enclosure it can govern, a classroom or a training suite with the door shut, where the only social facts present are the ones the facilitator has imported and arranged. Out in the world a child assigned to the lower caste might pass, on his way home, a person who lives the real version of what he spent the day performing, and the contrast would expose the performance for what it was. Inside the sealed room no such contrast can intrude. The facilitator builds a miniature society from scratch, populates it with proxies, assigns the ranks, and runs the whole thing under laboratory conditions the messy real world would never permit. Far from a limitation the exercise reluctantly tolerates, the artificiality is something the exercise requires, because only inside an enclosure scrubbed of the real can a manufactured hierarchy be made to

stand in for the genuine one without the genuine one walking past the window and giving the game away.

The pattern is not confined to Elliott's exercise. It runs through the whole family of dramatic interventions, and once you know to look for the absent subject you find it everywhere. Take the privilege walk, the exercise in which a line of people step forward or back as a facilitator reads statements about advantage and disadvantage, ending spread across the room with the privileged at the front and the rest behind. The walk is performed, overwhelmingly, in rooms of the relatively comfortable, students and salaried employees, and the disadvantages it dramatizes belong mostly to people not present, represented by a step backward that the person taking it can retrace the moment the exercise ends. In some versions the participants are handed fictional identities to walk on behalf of, which makes the substitution total, a room of people performing the disadvantages of imaginary others while the others, real and unconsulted, are somewhere else. Take *StarPower*, the trading game that secretly rigs its economy so an early advantage hardens into a ruling class and then hands the winners the power to rewrite the rules, a simulation meant to teach how the powerful entrench themselves; its subject, the powerless it gestures at, is represented by a deliberately weak starting hand that the player knows is a game and that will be reshuffled when the bell rings. Across the genre the constant holds, that the suffering being taught is the suffering of the absent, rendered by a proxy the present can put down.

The clearest case of all comes from outside the world of race, and it has the advantage of a settled verdict from the people it concerns. For decades, schools and workplaces have run disability simulations, blindfolding the sighted to teach them about blindness, putting the able-bodied in wheelchairs to teach them about paralysis, piping distortion into headphones to teach them about deafness. The structure is identical to Elliott's. A removable proxy is assigned, an ordeal is undergone, and the comfortable are sent away believing they now understand. And the people who live with these conditions have spent those same decades explaining, in the research and in the disability-rights movement, that the simulations teach close to the opposite of the truth. A sighted person who wears a blindfold for an hour meets the terror and clumsiness of sudden blindness, the helplessness of a skill not yet learned, and comes away with the feeling that blindness is a catastrophe of constant fumbling. A person blind for years carries none of that helplessness, having built the skills and the life the blindfolded simulator will never glimpse in an hour, and the simulation, by staging only the shock of onset, produces pity and fear where it promised understanding and hardens the stereotype of the disabled person as helpless. The objection has been put to the test. In a pair of controlled experiments, researchers had sighted participants spend about half an hour blindfolded, pouring water and sorting coins and finding their way down a corridor, and then asked them to judge how capably a blind person could live and work.

The blindfolded group rated blind people markedly less able than a control group that had done no exercise at all, the simulation having taught its subjects to mistake the fumbling of the first half hour for the permanent condition. The result reads like the third claim caught in a laboratory, though a single experiment can illustrate a claim without sealing it. A method built to raise regard for a group lowered it, in a clean comparison against doing nothing, because it staged the shock of losing a capacity and called that the experience of living without one. This is one research line, and it enters as corroboration rather than foundation, the foundation being the long testimony of disabled people themselves, who have told the simulation industry for years that the blindfold misrepresents them and were saying so before any experiment confirmed it. The disability scholars have a phrase for the principle these simulations violate, nothing about us without us, and it names exactly the wrong at the center of this whole genre, the staging of a group's experience by and for people outside the group, with the group itself reduced to the thing depicted. What those scholars recommend instead is the thing the exercise is built to avoid, sustained contact with disabled people themselves, being with rather than pretending to be like. The disability case matters here because the verdict is in and the subject delivered it. When the people a simulation claims to represent are finally asked, they report that the simulation gets them wrong, and the same judgment waits for every other member of the genre the moment its subject is given a voice.

The same shape appears in the poverty simulations that nonprofits and universities run, in which the comfortable are handed a fictional low-income identity and a set of impossible monthly choices, the rent or the medicine, the bus fare or the meal, and made to feel the squeeze for an afternoon before returning to their own solvent lives. It appears in the tunnel of oppression, the walk-through installation in which participants pass through staged scenes of bigotry and violence, watching dramatized suffering that belongs to people who are, once again, not the ones walking through. In each, the proxy is temporary, the subject is represented rather than present, and the comfortable are the intended audience and the measure of success. The variations run on without end, the trait swapped from eye color to income to a scripted scene, and the engine underneath never changes. A group's hardship is staged for the benefit of people outside the group, by means of a stand-in the participants will shed at the door, and the people whose hardship it is have no seat in the room and no hand on the meaning. Once the pattern is named it is impossible to miss, and it is the pattern, more than any single exercise, that this book indicts.

The exception proves the rule, and it proves it in the most painful way available. Every few years a school somewhere stages a mock slave auction, a lesson in which children enact the buying and selling of human beings, and the story reaches the news because it has gone wrong, with children, sometimes the Black children in the class,

made to play the enslaved while others bid, and their families discovering afterward what had been done. These episodes are treated as aberrations, the work of a thoughtless teacher, and they are something more instructive than that. They are what happens when the genre's usual condition fails, when the real subject is present rather than absent, assigned his own historical role to his face. The result is never the tidy catharsis the method promises. What follows instead is injury and outrage, a Black child made to stand while his classmates bid on him, a family demanding to know how a school could have thought this was education. The auction erupts precisely because it violates the rule the other exercises obey. It puts the real subject in the room, and the lesson that depends on his absence cannot withstand his presence. The eye-color exercise and the privilege walk avoid the explosion by avoiding the subject, and their relative calm is purchased by the substitution the auction omitted.

The auction is the loud version of the collapse, the one that reaches the newspapers, and there is a quieter version that happens whenever a member of the depicted group is in a position to respond and does. Disabled people have told the simulation industry for decades that the blindfold misrepresents them, and the simulations continue. Black writers and educators have objected that having white children role-play slavery or segregation trivializes and distorts what their ancestors endured, and the exercises continue. When the subject does manage to speak, the genre's response is rarely to stop and listen, which would concede that the subject's judgment outranks the facilitator's, and is instead to absorb the objection as one more difficult feeling to be processed, or to wave off the objector as failing to appreciate the good intentions involved. The pattern confirms the thesis from a second direction. The exercises require the subject's absence to run smoothly and, beyond that, to survive criticism, and on the rare occasions when the subject is present and critical, the machinery treats his presence as the problem rather than his absence as the original error. A method seriously interested in the people it depicts would treat their objection as decisive. This one treats it as an obstacle, which tells you, again, whose experience the exercise was built to serve.

Hold the proxy up to the light and it gives the whole game away. The defining property of every stand-in these exercises use, the eye color, the armband, the step backward, the losing hand, is that it can be removed. At three o'clock the brown-eyed child takes off the collar and is a child again. The employee at the back of the privilege line walks back to the front and returns to his salary. A player with the losing hand pushes back his chair and goes to lunch. The proxy is temporary by design, and its temporariness is exactly what the real thing is not, because the Black child cannot take off his race, the poor family cannot return its disadvantage to the box, and the person living under contempt does not get to push back his chair when the lesson ends. The exercise trades in a suffering that switches off, and a suffering that switches off differs in kind from a suffering that does not, so far that calling them by the same name is the

central deception of the method. What the child learns to feel for is the removable version, the version he himself just wore and shed, and the lesson quietly teaches him that the real version is like that too, a hardship that could be set down, which is the opposite of the truth and a poor preparation for meeting the real.

Someone will object that simulation is a respectable way to teach, that pilots train in flight simulators and surgeons practice on mannequins and schoolchildren rehearse fire drills, and that no one calls those frauds. The objection helps, because the disanalogy is exact and it isolates what is wrong here. A flight simulator trains a skill, the handling of an aircraft, and the skill it trains is the real skill, transferable without loss to the real cockpit. A fire drill rehearses a procedure, the quick exit from a building, and the procedure rehearsed is the real procedure. Neither one claims to reproduce the inner experience of another human being, and neither one stages the identity of an absent group for an audience of outsiders. That is the line. Simulating a skill or a procedure is honest, because the thing simulated is a doing that carries over intact. Simulating a person's experience of being despised is something else, because the thing claimed to be reproduced is an inner life that does not carry over, that the simulation gets wrong in the specific way the disability research documents, and that belongs to real people who are not in the room and never agreed to the performance. The pilot in the simulator is practicing to become competent at a task. The child in the armband is being told he now knows what it is to be someone he has never met, on the authority of a person who has never met them either. One of these builds a real capacity. The other manufactures a false certainty about strangers, which is worse than ignorance, because ignorance at least knows itself.

Return to the switch-off, which does more damage than its convenience suggests. A child in the lower caste feels real fear, but underneath it sits a fact he never loses hold of, the knowledge that this ends at the bell, and that knowledge changes everything he feels, because fear with a known and near horizon is a different animal from fear without one. The person who lives under contempt has no horizon, no three o'clock, no certainty that tomorrow returns him to safety, and that absence is close to the whole of the suffering. What grinds a person down is rarely the single bad afternoon. It is the knowledge that the afternoons do not stop, and the exercise, by its mercy in ending, withholds the one feature that defines the thing it claims to teach. The child feels a sharp, bounded distress and is told it was a window into a boundless one, and the substitution of the bounded for the boundless is a mistake of category dressed as understanding.

There is a moral dimension to the absence that goes beyond the pedagogy, and it deserves naming plainly. To stage another group's suffering without that group present is to speak for people who have not been asked, to represent their pain in a form they did not choose, before an audience they will never address, and to certify the audience as having understood. The exercise takes the experience of the absent and converts it into a

teaching aid for the comfortable, a thing done to make white children or salaried adults feel something, with the feeling of the comfortable as the measure of success and the actual subject consulted at no point in the process. The people whose suffering is being borrowed cannot correct the portrait, cannot object to its use, cannot decline to have their lives turned into a fifty-minute lesson for someone else's growth. They enter the exercise as raw material rather than participants, present only as the thing being represented and absent as anyone with a say. A method that treats real people's real pain as a prop for the moral edification of others, without their knowledge or consent, has a name, and the name is not empathy.

The consent question deserves to be pressed, because it is where the method's self-image collapses most completely. These exercises present themselves as acts of solidarity with the marginalized, gestures of respect toward people who have been wronged. Solidarity, though, begins with asking, and the exercise asks no one. It takes a group's history of suffering, the most intimate and costly thing that group possesses, and converts it into curriculum without a word to the people whose history it is. Imagine proposing it to them directly, telling a Black family that their children's inheritance of fear will be reenacted by white third-graders for an afternoon so those children might grow, or telling a disabled adult that his daily reality will be approximated by a blindfolded colleague for a workshop. Put plainly and to the subject's face, the proposal reveals its presumption, the assumption that another person's pain is available for use, a resource the comfortable may draw on for their own improvement without permission or return. The marginalized in these exercises go unhonored. They are mined. Their experience is extracted, refined into a lesson, and consumed by an audience that walks away enriched while they themselves receive nothing, are asked nothing, and in most cases never learn the transaction occurred. A solidarity that never consults the people it claims to stand with deserves another name. It is appropriation wearing solidarity's clothes, and the absence of the subject is what lets the appropriation pass unnoticed, because the one party who might have named it was, as always, not in the room.

Notice who benefits, because the accounting reveals the exercise's true beneficiary, and it is not the absent subject. The person who gains from the ritual is the participant, who walks out carrying something he prizes, the sense of himself as someone who has confronted hard truths and grown. That sense is pleasant, even flattering, and it costs him nothing he cannot recover by three o'clock. He has paid an afternoon's discomfort and purchased a lasting credential of moral seriousness, a story he can tell about the day he learned what discrimination felt like, and the bargain holds because the discomfort was bounded and the credential is permanent. The absent subject, meanwhile, has gained nothing and risks something, the deeper entrenchment of a false picture of his life in the minds of people who now believe they understand it. So the exercise transfers a benefit from the people it claims to serve to the people it claims to instruct, flattering the

comfortable with a feeling of enlightenment while leaving the marginalized worse represented than before. A practice with that distribution of gains and losses is best understood by who it is for, and it is for the comfortable, who get to feel they have done something difficult and admirable without leaving the safety of a room their subject was never allowed to enter.

A defender of these exercises will say the absence is unavoidable, that you cannot bring a survivor of every hatred into every classroom, and that a simulation is the next best thing when the real subject cannot be there. The first half of that is true and the second half does not follow. The subject often cannot be physically present, granted, and it is false that a proxy is what his absence calls for, because there is a way to teach a group's experience that centers the group rather than erasing it, and it has been available the whole time. Its name is testimony. You bring the subject's own account, his words, his voice, his writing, his recorded face, the memoir and the interview and the documentary, and you let him narrate his own life to the children in his own terms, uncontrolled by the teacher and uncorrected by anyone in the room because he is the authority and they are the learners. Testimony keeps the subject present as a subject even when he cannot be present in body, and it asks the children to listen rather than to pretend, to receive a real account rather than to manufacture a false experience. The contrast with the proxy is total. The proxy replaces the subject with the comfortable; testimony hands the comfortable over to the subject. One stages the absent person's life for an audience that will never hear from him; the other is nothing except hearing from him. That the genre reaches for the proxy over the testimony, for simulation over the subject's own voice, tells you what it is for, because the testimony was always there to be used, and the only things it lacks are the drama and the flattering sense of having lived another's life in an afternoon.

This gives the book its sharpest diagnostic, the test that exposes the whole genre with a single question. Ask of any of these exercises whether it could run with its true subject in the room, present as a full author of the meaning, free to contradict the lesson and to refuse the role. The eye-color exercise cannot, because it needs the room white. The privilege walk could not survive a back row filled with people who live those disadvantages and who might describe them in terms the script never anticipated. And the auction cannot survive its subject at all, as the scandals show. A lesson that collapses the moment its subject walks in teaches nothing about that subject, staging instead a performance in the subject's absence, for an audience the subject is not part of, about feelings the subject does not get to judge. That is the third claim, and it joins the other two into a single indictment. The method is the oppressor's, it produces the harm it claims to cure, and it can only run where the people it claims to serve are not present to see it done.

The test also reaches the exercises not yet invented. Whatever the next well-meaning version turns out to be, if its design needs the subject gone or silent, it belongs to this family by the one mark it cannot lose, the empty chair where its subject should sit, and the moment someone fills that chair the lesson comes apart in the hands of the people running it.

There is a final symmetry worth drawing, between the absent subject of these exercises and the one real moral actor my own story produced. The exercises stage the suffering of people who are not there, and they reward the children who play along with the staging. The single act of conscience in either of my classrooms came from someone who was there, fully present, refusing. When I declined the armband, I simulated no one's experience; I exercised my own judgment, in my own person, about a wrong being done in front of me. The contrast holds the whole book in miniature. These rituals are built around an absence, the missing subject, and they punish a presence, the child who shows up as himself and says no. What they train children to value is the performance of borrowed feeling for people who are not in the room, and what they train children to suppress is the real feeling of a person who is. A pedagogy that honored the subject would do the reverse, would bring real people into real contact and would prize the honest response of a present conscience over the rehearsed response to an absent one. The exercise inverts that, as it inverts everything, and the inversion is why the absence at its center is the right place to end the case against it.

One irony remains, and it is the deepest. Empathy, the thing these exercises exist to produce, is a reaching toward someone real, a taking-seriously of a particular person whose inner life differs from your own, and that reaching needs the other present in the encounter as the author of his own portrayal. Presence does not require a body in the room. A reader can reach toward a person on the page, a listener toward a recorded voice, a grandchild toward an ancestor whose letters survive, because in each of these the other still authors his own account, still sets the terms on which his life is shown, and the empathy travels toward a self that is genuinely there in the testimony even after the person is gone. The thing empathy cannot survive is the other's replacement. A body absent from the room is no obstacle, so long as the person still authors what is shown of him, while a person swapped for a proxy the facilitator invents and controls leaves nothing real on the far end to reach toward. That swap is what the simulation performs. It removes the actual other, the one who could say you have misread me, and hands you in his place a figure scripted by the person running the room, a stand-in calibrated to produce the feeling the lesson wants. Reaching toward that figure is no empathy at all, only a reflexive self-stimulation, the warmth of your own moved feelings mistaken for contact with another mind, because there is no other mind on the far end of it, only the facilitator's puppet wearing a stranger's coat. This is why the absence of the subject is less one flaw among the genre's several than the flaw that swallows the rest. The

remedies the last chapters defend, the memoir assigned and argued over, the survivor brought in to testify, the neighbor met as an equal, all keep the other as the author of himself, present in the one sense empathy requires. The exercise alone replaces him, and a lesson in empathy that begins by replacing its object with a puppet is a contradiction in its first term, a reaching toward no one dressed up as a reaching toward someone.

Step back and the three claims of this book lock together into one mechanism, each naming a different face of the same object. The inversion explains the method, why a lesson in justice is built from the oppressor's tools. The iatrogenic frame explains the result, why the cure produces the disease. And the absence of the subject explains the conditions, why the whole thing can run only in a sealed room with its true subject shut out, and why it would shatter the moment that subject walked in. Read together they describe a ritual wrong in its means, counterproductive in its effects, and dependent for its operation on excluding the people it claims to serve, a closed and self-protecting thing from which the only witnesses who could expose it have been removed by design. What remains is to explain how a performance with these properties became an institution, a product bought and sold, a standard fixture of schools and corporations and government offices. That story, the turning of a classroom stunt into an industry, is the subject of the chapters that follow, and it begins with the woman who first discovered that a nation would pay to watch the lesson she had built in a room with no one in it who could tell her what she had gotten wrong.

PART FOUR

The Industry

CHAPTER TEN

From Stunt to Standard

A method becomes a standard in one of two ways. It can earn the position by working, by demonstrating across trials that it produces the result it promises, the way a drug earns its place in a formulary. Or it can arrive at the position by another route entirely, by serving the needs of the institutions that adopt it well enough that the question of whether it works never has to be answered. The dramatic exercise took the second route. In the half-century since Jane Elliott first divided a third-grade class by eye color, her improvisation grew into a personal career and then into the seed of an industry, a multibillion-dollar business that now reaches into most large American companies and nearly every corporation on the Fortune list. The growth was real and the spread was fast, and at no point along the way did the method clear the bar that a standard is supposed to clear, the demonstration that it does what it claims. This chapter follows the path from the classroom to the industry and asks how an unevidenced exercise became a fixture of American institutional life, and the answer, worked out over the pages that follow, is that it served its buyers in ways that had nothing to do with whether it taught anyone anything.

Begin with the woman herself, because her path is the industry's in miniature. Jane Elliott conducted the exercise for the first time on April 5, 1968, the morning after the murder of Martin Luther King, and she did it out of grief and anger, a rural Iowa teacher trying to make her all-white class feel something of what the country had just done. The compositions her students wrote afterward ran in the *Riceville Recorder* on April 18 under a headline about how discrimination feels, the Associated Press picked the story up, and the wire carried it to a national audience. Within the year she was a guest on *The Tonight Show*, where her account of the exercise lit up the switchboard with calls, many of them furious that she had tried such a thing on white children. The fame cost her at home. Her neighbors turned on her, fellow teachers walked out of the lounge when she entered, her children were harassed, and someone poisoned the family dog. None of that deterred her. On December 15, 1970, she staged the exercise for an audience of adults, educators and physicians and social workers and civic leaders, at a White House

Conference on Children and Youth, and the same year a film crew captured it for an ABC documentary called *The Eye of the Storm*, which won a Peabody and made her a national figure. A second documentary, the *Frontline* episode *A Class Divided*, reunited her original pupils in 1985 and carried the exercise to another generation. By the middle of the nineteen-eighties the demands of her growing reputation had pulled her out of the classroom for good, and she became what she has been ever since, a full-time professional in the business of running her exercise for a fee.

The students who went through the original exercise have testified, across the decades, to its lasting power, and their testimony deserves a careful hearing, because it is the strongest evidence the method has and it does not prove what it appears to prove. Former pupils describe the experience as something seared into them, a thing they carry through life and cannot unlearn, a day that became part of who they are. Those accounts are sincere and there is no reason to doubt them, and they establish beyond argument that the exercise produces a powerful and durable impression. What they do not establish is that the impression is an education in the structure of racism, which is a different thing from a vivid memory of a hard morning. A person can be marked for life by an experience and still have learned from it something other than what the experience was meant to teach, the way a frightening day can leave a lasting fear without leaving any understanding of what was feared. The testimony establishes that the exercise is memorable, which no one disputes, and the book has argued from the start that memorability is the wrong measure, the gauge that reads high while the thing that matters goes unmeasured. Millions who have watched the exercise on film or sat through it in a workshop carry away the same strong impression Elliott's third-graders did, and the strength of the impression has been mistaken, across the decades, for proof that the impression is true. The fame of the exercise rests on that mistake. A method that reliably produces an unforgettable experience will always look like it works, because the experience is the thing people remember and the experience is what they report, and the question of whether anyone's understanding or behavior changed at all disappears behind the vividness of the day.

The transformation that matters most is the one that happened to the exercise when it left the school. Elliott redeveloped the classroom activity for the corporate market, and the corporate version was sold on a promise the original never carried, that subjecting employees to a few hours of manufactured hierarchy would build teamwork, improve the workplace, and help the bottom line. Her client list became a roster of the American establishment, the companies General Electric, Exxon, AT&T, and IBM among them, and the agencies the FBI, the Internal Revenue Service, the Navy, the Department of Education, and the Postal Service. Consider what that list represents. An exercise born in a single act of moral witness, devised in one night by a grieving teacher the day after a civil-rights leader was shot, had become a service purchased by Fortune 500

corporations and federal bureaucracies and marketed as a route to profit and cohesion. The moral content did not survive the move intact. A lesson that began as an attempt to make white children reckon with the cruelty of racism became, in its corporate form, a team-building product, its political and historical substance drained out and the bare mechanism, the staged humiliation of one group by another, left to function as a workshop exercise. The thing being sold was the experience, the strong feeling the exercise reliably produces, repackaged as a corporate good and priced accordingly.

There is a particular violence in what the market did to the exercise, and naming it requires no sentimentality about the exercise itself, which this book has spent nine chapters criticizing. A thing devised as moral witness was turned into a commodity, and the turning changed its nature. When Elliott separated her class the morning after the assassination, whatever the flaws of her method, she was making a claim about justice, trying to force a reckoning with a specific historical wrong at a specific raw moment. When the same exercise is sold to a corporation as a way to improve teamwork and lift the bottom line, the claim about justice has been replaced by a claim about productivity, and the historical wrong has become a backdrop for a workshop on employee cohesion. The mechanism stayed the same while the meaning was hollowed out, and what remained was the part the book finds most objectionable, the staged humiliation, now stripped of even the serious moral purpose that had once been offered to justify it. A federal agency that buys a session of this exercise does no reckoning with the history of American racism; it checks a box, purchasing a few hours of manufactured discomfort because the manufactured discomfort has a market and a reputation and fits into a budget line. The exercise that began as one woman’s attempt to make a moral point became, in the end, a product whose selling point was the reliability with which it could make people feel something, the feeling itself the good on offer, severed from any claim about whether the feeling taught the truth. That severing is the signature of the commodity, and it is what the industry did to the method at scale, turning a moral gesture, however misguided, into a service rendered for a fee.

Before the exercise became a corporate product it became a television product, and the broadcast phase did as much as anything to turn the stunt into a standard. The ABC documentary that filmed Elliott’s classroom carried the exercise to a national audience and won the industry’s highest award, and the later *Frontline* treatment, which reunited her grown pupils, became one of public television’s most requested programs. Elliott appeared on the major talk shows, repeatedly, and versions of the exercise were staged for television in other countries, the format reproducing itself on screen after screen. Television found in the exercise exactly what television wants, a compact and emotionally violent spectacle with a clear arc, the sorting, the cruelty, the tears, the reversal, the lesson, all of it delivered in an hour and guaranteed to produce a strong reaction in the viewer. The exercise was, in this sense, made for the camera, and the

camera repaid it by carrying it everywhere. What the broadcasts could not do was establish that the exercise worked, since a viewer moved by an hour of televised cruelty has had the same experience the participants had, a powerful feeling that proves nothing about understanding, and the millions who watched and were stirred became, in effect, a vast audience certifying the method by their emotion. The television phase converted the exercise into a famous thing, and fame, in the institutional world that would later buy it, is its own credential. A practice that everyone has heard of and that reliably moves people does not have to prove it works to become a purchase a cautious manager can defend, and the broadcasts manufactured precisely that reputation, the reputation of a method too well known and too affecting to question.

What the corporate version does to the people in the room has been documented, and the documentation confirms that commercialization changed the packaging while leaving the cruelty intact. In the corporate exercise Elliott divides a mixed group by eye color and subjects the blue-eyed members to a sustained regime of humiliation and contempt, and the effect on them is swift and visible, the targeted workers growing distracted and despondent within a few hours, stumbling over instructions a person would ordinarily follow without difficulty. That is the mechanism the original exercise ran on, transferred without modification to adults on a company’s payroll, and the marketing wrapped around it, the language of teamwork and winning together, sits oddly against the documented reality of employees reduced to confusion and misery by lunch. The corporate buyer was sold a tool for building cohesion and received a procedure for manufacturing distress, and the gap between the two is the gap between what the industry says it sells and what it delivers. The appetite for that procedure has proven to be global. Versions of the exercise have been staged for television audiences in Britain and elsewhere, the format crossing borders as readily as it crossed the line from classroom to boardroom, and the international reach tells its own story about what is being bought. A method that travels that well satisfies no local need for a particular curriculum; it satisfies a universal institutional appetite for a strong, brief, defensible experience that looks like action on a hard problem, an appetite that exists in every country with workplaces to manage and liabilities to limit. The exercise spread across the world for the same reason it spread across American industry, because the manufactured experience it reliably delivers is a thing institutions everywhere want to be able to say they have provided.

Elliott is widely described as the forerunner of diversity training, and the description is accurate in a way that indicts the field rather than honoring it. Her exercise supplied the template the industry would follow, the conviction that the way to teach people about discrimination is to manufacture an experience of it, to put participants through a staged version of cruelty and trust the intensity of the feeling to do the educational work. When companies later built unconscious-bias trainings, the Blue Eyes and Brown Eyes

exercise was among the things they reached for, folding the classroom stunt directly into the corporate curriculum. The personal career and the industry are therefore one story rather than two, the same method propagating along two tracks, Elliott carrying it from stage to stage in person while the broader business of diversity training grew up around the same core assumption she had pioneered. To understand how the stunt became a standard, the chapter has to leave Elliott for a moment and trace the industry that formed in her wake.

The diversity-training business grew out of the civil-rights law of the nineteen-sixties, and its earliest form was a creature of compliance. Title VII of the Civil Rights Act of 1964 made it illegal for employers of any size above a small threshold to discriminate by race, color, religion, sex, or national origin, and as discrimination complaints reached the new Equal Employment Opportunity Commission, the remedy the agency often imposed was training, an order that a company instruct its workforce in non-discriminatory behavior. The training of that era was a legal exercise, a way to inform employees of the law and to fold women and minority workers into an unchanged corporate culture. The business changed character in the late nineteen-eighties. In 1987 a report titled Workforce 2000 forecast that the coming American labor force would be more female and more diverse, and the forecast gave employers a reason beyond compliance to think about difference, a reason framed in the language of business advantage. Around that prediction grew the concept of managing diversity, advanced in 1990 by Roosevelt Thomas in an influential article that recast diversity as a matter of competitive strategy rather than legal obligation. The field moved through a predictable sequence after that, from compliance to the management of diversity as a business asset, then to a broad sensitivity training aimed at all employees, then to the cultural-competence and unconscious-bias frameworks of the new century, each phase larger and more ambitious than the last.

Two features of that history are worth drawing out, because they shaped everything the industry became. The first is its experiential inheritance. The diversity training that grew up in the corporate world drew on an older tradition of sensitivity training, the encounter groups and workshops of the mid-century that aimed to change people by putting them through an intense interpersonal experience rather than by teaching them anything in the ordinary sense. That inheritance married naturally to Elliott’s exercise, since both rested on the same wager, that a manufactured experience would accomplish what instruction could not, and the industry that resulted carried the wager into every conference room it entered. The second feature is the shift in who the training was supposed to fix. The earliest programs targeted specific groups, the women and minorities who were to be assimilated into an unchanged culture, and the later programs, in the name of a more sophisticated understanding, expanded to target everyone, on the theory that all employees harbor biases requiring correction. That expansion looked like

progress and functioned like a market, since a training that everyone must take is a far larger sale than a training for a few, and the universalizing of the audience multiplied the industry's customers at a stroke. The unconscious-bias frameworks of the new century completed the move, supplying a scientific vocabulary for the claim that every person carries hidden prejudices a workshop can surface and correct. That vocabulary lent the enterprise an air of rigor it had not earned, since the science of unconscious bias is contested and the link between any measured bias and any actual behavior is weak, but the appearance of rigor was enough, and the industry wrapped its old experiential wager in the language of cognitive science and sold it as a discovery. The history is a story of a method in search of a justification, finding a new one in each decade, while the method itself, the manufactured experience, stayed what it had always been.

Underneath the shifting vocabulary, the engine driving the industry's growth was liability. After firms learned that discrimination suits could cost millions to settle, they began adopting training programs in earnest, and a great many of those programs were designed less to change anyone than to protect the company, to create a documented record that the employer had taken discrimination seriously and could point to in court. The training existed, in large part, to preempt lawsuits by demonstrating diligence, which had a decisive effect on how the training was judged. A program adopted to manage legal risk does not need to work in order to succeed at its real job, since its real job is to exist, to be on file, to be the thing the company gestures at when a complaint arrives. The industry's own metric gave the game away. Diversity training came to be evaluated, by the companies that bought it, by the number of employees who had been put through it rather than by any measured effect on bias or behavior, a box to be checked and a headcount to be reported. A practice measured by how many people attend it rather than by what attendance accomplishes is a practice that has stopped pretending to be education and become a ritual of compliance, and that is what much of the industry quietly is.

The legal logic deserves to be spelled out, because it explains the otherwise baffling indifference to results. Over the nineteen-nineties the courts developed a doctrine under which an employer could limit its liability for harassment by showing that it had taken reasonable steps to prevent and correct the conduct, and an anti-harassment training program, documented and mandatory, became the readiest proof of reasonable steps. The training thereby acquired a value entirely separate from its effect on behavior. A company that could show every employee had completed the program held a shield in court whether or not the program changed a single mind, and the shield worked precisely because it was documented, the certificate of completion functioning as legal evidence of diligence. Once the training's purpose was understood this way, its content became almost beside the point. The company needed the training to have happened, to be recorded, to be producible on demand, and whether it accomplished anything was a

question with no legal consequence and therefore no institutional urgency. This is the deepest reason the industry could grow for decades on no evidence. The buyer sought no reduction in bias, which it had no way to measure and no legal need to achieve; it sought a defense, a piece of paper with protective power, and a defense does not have to work in the way education works in order to work in the way the buyer needs. The training is bought to be filed, and a thing bought to be filed will be bought forever, regardless of what it does to the people filed through it, because the filing is the point and the people are the occasion.

A market organized this way breeds a particular kind of supplier. Once the buyers were purchasing documentation rather than results, the sellers had no reason to compete on results, which could not be measured and did not matter to the sale, and every reason to compete on the things that did matter, the strength of the experience, the polish of the materials, the ease of scheduling, the reassurance the program offered an anxious legal department. A whole profession of diversity consultants grew up to meet that demand, and the profession acquired, as professions do, a stake in the continuation of the practice that sustained it. No industry investigates whether its own product is necessary, and the diversity-training industry had less incentive than most, since the absence of any outcome measure meant there was no feedback loop that could ever have killed an ineffective program. A drug that does not work eventually fails its trials and leaves the market. A training that does not work faces no such reckoning, because no one is running the trial, and the buyer's satisfaction depends on the training's having occurred rather than on its having succeeded. The result is a sector insulated from the ordinary correction that removes useless products from the world, a sector that could and did sell the same unevidenced method for half a century, refining its delivery and expanding its reach while never confronting the question that would matter if results were the point. The consultants are not frauds, for the most part, any more than Elliott was a fraud. They believe in what they sell, as people who profit from a practice tend to believe in it, and their sincerity is part of what keeps the machine running, since a sincere seller is more persuasive than a cynical one and the buyer wants to be persuaded. The market and the method sustain each other, the method supplying the experience the market sells and the market supplying the money the method needs to survive its own lack of results.

The scale the business reached is plain. By recent estimate, American organizations spend something on the order of eight billion dollars a year on diversity training, more than half of all midsize and large companies run some version of it, and nearly every firm on the Fortune 500 reports using mandatory training of this kind. A method that never demonstrated it worked had become, within a few decades, a near-universal feature of corporate life, purchased at enormous expense by the most sophisticated organizations in the country. The puzzle is genuine. These are companies that measure everything, that subject their marketing and their logistics and their hiring to relentless

quantitative scrutiny, and yet on this one practice they spent billions for decades without the evidence that would justify the smallest of their other expenditures. The explanation cannot be that the training works, because the question of whether it works was never the one being asked.

That resolution is the one this book has been building toward. The diversity-training industry persists for the same reasons the classroom exercise persists, because it serves the needs of the institutions that buy it, and those needs have little to do with education. Such training is cheap relative to structural change, and far easier to arrange. It is portable, sold as a package and delivered on a schedule. The exercise is memorable, leaving participants with the strong feeling it reliably manufactures, a feeling that reads to a manager as impact. It is defensible, a documented program that answers the question of what the company is doing about discrimination. And it is countable, reducible to the headcount that the institution reports as evidence of effort. Every one of those is a virtue of the training as an institutional artifact, a thing that fills a requirement and generates a record, and not one of them is a virtue of the training as a way of changing how people think or behave. The industry scaled the classroom exercise's defects along with its method, carrying the inversion and the manufactured feeling and the absent subject into ten thousand corporate conference rooms, and it added the one ingredient the classroom lacked, a price, which gave a whole sector of consultants a stake in the method's survival regardless of its results.

There is a further function the training serves, and it may be the most important of all, the function of substituting for the thing that would in fact work. Changing the demographics of an institution is hard, slow, expensive, and politically fraught. It requires altering how people are recruited and hired and promoted, confronting the managers who resist, redistributing opportunity in ways that some will experience as loss, and sustaining the effort across years without the gratification of a clean finish. A training, by contrast, is none of those things. A training can be scheduled, completed, and reported in a quarter, and it leaves the institution's actual arrangements undisturbed while generating the appearance of action. That is its deepest appeal to the organization that buys it, that it lets the institution be seen to address a problem without undertaking the difficult changes that addressing the problem would require. The training functions, in this respect, like an indulgence, a payment that discharges the sense of obligation without altering the conduct that created it, and an institution that has run its workforce through a diversity workshop can tell itself, and the world, that it has done something, when what it has done is purchase the feeling of having acted. The displacement is the harm that compounds all the others. Every dollar and every hour spent on the workshop is a dollar and an hour not spent on the structural work that might have changed the institution, and the workshop's existence, by satisfying the demand for visible effort, removes the pressure that might otherwise have forced the harder change. Beyond

failing to help, the training occupies the space where help would have gone, and absorbs the energy that help would have required, which is why a sincere commitment to diversity is better served by no training at all than by the kind the industry sells.

The individual employee's experience of the training completes the parallel with the schoolroom. A worker sent to a diversity workshop undergoes, in compressed form, what the schoolchild undergoes, a staged and emotionally charged few hours that end on schedule, after which he returns to his desk and his life, the experience sealed off behind him. Whatever feeling the workshop produced switches off as the workshop ends, the same way the blue-eyed child's distress switched off when the collar came off, and the worker carries back into the corridor a memory of an afternoon rather than a changed pattern of behavior, soon filed with every other mandatory session he has sat through. The institution, meanwhile, has acquired its certificate, the record that the worker attended, and that certificate does the same quiet damage at corporate scale that the schoolroom version did. An organization that has run its people through the training believes, and announces, that it has addressed bias, and the belief licenses exactly the complacency the evidence describes, the conviction that the problem has been handled and the work is done. A company persuaded it has fixed its culture stops looking for the discrimination still operating inside it, because looking would contradict the certificate it has already issued itself. The training thus produces, reliably, the one outcome most corrosive to the goal it claims to serve, a false confidence that the problem is solved, distributed across the workforce and ratified by management. The schoolroom issued its participants a private certificate of conscience, a story each child could tell himself about the day he learned about racism. A corporation issues itself the same certificate at the level of the institution, a documented program it can point to as proof of virtue, and the certificate is worth exactly what the schoolroom's was, which is nothing, except as a shield against the suspicion that nothing has changed.

Those results, which the chapter on evidence examines in full, are not encouraging, and they are worse than merely disappointing. The most thorough study of the question, an analysis of more than eight hundred American firms across three decades, found that mandatory diversity training of the standard kind tended to produce no gains in the representation of women and minorities in management, and in a number of cases was followed by declines, the share of women and of black employees in management falling in the years after firms adopted it. Those same researchers found that the conditions under which the training usually fails are the ordinary conditions of its delivery, when it is made mandatory, when it invokes the law and the threat of liability, and when it is aimed only at managers rather than offered to everyone, and the great majority of firms run programs that fall into at least one of those categories. The mechanism is the one the book traced in the resistant schoolchild, now operating in a conference room full of adults who did not ask to be there. A worker compelled to

attend a session that treats him as a probable offender and frames the matter in terms of rules and legal exposure tends to respond with resentment rather than reflection, and the resentment can settle into a hardened hostility toward the groups the session meant him to regard more generously. Training built to reduce bias activates it. The federal agency charged with enforcing anti-discrimination law reviewed what was known and reported that it could find no proof such training prevents the harassment it targets, and a research literature reaching back to the nineteen-thirties has struggled to show that anti-bias instruction reliably reduces prejudice, the better-designed studies finding the smaller effects, which is the signature of a phenomenon that shrinks the more carefully it is measured. The preview is enough to sharpen the puzzle: the industry reached a near-universal scale without ever proving its worth, and did so while the evidence accumulated that it often does harm.

A fair objection has to be met before the indictment goes further, because the industry is not one thing. Much of what corporations buy under the name of diversity training is no staged hierarchy at all, only a lecture, an online module, a compliance briefing on policy and law, didactic and dull rather than dramatic, and it would overstate the case to call every such session a descendant of the eye-color exercise. The honest map has two strands. One is the dramatic strand proper, the privilege walk, the trading game, the immersive workshop, and Elliott’s own corporate exercise, which carry the staged humiliation directly into the conference room. The other is the larger didactic strand, the slideshow and the policy module, which stages nothing and merely instructs. This book’s sharpest charges, the manufactured cruelty and the engineered humiliation, fall on the first strand, and it would be dishonest to pin them on the second. What unites the two, and what lets the chapter treat the industry as a whole, is the structure they share whenever the training is imposed. Both strands are typically mandatory, both frame the attendee as a probable offender, and both provoke the reactance that mandatory accusation provokes, which is why even the dull compliance module shows the backfire the research found. The dramatic strand earns the harsher verdict and the didactic strand the milder one, but the coercion that drives the backlash runs through both, and a company serious about the goal has reason to be wary of the whole category, less because every session stages a cruelty than because the ones that do harm directly and the ones that merely lecture share the compelled, accusatory frame that turns an adult against the lesson he was assembled to learn.

One more thread ties the industry to the book’s argument, and it is the thread of the absent subject. The diversity training is conducted about the marginalized, and it is conducted, overwhelmingly, without them in any position of authority over it. Its subjects, the women and the minority employees still underrepresented in the management ranks, are present in the workshop as its occasions and its topic rather than as the architects of the institution’s actual practices, which remain in the hands of the

same managers the training processes and releases. The session addresses their disadvantage as a matter of other people's attitudes, a thing to be adjusted in an afternoon, and in doing so it diverts attention from the structural arrangements, the hiring and the promotion and the pay, where their disadvantage resides. Cruelest of the findings is the one already noted, that the training is sometimes followed by a decline in the representation of the people it was meant to help, which means the exercise performed in their name can leave them worse off than no exercise would have. The subject is absent from the diversity industry in the same way the subject is absent from every ritual this book has examined, present as the thing discussed and the reason given, absent as the authority who might say whether the discussion serves her, and the industry's occasional tendency to reduce her numbers while professing to raise them is the absence of the subject producing, at the scale of the economy, the same perverse result it produces in a single classroom.

The path from Elliott's classroom to the eight-billion-dollar industry is worth holding up as a general lesson, because the conditions that produced it are not unique to diversity training. An unevidenced practice becomes an entrenched standard whenever three conditions hold together. The practice must produce a strong and reliable experience, something participants feel and remember, since the felt experience is what will be mistaken for proof. There must be an institution with a need the practice can appear to meet, a requirement to be seen acting, a liability to be managed, a box to be checked. And there must be no measurement of outcomes, no trial that could ever return the verdict that the practice does not work, since measurement is the only thing that reliably kills a useless standard. Where those three conditions meet, a practice can persist indefinitely on no evidence at all, sustained by the experience it manufactures, the institutional need it serves, and the absence of any test it could fail. The dramatic exercise satisfies all three perfectly, which is why it spread as far as it did and lasted as long as it has, and the diversity-training industry is only the largest and most lucrative instance of a pattern the whole book has been describing. The stunt became a standard through the presence of those three conditions, all of them in place then and in place still, wherever the manufactured experience is for sale, and through nothing more demonstrable than that.

Seen from inside the institution, none of this is irrational, which is what makes it so hard to dislodge. A manager who buys the training is meeting a real need with a real product, acquiring the documentation, the visible action, and the defensible record his position requires, and he is doing so at a price the budget can bear. The irrationality appears only when the practice is measured against the goal it claims to serve, the reduction of discrimination, which is the one standard by which it fails and the one standard no one in the transaction is required to apply. That is the trap the whole genre sets. It aligns the incentives of the people who buy it with the continuation of a practice

that does not work, so that every individual decision to purchase the training is sensible and the sum of those decisions is a multibillion-dollar industry built on a method the evidence condemns.

None of this is an argument against the goal the industry was built to serve. Workplaces should be free of discrimination, the civil-rights law that pushed companies toward that freedom was right and overdue, and the integration of American institutions remains an unfinished and necessary project. The failure of the dramatic method is a reason to pursue that project with tools that work rather than a reason to abandon it, the approaches a later chapter will describe, which engage people in solving the problem rather than processing them through a workshop, expose them to one another across the lines of difference, and hold them accountable for outcomes rather than attendance. Jane Elliott was sincere and she was brave, and she paid a real price for confronting her town with its racism, and none of that is in question. What is in question is the method she pioneered and the industry that monetized it, a method that scaled into a standard without ever earning the title, sustained by a business with every reason to keep selling it. The industry sells something more flattering than a method, though, and that is the subject the next chapter takes up. It sells a conscience, a way for an institution to be seen to care, purchased by the session and filed away as proof, and the manufacture of that purchased conscience is where this account turns next.

CHAPTER ELEVEN

The Rented Conscience

The previous chapter ended on a claim that needs defending, that what the diversity industry sells is a conscience. The word is exact and the transaction is old. An institution that buys the training acquires a way to be seen to care, a documented act of moral seriousness it can display to anyone who asks, and the displaying is the product, severed from any change in how the institution behaves. There is an ancient name for the manufacture of a response that is paid for rather than felt, and the name is the claque, the body of hired applauders who filled the theatres and opera houses of the old world with acclaim that no one in the seats felt. The claque sold the appearance of enthusiasm to performers who could not be sure of earning it, and the rented conscience sells the appearance of virtue to institutions that have not done the work of deserving it. The parallel runs deep, and it carries a warning the rest of this chapter will develop, that a purchased conscience is worse than empty in the way purchased applause is empty. It is worse, because the evidence on how people use a credential of virtue shows that the purchased conscience licenses the wrong it was bought to prevent.

The claque is among the oldest devices for manufacturing a public response. When the emperor Nero performed, he was attended by a body of some five thousand soldiers who chanted their approval on cue, an organized claque assembled to drown the silence his acting would otherwise have met. By the early nineteenth century the practice had been industrialized, an agency opening in Paris in 1820 to supply applauders on demand, organized under a chef de claque who studied the work in rehearsal and drilled his hired forces to deliver their effect on the right lines. His deeper function was to lead rather than to applaud, to show the real audience around him when to clap and when to weep, so that the paid response would seed an unpaid one and the genuine spectators, taking their cue from the planted ones, would supply the enthusiasm the planted ones had only simulated. This is the most ingenious and the most corrupting feature of the device, that it altered the real audience’s response rather than merely adding false acclaim, teaching people who had come to judge for themselves to judge instead by the manufactured signal around them. A performer served by a good claque lost the one

thing the audience was there to give him, an honest reaction, because he could no longer tell which part of the applause was real, and the rented conscience corrupts the feedback an institution needs in exactly the same way, replacing the signal that would tell an organization whether it was doing right with a manufactured one that says it is.

What the claque sold was never applause, exactly, since applause is cheap and anyone can clap. What it sold was the appearance of spontaneous acclaim, the impression that an audience had been moved of its own accord, manufactured for performers who needed that impression and could not rely on producing it honestly. The whole value of the device lay in its being mistaken for the real thing. A claque that the audience knew to be paid was worthless, because the acclaim it produced derived its force entirely from the assumption that it was unpaid, the assumption that those cheering hands belonged to people who cheered because they were moved. The claque was, in this precise sense, a counterfeit, a manufactured copy of a genuine response, passing as the genuine article and losing all its worth the moment the difference was detected. Hold that structure in mind, because it is the structure of the rented conscience exactly. The institution that buys a conscience is buying the appearance of a genuine moral response, the impression that it has been moved to care, manufactured for an institution that needs that impression and has not produced it honestly, and worth something only so long as the difference between the appearance and the reality goes undetected.

The dependence on concealment is the defining property of a counterfeit, and it separates the rented conscience from the real thing as sharply as a forged note is separated from a true one. A counterfeit must keep its books closed. Its worth survives only while the gap between what it appears to be and what it is remains hidden, which means the counterfeiter’s central task, after the forgery itself, is the prevention of inspection. The claque depended on the audience never being certain which hands were paid, and the rented conscience depends on no one comparing the institution’s professions to its conduct, on the workshop standing in for the audit that would expose it. A real conscience has the opposite relation to inspection. An institution that had in fact changed its conduct would have nothing to fear from an audit and every reason to welcome one, since the audit would confirm the change and convert the appearance of virtue into the demonstration of it. The test is therefore simple and revealing. Offer an institution two ways to establish that it cares about discrimination, a workshop that produces a certificate and an audit that measures whether its hiring and promotion have changed, and the institution’s choice tells you which kind of conscience it has. The one that reaches for the certificate and declines the measurement has a rented conscience, a thing that cannot survive the daylight of inspection, and the reaching for the certificate over the measurement is the surest sign that the appearance is all there is. That industry exists to supply the certificate, and its existence, alongside the relative scarcity of the audit, is evidence of how many institutions prefer the appearance they can buy to the

reality they would have to build.

The diversity training is the claque of corporate virtue. An institution runs its workforce through the workshop and acquires thereby the appearance of conscience, a documented demonstration that it takes discrimination seriously, and the appearance is the entire product, detached from any change in the institution's conduct exactly as the claque's acclaim was detached from any genuine enthusiasm in the hall. The previous chapter described the certificate the institution issues itself, the record that the training occurred, and named it a false certificate of conscience. This chapter can now say what the certificate is for. It is the receipt for a rented conscience, the proof that the institution has paid for the appearance of caring, producible on demand for any audience that wants to see it. The training manufactures, reliably and on schedule, the impression of an institution that has been moved by the wrong of discrimination to act against it, and it manufactures that impression for institutions that have not, in any measurable way, been so moved, since the conduct that would show the caring real, the hiring and the promotion and the redistribution of opportunity, remains exactly where it was. Rented, worn for the performance and returned when the performance ends, the conscience is worth something only to those who do not look behind it, like the claque's acclaim.

The claque is the rented conscience's ancestor in form, the manufacture of a paid response that passes for a felt one, and there is a second ancestor that supplies the rest of the resemblance, the medieval sale of indulgences. An indulgence was a document, purchased from the Church, that remitted the penalty owed for sin, and at the height of the trade it could be bought for a fee by a person who wished to discharge a moral debt without the harder labor of penance and amendment. The seller offered absolution as a product, the buyer paid for the appearance of a settled account with heaven, and the transaction left the buyer's conduct exactly where it had been, the sin unrepaired and the disposition unchanged, the only alteration a piece of paper certifying that the matter was closed. The parallel to the diversity training is close enough to be uncomfortable. An institution buys a workshop the way a sinner bought an indulgence, paying a fee to a seller of absolution for a document that discharges a moral obligation without requiring the change the obligation was supposed to compel. The workshop is the modern indulgence, the certificate of completion the modern letter of pardon, and the logic is identical across the five centuries that separate them, the substitution of a purchase for a penance, the conversion of a moral demand into a market transaction that leaves the demand unmet. The sale of indulgences provoked, in its time, the most consequential objection in the history of the Western church, when a monk argued that absolution could not be bought because the thing it purported to sell, a right relation to the good, was not the kind of thing a payment could produce. That objection applies without modification to the rented conscience. A right relation to the cause of justice cannot be purchased by the session, because the thing being sold, an institution committed in

earnest to ending discrimination, is not a thing a workshop can manufacture, and the certificate that says otherwise is selling what the indulgence sold, the false assurance that a debt has been paid when only a fee has been. The diversity industry is the indulgence trade reborn in the corporate age, and it stands open to the same charge, that it takes money for absolution it has no power to grant and leaves its buyers worse off for the false peace it sells them.

The purest form of the rented conscience costs even less than a workshop. It is the public statement of values, the diversity mission posted to the website, the annual report on inclusion, the corporate pledge issued after a season of unrest, each of them a document that performs conscience for an audience while committing the institution to nothing it can be held to. A statement of values is the indulgence stripped of even the pretense of a procedure, pure appearance at near-zero cost, a paragraph of virtue that changes a company's conduct not at all and its image considerably. The training at least gathers people in a room for an afternoon. The statement does not trouble anyone for even that long, and it buys the same thing, the look of an institution that cares, produced for the public and the press and the institution's own sense of itself, and worth, like every other rental, exactly nothing once a reader thinks to ask what behind the words has changed. These statements arrive in floods after a season of public reckoning, every institution issuing its pledge within days, the uniformity of the wording its own confession, since a sentiment arrived at in earnest would not emerge identical from a thousand legal departments at once. A flood of identical pledges is a claque of conscience, a manufactured chorus of institutions all professing the same feeling on cue, and the public that reads them is the audience the chorus was assembled to move.

A rented conscience, like a claque, is performed for an audience, and the diversity industry performs for several at once. There is the workforce, to whom the training says that the institution cares, reassuring the employees who might otherwise doubt it. The public forms a second audience, to whom the documented program says the same thing at larger scale, protecting the institution's reputation in a market that rewards the appearance of virtue. A court forms a third, to whom the training is offered as the evidence of diligence the previous chapter described, the legal shield that turns a workshop into a defense. And there is the institution itself, the strangest and most important member of the audience, to whom the training says that it is good, allowing the organization to reassure itself that it has met its obligations and may rest. Each of these audiences is sold the same counterfeit, the appearance of a conscience the institution has purchased rather than earned, and the performance succeeds with each of them to the exact degree that none looks behind the appearance to the unchanged conduct it conceals. The most consequential deception is the last, the institution's deception of itself, because an institution persuaded of its own virtue stops trying to acquire it.

The institution's deception of itself is where the rented conscience does its quietest and worst work. Its other audiences are external and can in principle see through the performance, but the institution as its own audience has no such distance. It wrote the script and watched the performance knowing it was staged, and the performance works on it anyway, because an organization, like a person, would rather believe well of itself than badly, and a documented program is exactly the evidence a willing believer needs. This is the self-certification the book traced in the individual, raised to the level of the institution and given a paper trail: the corporation points to a thing it did, the thing it did stands in for the structural change it declined, and the institution closes its own case, losing the discontent that might have driven it to improve.

The rented conscience belongs to a larger family, the manufacture of a counterfeit public that runs from Nero's hired soldiers to the modern bot farm, the plaza filled with paid demonstrators, the campaign of letters written to order, the network of automated accounts simulating a groundswell no living person feels. That descent is the subject of a separate book of mine, *The Rented Crowd*, and what matters here is only that the corporate performance of conscience works by the same logic as the rest of the family, an appearance produced well enough to pass for the reality and to do the reality's work until someone troubles to check.

Here the rented conscience parts company with the claque, and parts from it in the direction of greater harm. A claque's counterfeit applause is empty, a thing of no value once detected, but it does not ordinarily make the performance worse. The rented conscience does something the claque does not, because of a feature of human moral psychology that has been studied for a quarter-century under the name of moral licensing. That finding, established in a series of studies beginning with the work of Benoit Monin and Dale Miller at the start of this century, is that establishing one's credentials as a good person can free one to behave worse, that a record of virtue, once secured, licenses the transgressions it would seem to rule out. The foundational paper concerned prejudice precisely, and it carried a title about moral credentials and the expression of prejudice, because the effect showed up first and most clearly in exactly this domain. People who had been given a chance to establish that they were not prejudiced, by endorsing a minority candidate or affirming an egalitarian view, became more willing afterward to make decisions that favored their own group, the credential of tolerance freeing them to act in ways that would otherwise have looked intolerant. The effect has two mechanisms that the literature distinguishes, a credits model in which good deeds accumulate like a balance to be spent on later lapses, and a credentials model in which the prior good deed changes how the later bad one is read, so that an ambiguous act of bias is construed as something innocent because the actor has already proved he is not biased. That effect is real in the sense the argument needs, established well enough to name as a mechanism, though it carries the same cautions the next part

of this book will bring to its whole field. The replication era has been hard on social psychology, and the licensing literature has not been spared, with a pooled effect that is modest and some particular versions that have failed to repeat, so a fair reading treats it as a credible tendency rather than an iron law. The rented conscience does not need licensing to be a landslide, only that a secured reputation for virtue can, on the margin and with fair regularity, loosen the restraint it appears to certify, and on that modest claim the evidence is steady enough to build.

Of the two, the credentials model is the more dangerous here, because it operates where modern workplace bias operates, in the ambiguous act rather than the open slur, the candidate passed over for a reason that can always be given an innocent name, the review marked a little lower. A prior credential of tolerance is exactly what lets such an act be read as innocent, and the research even finds people assembling the credential in advance of the transgression it will excuse, the individual version of an institution buying a training before the complaint it expects.

Apply the finding to the workshop and the result is damning. The training issues every attendee the exact thing the licensing research identifies as the source of the effect, a documented credential of non-prejudice, in the one domain where such credentials are known to loosen restraint. It works no transformation of tolerant people into bigots; it hands everyone a document attesting to their tolerance, and such a document, by the evidence, can make the bias it names somewhat easier to express.

The combination of findings produces a conclusion sharper than either finding alone. From the evidence reviewed in the previous chapter, the training does not reliably reduce the bias it targets. From the licensing research, the credential the training confers can make the expression of that bias somewhat easier. Put the two together and the diversity workshop occupies the worst available position, a practice that fails to deliver the benefit it promises while delivering a harm it never mentions, the issuance to every participant of a certificate that the science associates with looser, not tighter, restraint on prejudice. The defender of the training will object that the effect is modest, and the objection is fair as far as it goes, since the licensing tilt is real but small and no single workshop turns a fair-minded person into a bigot. That modesty, though, has to be weighed against the modesty of the training’s benefit, which the evidence puts at or near zero, and a practice whose benefit is approximately nothing and whose harm is small but real does not come out ahead by being only slightly harmful. The arithmetic is unforgiving. An intervention that does not work and may do a little damage is worse than no intervention at all, which fails just as surely while doing no damage and costing nothing, and the eight billion dollars a year buys, on this accounting, a small net harm dressed as a large public good. The rented conscience does more than fail as a purchase; it leaves the buyer worse off than the purchase was meant to make him, which is the one thing a conscience is never supposed to do.

The same logic operates at the level of the organization. A company persuaded by its own training that it has addressed discrimination has acquired an institutional credential it can spend, and it spends it by declining the harder things, treating the workshop as the discharge of an obligation rather than the beginning of one, which is the complacency the previous chapter named, now given its mechanism.

A further finding from the same body of research explains why the rental is preferred to the real thing even by people of goodwill. The psychologists who study moral behavior have documented what they call the rejection of moral rebels, the tendency of a group to resent the member who makes the harder, more virtuous choice, because his choice implies a judgment on the easier choices the others have made. The person who does the right thing in earnest, and who thereby shows up the people content with the appearance of it, earns dislike rather than thanks, experienced as a reproach rather than a model. This dynamic favors the counterfeit at every turn. An institution that buys the rented conscience joins a comfortable majority of institutions that have done the same, and threatens no one, while an institution that does the real and difficult work of structural change implicitly indicts its peers and invites their resentment. The rented conscience is socially safe in a way the real one is not, since the rental asks nothing of anyone and flatters everyone who buys it, whereas the genuine article exposes the gap between what most institutions do and what they could do. The market thus rewards the counterfeit twice, once because it is cheaper and again because it is more comfortable, and the institution that might have chosen the real thing finds the social weather running against it, the rented conscience approved on every side and the real one quietly resented as a rebuke.

Hold the two harms of the rented conscience together, because they compound rather than merely add. The first harm, described in the previous chapter, is displacement, the way the workshop occupies the place where real change would have gone, absorbing the budget and the attention and the sense of obligation that the structural work would have required. The second harm, described in this one, is licensing, the way the credential the workshop confers loosens the restraint on the conduct it names. These two do not sit side by side as separate costs. They feed each other. The displacement ensures that the real work is not done, and the licensing ensures that the people who might have done it feel no pressure to, because the credential has told them, individually and collectively, that the matter is handled. An institution thus arrives, by way of the rented conscience, at a stable and self-sustaining failure, having spent money to acquire a document that both prevents the cure and suppresses the symptoms that would have signaled the disease was still present. The genius of the arrangement, if a thing this destructive can be called ingenious, is that it leaves everyone satisfied, the institution persuaded it has acted, the employees relieved of the discomfort of being suspected, the lawyers supplied with their defense, and the discrimination itself

undisturbed beneath the whole contented structure, protected now by the ritual that was sold as its remedy.

Everything about the rented conscience conspires to make it the rational choice, which is why naming it as a counterfeit does not, by itself, end its reign. Set the incentives side by side. The rental is cheaper than the real thing, a workshop against a restructuring, and it is faster, a quarter against years. It is more defensible, a documented program against the uncertain promise of a changed culture, and socially safer, the comfortable choice of the majority against the resented choice of the rebel. Most flattering of all, it offers a certificate of virtue in place of the painful admission, implicit in any serious reform, that the institution had been failing. Every axis on which an institution might weigh the two options favors the rental, and the single axis on which the real thing wins, its actual reduction of discrimination, is the one axis no one in the transaction is required to measure and few choose to. A choice structured this way will come out the same way nearly every time, the institution selecting the counterfeit through the ordinary operation of incentives that all point one direction rather than through any corruption or indifference. This is the deepest reason the diversity industry will outlast every exposure of its emptiness. The exposures address the one consideration that was never driving the purchase, the question of whether the training works, while leaving untouched the five considerations that were, and a product chosen for reasons that have nothing to do with its effectiveness is not threatened by the news that it is ineffective.

The strongest case for the training should be stated, since the book owes it that. A defender would argue that the workshop raises awareness, that even an imperfect session puts the question of discrimination in front of people who might otherwise never consider it, that it signals an institution’s values in a way that shapes the culture over time, and that a motivated participant can take from it something real. The case is effective in the sense that some of it is true, since a workshop can convey information a person did not have, and an institution’s public commitment to a value can, under the right conditions, influence the behavior of those who already share it. The case is not effective as a defense of the practice, for reasons the book has assembled across two chapters. Awareness is not the bottleneck, since the evidence finds that information about bias does not reliably change behavior and that the people most in need of changing are the least moved by being told. The signaling of values cuts both ways, since the same signal that might inspire the committed licenses the complacent, and the licensing research suggests the second effect is at least as real as the first. And the benefit to the motivated participant, even if granted, is purchased at the price of the credential issued to everyone else, the certificate that the science associates with looser restraint on prejudice. Weigh the modest and conditional good the training might do for the already-willing against the displacement of real change and the licensing of bias

across the whole workforce, and the recommendation follows without strain. The honest course is to stop renting the conscience and to build one, by the means the next chapters describe, which cost more and flatter less and have the single advantage the rental lacks, that they work.

Step back and the rented conscience reveals itself as the industrial form of everything the book has been tracing. The genre began with a single teacher manufacturing an experience of cruelty and trusting the feeling to teach, and the industry ends with institutions manufacturing an experience of virtue and trusting the appearance to suffice, the same three failures running through both. All three claims reappear in industrial form. The manufactured conscience is the oppressor's counterfeit of the real one, the form of caring worn by those who do not care enough to change; the credential meant to reduce bias instead licenses it; and the marginalized in whose name the conscience is purchased are nowhere in the transaction that claims to serve them, present only as the occasion for a performance staged for other audiences. A rented conscience is what the false certificate becomes when a market gets hold of it, the private lie a person tells himself after a workshop scaled into a public lie an institution tells the world and itself, sold by the session and filed as proof.

None of this impugns the wish to have a conscience, which is the one thing in the whole transaction worth honoring. An institution that wants to be free of discrimination wants something good, and the tragedy of the rented conscience is that it preys on exactly that wish, offering the sincere institution a way to feel its wish satisfied without the labor of satisfying it. The answer is to make the conscience real rather than to abandon it, to do the structural work that an actual conscience would demand, the work the final chapters of this book describe, which changes an institution's conduct rather than its records. A real conscience cannot be rented, only built, and it is built in the slow alteration of how people are hired and judged and advanced, not in the afternoon a workshop fills. The diversity industry sells the rental because the rental is what institutions will buy, a conscience by the session, worn for the audience and licensed to excuse the inaction behind it, and the industry will go on selling it for as long as the appearance of virtue is cheaper than the substance and serves the same purposes in the same markets. Whether the substance can be measured at all, and what the measurements show when they are taken, is the question the next part of this book confronts, turning from the industry that sells the method to the science that has tried, for decades, to find out whether the method works.

PART FIVE

The Science

CHAPTER TWELVE

What the Evidence Says

Everything to this point has been an argument from logic, from history, and from the shape of the thing. The exercise inverts the relation it claims to teach. Its cure manufactures the disease it treats. The ritual needs a sealed room and an absent subject, and dies on contact with the real one. The industry sells a conscience by the session, a counterfeit that licenses the wrong it names. These are claims a reader can test against reason and against the record, and a reader who has come this far may have tested them and found them sound, and may still want the one thing the argument has so far withheld, which is proof of the ordinary kind, the kind that comes from counting.

This chapter supplies it, and the proof has a shape worth naming before the numbers arrive. The case rests on a convergence. The same method, tried in field after unrelated field against problems that have nothing to do with one another, measured each time with the full apparatus of the randomized trial and the meta-analysis, fails in all of them and fails in the same way. That is the argument, and its order is deliberate. The one place a reader would expect the proof, the diversity workshop that is this book’s nominal subject, is the place the evidence runs thinnest, where the research is poor and an honest critic must concede the firm evidence is scarce. So the book does not rest the case there. It rests on the wider pattern, harder to wave away than any single field could be. Precision matters in drawing that pattern, because the category can be stretched too wide and a fair critic will stretch it to break it. The thing that converges is the coerced attempt to change a person by staging or inducing a feeling in him, imposed by an authority or a category rather than chosen, with the engineered experience standing in for the real engagement it imitates, and it appears in the two tiers Chapter Three set out, the full five-part exercise at the center and the looser ring around it that runs on the same engine while missing a part or two, the fear tour without its reversal, the drug lesson without its caste. A hard practice moves a person and a well-made memoir moves him, and neither belongs to either tier, because each is entered freely and pays off in a real task or a real encounter, where the staged exercise only assigns a feeling and calls the feeling the lesson. That shared structure, coercion in place of choice and a proxy in place of the

thing, is the natural kind under all the different surfaces, and it is what has failed in field after field where it has been measured, whether by the harm of its harshest versions or the futility of its mildest. A method that fails once has a flaw. A method that fails across crime and drugs and poverty and prejudice, in studies run by different people for different ends over fifty years, has a signature, and the signature is the argument of this chapter. The convergence is the proof. No single program carries the weight. The pattern across all of them does.

One distinction has to be drawn inside that convergence before the cases arrive, because the programs do not all fail with equal severity, and a reader is owed the precision that the word failure can blur. Two claims run through this chapter, supported to different degrees. The stronger and narrower one is that an engineered aversive experience can actively harm, can leave the person worse than no program at all would have, and it is proven cleanly by the prison-tour studies, where the frightened boys went on to offend at higher rates than the untouched controls, and it holds for the mandatory workshop that breeds the resentment it set out to dissolve. The broader and looser one is that interventions working by induced feeling tend not to change behavior for long, which is the lesson of the drug curriculum that bought no lasting abstinence and the poverty simulation whose softened attitudes are real but short-lived. Keeping the two apart matters, because their force differs. A program that merely fails to last has wasted money and time. A program that backfires has done damage, and only some of the cases reach that second bar. The word iatrogenic, the cure that produces the disease, belongs to the harsh end of the family, the staged fright and the forced confrontation, while the gentler members earn the milder verdict of futility. The dramatic method is convicted on both counts at once, of frequent backfire in its cruelest forms and of routine waste in its mildest, and the reader is owed the knowledge of which charge each case proves, since the difference between a treatment that does nothing and a treatment that wounds is the whole difference the harsher word is meant to mark.

There is a standard for judging whether an intervention works, and it is the standard medicine uses. A treatment is tested against a control, the assignment made at random so that the two groups differ only in whether they received the treatment, and the results are pooled across many such trials in a systematic review that weighs the strong studies and discounts the weak. A drug that failed this test would never reach a pharmacy, and a drug that made the disease worse would end a career. Social interventions are rarely held to it. They are adopted because they seem sensible, defended because they feel powerful, and sustained because no one looks. The dramatic method has mostly escaped the standard for the same reason it spreads, because the experience is so convincing that measurement seems beside the point. Where the standard has been applied, by the Campbell Collaboration, the social-policy counterpart of the Cochrane reviews that brought the discipline of the drug trial to medicine, the dramatic method has not

survived it. A practice that would be pulled from the market as a drug has been kept in the schools and the prisons and the corporate calendar as a ritual, and the chapters of evidence that follow are the record of what happened when someone finally ran the trial.

Begin with the purest case, because it is also the most disturbing, and because it removes at the outset any comfortable assumption that a powerful experience is at worst a harmless one. The case is Scared Straight. The program took juvenile delinquents and children thought to be heading toward delinquency and brought them inside a real prison to be confronted by real inmates, men serving long sentences who shouted at them, described the violence of the place in detail, and painted the future that waited for them if they stayed on their course. It is the dramatic method in its most concentrated form, a staged encounter engineered to terrify, on the theory that a young person frightened badly enough will turn away from the path that leads to the cell. The theory is intuitive, the experience is overwhelming, and the program became famous on the strength of a 1978 documentary about a prison program in New Jersey that won an Academy Award and persuaded a generation of officials that the thing worked. It spread across the United States and was exported abroad. It felt like the answer.

The documentary deserves a closer look, because it is the mechanism of the disaster in miniature. A film crew followed a group of hardened convicts as they berated and frightened a busload of teenagers, and the footage was overwhelming, the inmates towering and furious, the children visibly shaking, the change seemingly written on their faces by the closing minutes. It was riveting television and it carried an unmistakable message, that this works, look at it working. The film won its Academy Award and its audience drew the conclusion the images compelled, and the program spread on the strength of how convincing the spectacle had been. The persuasive power of the documentary and the failure of the program were the same fact seen from two sides. What made the film so moving, the rawness of the fear and the apparent break in the children, was the felt impact that everyone mistook for proof, and the mistake was carved into celluloid and broadcast to a nation that adopted a harmful program because a camera had made its cruelty look like medicine.

The program was measured, and the measurement is among the most important in the history of social policy, because it did not merely fail to find a benefit. The Campbell Collaboration, an international body that produces systematic reviews of social interventions to the standard medicine uses for drugs, examined the controlled studies of Scared Straight and programs like it. Anthony Petrosino and his colleagues found nine trials that had randomly or nearly randomly assigned youths to attend the program or to a control group that did not, a total of 946 young people between the ages of fourteen and twenty, across eight American states. The combined result was not a small positive effect, and not a null effect, and not a wash. Those who went through the program committed more crime afterward than the ones who did not. The seven studies

that tracked reoffending showed that the experience significantly raised the odds of offending. In the preliminary version of the analysis, the increase ran between one and twenty-eight percent over a no-treatment control. The reviewers stated the conclusion plainly, that programs of this kind do more harm than doing nothing at all, and that they could not recommend the approach as a way to prevent crime.

Sit with the structure of that finding, because it is the hinge of the chapter. An institution wanted to reduce crime among at-risk youth. It chose the most dramatic intervention available, an intervention whose force as an experience no one disputes, and it applied that intervention to thousands of children with the certainty that an experience so strong could only help. The intervention made the problem worse. It did not waste the children’s time in a neutral way. It increased the behavior it was built to prevent, and it did so reliably enough that the most rigorous review in the field concluded the children would have been better off left alone. The title Petrosino and two colleagues gave to an earlier paper on the subject states the lesson the whole of this book is trying to establish, that well-meaning programs can have harmful effects. The drama is not inert. A staged experience powerful enough to move a person is powerful enough to move him in the wrong direction, and the strength of the experience is no guarantee of the direction of the result. That is the first thing the evidence proves, and it is fatal to the assumption on which every dramatic method rests, the assumption that an intense experience in a good cause is at worst a waste and at best a cure, and never itself the injury.

Why a fright should produce more crime rather than less is a question the research answers only in part, and the partial answers matter, because they recur in every other field where the method has been tried. One mechanism is reactance, the human tendency to push back against a force that tries to control us, so that a heavy-handed attempt to scare a young person into compliance produces defiance instead, a hardening of the identity the program meant to dissolve. Another is labeling, the way an official ceremony that treats a child as a future convict can confirm him in that role, supplying the self-image that the behavior then grows to fit. A third is the simple glamour of the forbidden, the possibility that a vivid tour of the criminal world reads to some adolescents less as a warning than as an initiation. The mechanisms are debated. The result is not. That strong experience backfired, and it backfired in a direction the planners never imagined, because they had stopped at the strength of the experience and never asked what the experience would do.

The most telling part of the Scared Straight story came after the evidence was in. A method shown by the most rigorous review in its field to increase the crime it targeted might be expected to vanish, quietly retired by the institutions that had championed it. It did the opposite. Decades after the trials, the format returned as popular television, a new series built on the same premise, the same busloads of frightened teenagers and the

same shouting inmates, broadcast as entertainment and by implication as instruction. The federal office responsible for juvenile justice felt compelled to object in public, warning that the approach did not work and could do harm, a government body reduced to pleading with a television audience not to believe what the cameras seemed to show. The spectacle had outlived the science because the spectacle was the point. A busload of weeping children is a powerful thing to watch, and the power of the watching was always going to beat a table of reoffending rates, which is the whole disease of the dramatic method stated as a programming decision.

Carry that lesson to the second case, which trades the prison for the classroom and fear for resistance, and which reached almost every child in America. Drug Abuse Resistance Education, known by its acronym, was founded in 1983 by the chief of the Los Angeles police, Daryl Gates, in partnership with the Los Angeles school district, and it sent uniformed police officers into elementary classrooms to teach children to refuse drugs through a curriculum of role-play, pledges, and the rehearsal of resistance to peer pressure. It was the experiential method in a gentler register, the lesson delivered as a felt rehearsal rather than a lecture to be forgotten, a child standing up in front of the class to practice saying no, a uniformed authority making the danger vivid and personal. At its height the program ran in roughly half of American school districts, and by some counts a large majority, with a cartoon lion for a mascot and a brand recognition most public health campaigns can only envy. It felt, again, like the answer, and the feeling was widely shared and deeply held.

It was measured as thoroughly as any program of its kind, and the measurements form a wall. In 1994 a team led by Susan Ennett published a meta-analysis in the *American Journal of Public Health* that pooled eight methodologically rigorous evaluations of the program. The effect on drug-use behavior ranged from nothing to almost nothing, with a weighted average that rounded to negligible, and the effect was reliably smaller than the effect of the interactive programs the children might have received instead. Longer studies found the same. A five-year follow-up found no significant effect on drug use, with the small early gains in attitude fading as the children aged into the years when the drugs were on offer. The early evaluations had already found something worse than nothing in places, a slight increase in use among some who had taken the program, the boomerang that fear-based prevention keeps producing. The federal government drew the obvious conclusion. In 1998 the Department of Education required that the drug-prevention programs it funded be supported by evidence, and the most famous drug-prevention program in the country did not qualify, and by 2001 had lost its eligibility, after which it adopted a new and different curriculum built with prevention specialists and released over the following decade.

Why a program built to suppress drug use should sometimes raise it repeats the puzzle of the prison, and the likely answer rhymes with that one. A curriculum that spends hours warning ten-year-olds about the drugs all around them, and rehearsing the refusals they will need when the pressure comes, delivers a second message under the first, that drug use is common, expected, a normal feature of the adolescence ahead. A child told repeatedly that everyone will offer him drugs learns that drugs are everywhere and everyone is using them, and the lesson meant to build resistance can plant the expectation instead. The vividness compounds the error. A police officer describing the effects of a drug in detail, however grim the framing, has introduced the drug, named it, made it concrete and a little glamorous, and for some fraction of children the introduction outweighs the warning. The mechanism is the prison's mechanism wearing a friendlier face, the staged lesson supplying the idea it meant to forbid, and the result is the boomerang the evaluations kept finding, a prevention program that in places functioned as advertising.

What belongs most precisely to this book in the DARE story is less the failure, which it shares with Scared Straight, than the defense its champions offered once the failure was known. Pressed on the absence of evidence, a program official is recorded in a 1998 report to a federal justice institute making the argument that the program's strongest numbers were the ones that never showed up, the drug use prevented and therefore impossible to count, and that the strong public support the program enjoyed was a better measure of its worth than the scientific studies. Read that defense slowly. It is the whole pathology of the dramatic method stated by one of its own. The evidence says the thing does not work, the feeling says it does, and when the two collide, the institution sides with the feeling and recasts the absence of measurable benefit as a benefit too subtle to measure. The program survived for years after the science had condemned it, kept alive less by any showing that it saved a single child than by the satisfaction it gave everyone involved, the officials who funded it, the officers who taught it, the parents who watched their children pledge, and a public that wanted to believe a vivid lesson could inoculate a child against a hard world. Felt impact had replaced demonstrated impact as the standard of judgment, and the substitution kept a useless program alive long after its uselessness was a matter of published record.

Two unrelated fields, crime and drugs, have now produced the same verdict on the same method, and a reader is entitled to ask whether the diversity workshop, this book's actual subject, behaves any differently when it is measured with equal care. It does not, and the most rigorous look at it comes from a pair of sociologists who set out to find what the corporate apparatus of diversity programming had accomplished. Frank Dobbin and Alexandra Kalev assembled three decades of data from more than eight hundred American firms, tracking what happened to the representation of women and minorities in management after companies adopted the standard tools, the diversity

training chief among them. The training, they found, did not deliver. In the firms that made it mandatory, the form nearly all large firms use, the numbers that were supposed to rise often fell. The share of Black men in management dropped. The share of women in management dropped. Similar declines appeared for other groups the programs were meant to help. The training that was supposed to open the institution to people it had excluded was followed, on average, by an institution slightly more closed to them than before. One limit on that finding belongs here, since the book holds others to the same standard. This is firm-level data tracked over time, an association rather than a controlled experiment, so the decline that follows the training cannot be charged to it with the confidence a randomized trial would give, and the honest version of the claim is the narrower one, that the most thorough study of the question found the promised gains missing and, in many firms, a slide where the gains were meant to be.

The reason Dobbin and Kalev give for the backfire is the reason the prison and the classroom already supplied, and the recurrence is the point. Mandatory training that tells managers their judgments are biased and orders them to correct themselves produces reactance, the same push-back that hardened the delinquents and the schoolchildren. A manager compelled to attend a session that casts him as a bigot in need of correction does not emerge reformed. He emerges resentful, his defenses raised, and the bias the session was meant to lower can rise instead, activated by the resentment the compulsion produced. The conditions that make training most likely to backfire are the conditions under which most of it is delivered, made compulsory, justified by the threat of lawsuits, and aimed at managers as a class to be fixed. Roughly three-quarters of firms train under one or more of those conditions. The federal commission charged with enforcing the civil rights laws reviewed the evidence in 2016 and reported that it could find no demonstration that the standard diversity training reduces bias or stops harassment, the same official verdict the Department of Education had delivered on DARE sixteen years earlier and the Campbell reviewers on Scared Straight before that.

The same study that condemned the training pointed to the things that worked, and the contrast sharpens the verdict. When Dobbin and Kalev looked for the practices that moved the numbers, they did not find them among the lectures and the modules. They found them in structures, in targeted recruitment that widened the pool, in mentoring that attached a senior figure to a junior one across lines of difference, in the diversity task forces that gave managers ownership of the outcome and a reason to notice it. The tools that worked shared a feature the training lacked. They engaged people as agents rather than treating them as defective, they brought groups into sustained contact rather than lecturing them about their biases, and they made someone accountable for a result rather than for attendance. Voluntary training, offered to people who chose it rather than forced on people who resented it, did better than the mandatory kind for the same reason. The finding points two ways at once, that the dramatic confrontation hurts and

the patient structural work helps, which is the opposite of what the industry sells, and the later chapters return to the structures that survive the evidence.

There is a milder version of the failure that deserves its own treatment, because it accounts for the experiences that keep the method alive in the minds of people who have been through it. Not every dramatic exercise backfires. Many of them work, in the narrow and temporary sense that they move the people who undergo them, and the trouble is only that the movement does not last and does not reach behavior. The poverty simulation is the clearest instance, the staged exercise examined in an earlier chapter in which the comfortable spend an afternoon impersonating the poor, failing to pay simulated rent and waiting in simulated lines. The studies of it find what an honest observer would expect. Attitudes shift in the hours after the exercise. Participants report more sympathy for the poor, score higher on measures of concern, and leave moved. And then the effect decays. Follow-ups find the gains fading toward the baseline, and some find a ceiling effect, the participants having arrived already sympathetic, with the exercise adding a temporary intensity to a feeling that was there before and changing nothing they will do once the feeling cools. The empathy is real. The empathy is also a mood, and a mood is not a reform, and an afternoon that produces a strong feeling and no lasting change is a particular kind of failure, the kind that feels like a success to everyone in the room and thereby escapes the scrutiny a visible failure would attract.

Step back from the individual programs to the state of the science as the scientists themselves describe it, because the discipline that studies prejudice reduction has rendered a verdict on its own field that the industry built on that field does not advertise. Elizabeth Levy Paluck and Donald Green, working at Harvard and Yale, surveyed the whole research literature on reducing prejudice in 2009, hundreds of studies, and concluded that only a small fraction were rigorous enough to support any confident claim, and that the causal effects of the most widespread interventions, diversity training and media campaigns among them, simply remained unknown. That is a remarkable sentence to find at the center of a multibillion-dollar enterprise. The people who study the question hardest reported, after examining everything, that the field did not know whether its most popular tools did anything at all. A dozen years later the same scholars and their colleagues looked again, this time at 418 experiments reported across more than three hundred manuscripts from 2007 to 2019, and the picture had grown more rigorous without growing more reassuring. Three-quarters of the studies tested light-touch interventions whose lasting effects were unclear. The most common approach asked people to imagine their way into another’s mind. The optimistic results that such studies reported came shadowed by signs of publication bias, the systematic tendency of encouraging findings to reach print while discouraging ones stay in the drawer, which inflates the apparent effect of everything the field believes. And the enthusiasm for interventions aimed at unconscious bias, the reviewers noted, was not

matched by evidence, since most of the recent work did not even measure the implicit attitudes those interventions claim to change.

A pattern runs through this whole literature, and once seen it is hard to unsee. The weaker the study, the larger the effect; the stronger the study, the smaller it shrinks. A poorly controlled evaluation with a handful of subjects and a questionnaire handed out minutes after the session finds dramatic improvement. A large randomized trial that follows people for a year and measures what they do finds almost nothing. This is the fingerprint of an effect that is not real, or barely real, an effect that lives in the noise and the hope of small studies and evaporates under rigorous light. Combine it with the publication bias the reviewers flagged, the tendency of encouraging results to reach print while null results die in a drawer, and the apparent record of the field tilts upward by construction, a literature that looks more promising than the studies inside it warrant. The methods that feel the most powerful in the room are the ones whose measured effects collapse the hardest when the room is replaced by a controlled trial, which is what one would expect of a method that confuses the intensity of an experience with the persistence of a change.

The fashion of the present moment is the unconscious-bias training, and it fits the pattern exactly. Its premise is appealing, that discrimination now lives below awareness, in automatic associations a person cannot feel and would disavow, and that a session surfacing those associations can loosen their grip. The trouble is that the evidence for the second half has never matched the enthusiasm for the first. The instrument used to reveal implicit bias predicts individual behavior weakly at best, the reviewers of that question have found, and the interventions that claim to reduce implicit bias tend to produce shifts that are small, that fade within days, and that have not been shown to change what people do. Paluck and her colleagues noted the gap with some understatement, that the appetite for these programs in the world is not justified by the research, and that most of the recent studies did not even measure the implicit attitudes the programs are sold to fix. The unconscious-bias workshop is the newest dramatic method, a staged encounter with one’s own hidden sins, and it is failing the way the older ones failed, by feeling revelatory and changing nothing that can be counted.

Assemble the verdicts and the shape of the thing is unmistakable. A fear-based program to prevent crime increased crime. A resistance-based program to prevent drug use did nothing and sometimes increased use. The training built to raise the representation of women and minorities was followed by a decline in it. An empathy-based exercise to build lasting concern for the poor produced a feeling that faded by the following week. And the science that studies the whole enterprise of changing hearts through designed experience reports that it cannot, after decades of trying, confirm that the popular methods work, and suspects that the encouraging results it does have are inflated by the bias of its own publication. Five failures of five unrelated

programs would be a coincidence worth shrugging at. Five appearances of one failure is a diagnosis, and the diagnosis is plain. The method is the staged experience built to change a person by making him feel, and in the cases examined here it fails in the same way, by mistaking the strength of the feeling for evidence of the effect, by provoking the resistance that strong pressure provokes, and by producing changes that live as long as a mood and die as fast as one.

The convergence does more than condemn a list of programs. It confirms from the outside, with numbers, what the book argued from the inside. The inversion shows up as backfire, the staged degradation producing more of the harm it opposed and the confrontation breeding the resentment that deepens the bias. Scared Straight is the manufactured disease made literal, a treatment that produced the illness it was prescribed against. And the cost of the absent subject shows up in who pays when the stand-in is real, the delinquent who reoffends, the child who tries the drug, the employee passed over after the training that was supposed to protect her. Argument and evidence reach the same verdict from opposite directions, which is the strongest position a case can occupy, the logic predicting what the measurement then finds.

That leaves the question the DARE official answered without meaning to, the question of why methods this thoroughly discredited persist. The answer is the one the previous two chapters supplied, and the evidence now confirms it from the outside. The programs survive because they were never selected for their effect on the problem, only for their effect on the institution. Scared Straight gave officials a way to be seen acting decisively against juvenile crime. DARE gave a city a way to put a caring officer in every classroom and a photograph in every paper. The diversity workshop gives a corporation a documented act of concern producible in court and persuasive to its own conscience. Each of them satisfied a felt institutional need, and each was judged by whether it satisfied that need rather than by whether it moved the outcome it named, which is why published proof of failure changed so little for so long. A practice chosen for the comfort it provides will not be abandoned when it turns out to provide no cure, because the comfort was the function and the cure was the cover story. The numbers that never showed up, in the DARE official’s phrase, were never the point. The showing was the point.

A deeper question is why the showing persuades at all, why intelligent people trust a vivid afternoon over a controlled study, and the answer lies in how human judgment weighs evidence. A single concrete scene, charged with emotion and witnessed firsthand, lands with a force that a page of statistics cannot match. The mind remembers the shaking child and the shouting inmate; it does not remember a reoffending rate, and it does not feel one. Psychologists call this the pull of the vivid, the way an available and emotionally loaded instance crowds out the pallid abstraction of an average, and it is a permanent feature of how people reason rather than a failing of the careless. An

administrator who has watched a roomful of employees leave a workshop visibly moved has an experience that feels like knowledge, and the feeling is honest even when the knowledge is false. The dramatic method exploits this pull by design, since its whole theory is that the strong experience is what changes people, and the same strength that is supposed to do the changing is what fools the observer into believing the change occurred. The method is built to be convincing whether or not it is effective, and those are separate properties that the vividness of the thing makes almost impossible to tell apart.

This is why the defenders' strongest card, the personal testimony, settles nothing. Someone always stands up to say the program changed his life, that he was headed for prison or addiction until the day the inmates screamed at him or the officer warned him, and the testimony is sincere and moving and worthless as proof. It is worthless because it has no counterfactual, no way to know what the same person would have done without the program, and the whole point of the controlled trial is to supply the counterfactual the testimony lacks. It is worthless, too, because of who gets to testify. The people for whom the program backfired, the youths who left more likely to offend and then did, do not return to the reunion to describe their relapse, and the children whose drug use the lesson nudged upward do not write letters of thanks. Testimony is gathered from survivors and volunteers, a sample selected for the outcome it is meant to demonstrate, and a method can harvest glowing testimony while harming its average participant, which is what the dramatic programs do. The plural of anecdote is not data, and a thousand moving stories from the people a program helped tell you nothing about the people it hurt, who are not in the room and were never asked.

Who bears the cost is the same in each of these cases, and the pattern names the book's third claim in the language of harm. The institution that runs the dramatic program is not the party that pays when it fails. An official who funds Scared Straight feels resolute; the youth who reoffends pays. The city that puts an officer in the classroom feels caring; the child nudged toward the drug pays. A company that mandates the workshop feels responsible; the woman passed over by the manager the session antagonized pays. In each pairing the institution buys a feeling and the subject absorbs the consequence, and the subject is the person the ritual treated as a prop, the delinquent and the child and the excluded employee, present in the world but absent from the calculation that selected the program. The dramatic method distributes its benefits and its costs to different people, comfort to the institution that performs it and damage to the subject performed upon, and a method with that distribution will persist no matter how badly it fails the people it claims to serve, because the people it serves and the people it fails are not the same.

A failed program is not free, and its true price is the cure it displaces. Money and time are finite, and every hour a child sits through a police officer's drug lecture is an

hour not spent on the tutoring or the counseling or the after-school program that might have done some good. DARE ran for a generation in roughly half the school districts in the country, taught by sworn officers whose hours carried a real price, and the bill for all of it purchased a measured effect indistinguishable from zero. Every dollar a company spends on a mandatory workshop that lowers its diversity numbers is a dollar not spent on the targeted recruitment or the mentoring that raises them. The afternoon the comfortable spend role-playing poverty is an afternoon, and a budget, that could have gone to the rent assistance the actual poor need now. The dramatic method does its deepest damage here, in the help it crowds out, because a society convinced it has addressed a problem stops looking for the solution that works. A useless program is worse than idleness when it occupies the space and consumes the resources a useful one would need, and it buys, with that space and those resources, the false belief that the problem is being handled, which is the most expensive thing a failed program sells.

Underneath the institutional incentives sits a bias deeper than any of them, the human preference for action over inaction even when the action is worse. A leader who does nothing about a visible problem is blamed for indifference, while a leader who does something useless is credited with effort, and the asymmetry holds even after the something has been shown to fail. Scared Straight was chosen over leaving at-risk youths alone, and the leaving-alone turned out to be the safer course, yet no official ever lost a position for declining to traumatize children, while many would have been attacked for standing by as juvenile crime climbed. The same logic shields the useless workshop and the failed drug program. Inaction reads as neglect and action reads as care, regardless of which one helps, and a public that judges its institutions by visible effort rather than by measured result will reward the dramatic gesture over the quiet method every time. Economists call this a preference for commission over omission, and it warps policy wherever the cost of a program falls on someone other than the person who approves it. The institution collects its reward for the gesture in the currency it values, the appearance of concern, and pays for the failure in a currency it discounts, the welfare of the absent subject, so the ledger always favors acting badly over not acting at all. The finding that a program does more harm than nothing changes little against a force like that, because nothing was never a choice the institution could be seen to make, even when nothing was the choice that would have spared the people the program injured.

A fair statement of the other side is that not all of this is hopeless, and the honest critic must say so plainly. The research that buries the dramatic method also finds that some interventions help, and the line between them is instructive. Sustained contact between groups, of the kind that unfolds over weeks of shared work toward a common goal, has the best evidence behind it of anything in the field. Interactive programs that build real skills, rather than staging a single emotional event, outperformed DARE in the

same analyses that sank it. A workshop chosen freely by a willing participant behaves differently from one forced on a resentful one. The failure documented in this chapter is the failure of a specific form, the short, intense, staged confrontation designed to change a person by overwhelming him, and it would be dishonest to stretch the indictment into a claim that nothing can reduce prejudice or that education itself is futile. What is on trial is the dramatic method. The quieter methods, patient and structural and sustained, are a separate matter, and they are the subject of a later chapter, which is the constructive answer this critique owes and intends to pay.

None of this argues that the problems are not real or that the goals are not worth pursuing, and the distinction has to be held with both hands, because the failure of these methods is the favorite weapon of people who would be glad to see the goals abandoned along with the methods. Crime among the young is a real harm and worth preventing. Drug addiction wrecks real lives and is worth preventing. The exclusion of women and minorities from the institutions that distribute opportunity is a real injustice and worth ending. The poverty that the simulation trivializes is a real condition and worth relieving. Nothing in this chapter touches the worth of any of those ends. It touches only the means, and it delivers a hard and specific message about the means, that the dramatic experiential method does not achieve them and sometimes sets them back, and that the institutions devoted to those ends owe it to the people they claim to serve to stop spending on a method that fails and to spend instead on the methods that the same body of research, in its quieter findings, suggests can work. The intergroup contact that Paluck and Green found promising, the interactive skills-building that outperformed DARE in the same meta-analysis that buried it, the structural changes in hiring and promotion that Dobbin and Kalev found to move the numbers when the training did not, the booster sessions and sustained engagement that the drug researchers recommended in place of the one-time fright, all of these point toward a different way, and the chapters ahead take it up. What this chapter proves is narrower, and it is final within its bounds. The method on trial has been measured, in field after field, by people with every motive to find it working, and it does not work, and the children and employees and citizens subjected to it have been paying, in wasted effort and occasional real harm, for the institution’s need to be seen to care.

One question remains, and it reaches beneath the programs to the ideas that justified them. The dramatic method did not arise from nothing. It drew its authority from a handful of famous experiments that seemed to prove how readily ordinary people could be made to do extraordinary things, the studies of obedience and conformity and the corrupting power of roles that every educated person can name, the demonstrations that taught a generation the human being is a puppet of his situation and can be reshaped by staging the right one. Those experiments are the intellectual foundation under the exercises, and that foundation has been crumbling. The most celebrated of the

demonstrations have failed to replicate, or have been found on inspection of their own archives to have been staged in ways their fame never disclosed. The next chapter turns to them, because a method already condemned by the evidence of what it does is condemned a second time by the collapse of the evidence it came from, and the two condemnations together close the case.

CHAPTER THIRTEEN

The Collapse of the Dramatic Paradigm

The previous chapter condemned the dramatic method by its results, the record of what happened when someone counted. This chapter goes beneath the results to the ideas that licensed them, because a method does not spread on intuition alone. It spreads on the authority of a science that seems to vouch for it, and the dramatic method had such a science, or believed it did. Every teacher who staged a cruel situation to teach a lesson was drawing, knowingly or not, on a body of famous experiments that had taught a generation a single electrifying proposition, that ordinary people are clay and the situation is the hand, that the decent become cruel and the cruel obedient when the setting demands it, and that a person can therefore be reshaped by staging the right scene around him. That proposition is the intellectual foundation under the exercise. The eye-color drill, the privilege walk, the prison role-play, the workshop that breaks a person down to build him back, all of them rest on the belief that psychology proved the power of the engineered situation to transform the people placed inside it. That foundation has been crumbling for a decade, and this chapter is the report of the collapse.

The dependency has to be explicit before the demolition begins, because the exercise cannot stand without it. Jane Elliott took up an idea she had not invented, that a day in a rigged room could turn ordinary children into oppressors and victims, a premise the science of her era was busy certifying, the premise that situations are sovereign and selves are soft. The whole pedagogy of staged cruelty depends on two claims that the famous experiments were taken to establish. One is that a strong situation can rapidly produce cruelty or submission in people who arrived decent, which licenses the staging. A second is that the state so produced reveals something true and teaches something lasting, which licenses calling the stage a classroom. Strip away the experiments and the exercise is left asserting, without support, that humiliation is educational and that the behavior it induces is a window into the soul. The experiments were the support. Examine them.

Begin where every textbook begins, with the Stanford Prison Experiment, the single most cited demonstration that good people turn bad in a bad place. The story is known even by people who know no other piece of psychology. In the summer of 1971 a Stanford professor named Philip Zimbardo recruited college students, screened them for normality, and divided them at random into guards and prisoners in a mock jail built in a campus basement. Within days, the story goes, the situation took over. The guards, ordinary young men a coin-flip away from being prisoners themselves, grew spontaneously sadistic, humiliating their charges, inventing cruelties, descending into brutality so fast and so far that the two-week study had to be halted after six days. The lesson was delivered to the world with enormous force. Put a decent person in the role of jailer and the role will eat him; evil is situational; the line between a good man and a torturer is a uniform and a set of keys. For half a century that lesson sat in the textbooks as settled fact and supplied the moral physics for every exercise that followed.

The lesson is no longer tenable, and the experiment's own records are the reason. Beginning in 2014, a French researcher named Thibault Le Texier obtained access to Zimbardo's archive, the tapes and transcripts and memos the famous account had always rested on but few had examined, and what he found, published in 2018 as a book whose title translates as the history of a lie and in 2019 in the journal *American Psychologist*, dismantled the canonical story piece by piece. The guards did not descend into cruelty on their own. The experimenters instructed them. In an orientation the day before the study began, the experimenters told the guards what behavior they wanted, coaching them to be tough, to deny the prisoners a sense of individuality, to create an atmosphere of fear, and the specific torments later described as spontaneous, the middle-of-the-night counts and the petty degradations, were techniques the staff had supplied. The design itself was borrowed, lifted from a classroom exercise a student had run in one of Zimbardo's courses three months earlier, an exercise with a conclusion already in hand, so that the study was built to confirm a result rather than to discover one. The guards were never told they were the subjects of the experiment, and they believed they were research assistants helping to study the prisoners, which means the famous finding about how subjects behave was measuring people who thought they were staff. And the immersion the legend depends on was largely theater. Le Texier found participants who later described performing, playing up distress or brutality because they understood that was the assignment, with the data collected too selectively and incompletely to support the sweeping claims built upon it.

The contamination ran from the top of the study downward, because the man who designed the experiment also played a role inside it. Zimbardo cast himself as the prison superintendent, installing himself within the fiction he was supposed to be observing from outside, and his graduate assistant took the part of the warden who briefed the guards. The experiment had no neutral observer at all. It had a director with a thesis,

stationed in the middle of his own production, cueing the cast and shaping the action toward the conclusion he had announced in advance, and a research team that wanted the prison to turn brutal and signaled that want at every turn. A study in which the lead scientist becomes a character, gives the antagonists their motivation, and roots for the outcome stages a play and reviews its own performance, and the rave review it gave itself became, for fifty years, a finding.

The most dramatic single moment in the whole affair illustrates the performance at work. Its emotional peak is the prisoner who broke down within thirty-six hours, screaming and raging until the experimenters released him, his collapse offered as proof that the mock prison had become psychologically real enough to shatter a healthy young man. The prisoner himself has since told a different story. He has said that he was acting rather than breaking down at all, faking distress because he wanted out of a study he had found boring and confining, and that he was startled the experimenters took his performance for a real psychiatric emergency. Whatever the exact mix of true strain and deliberate theater, the centerpiece of the evidence turns out to be a young man playing a part, and an experiment whose most famous data point is a subject's improvised performance has documented the power of role-play to convince an audience rather than the power of a situation to transform a soul.

The corroboration arrived from another quarter the same year. The psychologists Alexander Haslam and Stephen Reicher, examining the same archive, concluded that the guards had been actively encouraged to adopt the role, rather than conforming to it of their own accord, by experimenters exercising a kind of leadership who persuaded the guards that harshness served a worthy scientific cause. Years earlier the two had run a prison study of their own for the BBC in which they pointedly declined to coach the guards or to take the commanding role Zimbardo had taken, and the result looked nothing like Stanford. The guards did not spiral into tyranny, the group dynamics ran in unexpected directions, and the contrast pointed at the obvious culprit. The brutality at Stanford came from being told what the experimenters wanted and rewarded for producing it, rather than from the bare fact of being placed in a situation. Zimbardo defended his work after the criticisms appeared, insisting the demonstration still revealed something real about dehumanizing environments, and the debate continues. The plain reading of the records, though, is hard to escape. The most famous proof that situations manufacture cruelty was a situation in which the cruelty was manufactured by the experimenters.

What allowed the legend to stand for half a century is a question that indicts more than one study. The Stanford Prison Experiment was never the rigorous trial its fame implies. It had no control group, a handful of subjects, and no proper comparison against which to measure what the situation supposedly caused, and its results appeared less in the hard channels of peer review than in Zimbardo's own retellings, a slide presentation,

a documentary, and a long career of vivid lectures. The study persuaded the public the way a film persuades, through narrative and image and the charisma of its author, and it entered the textbooks on the strength of how unforgettable the story was. A weak experiment with an indelible story will beat a strong experiment with a forgettable one every time, in the marketplace of attention that decides what gets taught, and the prison study won that market for decades on showmanship while the questions about its rigor sat in plain sight, asked occasionally and waved away. The same force that keeps a useless workshop running kept a broken experiment canonical, the preference of audiences for a gripping demonstration over a careful one.

Sit with the recursion in that finding, because it is the heart of the chapter. The Stanford Prison Experiment was itself a piece of staged moral theater. It built a fake institution, assigned people to a caste by the flip of a coin, licensed one caste to degrade the other, demanded the cruelty it claimed to observe, and produced a vivid spectacle that everyone mistook for a truth about human nature. What looks like an experiment studying the dramatic method is the dramatic method itself, staged by a professor instead of a schoolteacher, in a basement instead of a classroom, and mistaken for science because it wore the costume of science. The foundational proof of the pedagogy of staged cruelty turns out to be an instance of the pedagogy of staged cruelty, a coached performance with a foregone conclusion, and when the exercise cites Stanford it reaches toward an earlier performance of the same method rather than toward an external authority that validates it, an echo mistaken for a confirmation.

The second pillar is Milgram, and it cracks along the same line. Stanley Milgram ran his obedience studies at Yale in the early 1960s, and the result became one of the most quoted numbers in the history of social science. Volunteers were told to deliver electric shocks of rising voltage to a man in another room whenever he erred on a memory task, the man was an actor who screamed and begged and eventually fell silent, and Milgram reported that sixty-five percent of his subjects obeyed an experimenter’s instruction all the way to a final switch marked with a lethal voltage. The finding landed in a culture still trying to comprehend the Holocaust, and it was received as the explanation, the proof that ordinary people will commit atrocities on the order of an authority, that the guards at the death camps were men in a situation rather than monsters, that any of us would do the same. That shock machine became shorthand for the darkness in the average soul, and it joined Stanford as the empirical bedrock beneath the claim that a staged situation can turn decent people into instruments of harm.

The number was an artifact of selection, and the fuller record tells a different story. A psychologist named Gina Perry, working through Milgram’s own archive at Yale and tracking down his surviving subjects, established what the famous figure concealed. The sixty-five percent came from a single version of the experiment, twenty-six of forty people in one particular setup, while Milgram ran more than two dozen variations

involving over seven hundred subjects, and across those variations the rate of full obedience swung wildly, from near nothing to almost everyone, depending on how the scene was arranged. Choose a different condition and the headline could as easily have been that most people refuse. Across the conditions taken together, defiance was a common outcome, often the dominant one, so a body of data that could honestly have been summarized as a study of how people resist authority was instead sold as proof that they surrender to it. Nor was the procedure the clean, standardized thing the writeups implied. The experimenter improvised, pressing reluctant subjects far past the four scripted prompts the method described, cajoling and ordering and inventing pressure on the spot, so that what was measured was less an iron law of obedience than the persistence of one man determined to extract a result.

The variations Milgram ran, mostly left out of the famous retelling, are where the idea of blind obedience finally dies. When the experimenter delivered his orders by telephone from another room instead of standing at the subject’s shoulder, compliance collapsed, and many subjects who had seemed obedient began secretly administering lower shocks than instructed, lying to the absent authority. When the subject had to hold the victim’s hand onto the shock plate, forcing direct physical contact with the suffering, obedience fell further still. Bring two experimenters into open disagreement about whether to continue, and almost no one went on. Let the subject watch a peer refuse, and refusal turned contagious, most following the rebel rather than the authority. Read as a set, the conditions describe a person constantly weighing proximity, contact, social cues, and the legitimacy of the command, adjusting his behavior to each, which is the portrait of a moral agent doing visible if imperfect ethical work. The single number that made Milgram famous froze one frame of that process and sold it as the whole film, and the whole film shows people reasoning, resisting, and often refusing, the opposite of the automatons the legend requires, under conditions it conveniently forgets.

The detail that should have buried the legend concerns who obeyed and who refused. Perry found, in Milgram’s own materials, that the subjects most likely to defy the experimenter and stop were the ones who believed the shocks were real, while compliance was concentrated among those who suspected, correctly, that the whole thing was a setup and no one was being hurt. The de-hoaxing meant to reassure participants was so thin that perhaps six hundred people left Yale still believing they had electrocuted a stranger, one of them checking the New Haven obituaries for two weeks afterward to see whether he had killed a man. Read those facts together and the moral of the shock machine inverts. The experiment did not show that ordinary people will torture an innocent on command. It showed, to the extent it showed anything stable, that people who grasped a real human being was suffering tended to refuse, and that the obedience Milgram celebrated leaned on subjects who had seen through the theater and were playing along. Perry’s judgment, after years in the archive, is that Milgram wanted

a dramatic figure to make his name and shaped the presentation of his findings to deliver it.

A more recent reading of the obedience studies turns them against the pedagogy even more directly. Haslam and Reicher, the same researchers who reexamined Stanford, have argued that what Milgram captured was less blind obedience than what they call engaged followership, the willingness of people to go along with an authority when they have been persuaded that the authority serves a cause they value, in this case the advance of science. People did not switch off their judgment and submit to a uniform. They reasoned, they weighed the experimenter's appeals to the importance of the research, and they cooperated to the degree they were convinced the cause was worthy and balked when they were not. Whatever else this revised picture is, it is fatal to the premise the exercise requires. The exercise needs human beings to be clay, passively reshaped by the situation pressed around them. The corrected science describes human beings as agents who interpret, resist, see through manipulation, and decide, which is the opposite of clay, and a pedagogy built on the belief that people are putty in the hands of a strong situation is building on a reading of Milgram that the evidence no longer supports.

Much of the cultural force of the obedience studies came from a single electrifying analogy, the claim that Milgram had found the psychology of the Holocaust, the ordinary mechanism that turned clerks and soldiers into accessories to genocide. Milgram encouraged the comparison, and it fused in the public mind with the philosopher Hannah Arendt's account of Adolf Eichmann as a thoughtless bureaucrat who organized mass murder out of careerism rather than conviction, her phrase about the banality of evil becoming the motto of the whole idea. Both halves of that synthesis have since been contested by careful scholars. Later historians working from Eichmann's own words, written and recorded before his capture, have argued that he was a committed ideologue who believed in the killing, not a hollow functionary merely following orders, which dismantles the figure on whom the banality thesis rested. The dramatic experiment and the dramatic historical claim propped each other up, the lab seeming to explain the genocide and the genocide seeming to validate the lab, and both turn out to have flattened a complicated reality into a single memorable shape. What falls here is the cartoon, the notion that one man's careerism or one lab's shock generator captures the whole psychology of genocide, and an earlier chapter took care to lean on neither cartoon. The durable point survives the correction intact, that ordinary people can be brought to do terrible things when an authority they credit assures them the cause requires it. That claim never depended on Eichmann being thoughtless or on Milgram's sixty-five percent being exact, drawing instead on the larger and quieter record of how atrocities are actually staffed, a record no revision has overturned. A frightening story that explains atrocity in one stroke is satisfying in the way the dramatic

method is always satisfying, and it is unreliable in the same way too.

Step back from the two pillars to the ground they stand on, because their fall was part of a wider reckoning that has shaken the whole discipline rather than an isolated pair of scandals. Beginning around 2011, psychology entered what is now called its replication crisis, a period of discovering that a large share of its published findings, including many of its most celebrated, could not be reproduced when independent teams ran the studies again with adequate rigor. The signal event was a project organized by the Open Science Collaboration and published in the journal *Science* in 2015, an effort that repeated one hundred studies drawn from three leading psychology journals, using the original materials and larger samples. Only about a third of the replications produced a result in the same direction with statistical significance, and the effects that did survive shrank on average to roughly half their original size. A field that had assumed its published record was largely sound learned that something close to two out of three marquee results would not come back when tested honestly, and that the published effect sizes were inflated even where the effects were real.

The interpretation of that project has been contested, and fairness requires saying so. A group of prominent researchers led by Daniel Gilbert argued in 2016 that the replication effort had underestimated reproducibility, that statistical and procedural choices had biased its results toward failure, and that the data were consistent with a far healthier picture of the field. The original authors answered, the argument over the precise rate continues, and an honest account does not pretend the single number is settled. What is not seriously contested is the larger pattern the crisis exposed, the routine reliance on tiny samples, the quiet flexibility in analysis that lets a researcher hunt through a dataset until something crosses the line of significance, the bias of journals toward surprising positive results and against the null findings that would have corrected them. Those practices were standard for decades, and they were exactly the practices that manufacture dramatic findings out of noise. The crisis is real whatever the exact reproducibility rate, and it falls hardest on the kind of result that made the dramatic paradigm famous.

How noise becomes a dramatic finding is no longer a mystery, because the methodologists have shown the machinery in detail. In 2011 a trio of researchers published a paper on what they called false-positive psychology, demonstrating that the ordinary freedoms a scientist enjoys, the choice of when to stop collecting data, which measures to report, which conditions to compare, which outliers to drop, are together so powerful that a determined analyst can manufacture a statistically significant result for almost any hypothesis, however absurd. To prove it they ran a real experiment with real subjects and reported, with a straight face and a passing grade from the standard statistics, that listening to a particular Beatles song literally made people younger in age, an impossibility produced by nothing more sinister than the routine flexibility every

researcher had been using. The point was that the tools of the field, used in good faith and without any sense of cheating, generate false discoveries as a matter of course, and that the brighter and more surprising the claimed effect, the more likely it was assembled this way. The dramatic findings were the field's most natural errors rather than its triumphs, the places where wishful analysis and a hunger for astonishment met and produced a result too good to be true and too vivid to question.

A single thread ties the prison, the shock machine, and the failed replications together, and it has a name in the discipline, demand characteristics. Decades ago the psychologist Martin Orne pointed out that people who agree to be studied stop behaving naturally, working instead to figure out what the experimenter is after and, being cooperative creatures, to supply it. A subject reads the cues, forms a hypothesis about the desired result, and performs accordingly, so that an experiment which telegraphs its wishes will tend to receive them regardless of any real effect. This is the mechanism beneath every story in this chapter. The Stanford guards gave their director the brutality he plainly wanted. Milgram's compliant subjects, the ones who suspected the setup, gave the experimenter the obedience the situation signaled was expected. The priming and posing effects appeared most reliably in the laboratories of the believers who expected them and faded in the hands of skeptics, which is the fingerprint of demand rather than discovery. A dramatic study is the most vulnerable of all to this contamination, because its whole design is a broad hint, a stage so heavily dressed that the actors can read their parts off the scenery, and the result it harvests is the audience's cooperation mistaken for a law of the mind.

The marquee casualties make the point concrete, because the flashiest effects fell the hardest. Consider power posing, the claim that holding an expansive, confident posture for a couple of minutes raises testosterone, lowers stress hormones, and makes a person bolder and more successful. It came from a 2010 paper and was carried to global fame by a 2012 conference talk that became one of the most watched ever given, followed by a bestselling book, the whole package telling anxious people that they could remake their fortunes by standing like a superhero before a job interview. Then the replications came. A larger study in 2015 found that the pose changed how powerful people felt and did nothing measurable to their hormones or their behavior, and a reanalysis of the wider literature concluded that the effect was either absent or too small to have been detected by the studies claiming it. In 2016 the lead author of the original paper posted a public statement that she no longer believed the power-pose effect was real and acknowledged the analytic practices that had produced the illusion. The same fate met ego depletion, the influential idea that willpower is a finite fuel that drains with use, a notion that launched hundreds of studies and a popular literature, until a large preregistered effort across many laboratories in 2016 ran the standard test and found an effect indistinguishable from zero. Story after story followed the arc, the dramatic and

counterintuitive result, the media triumph, the failed replication, the quiet retraction or the loud defense.

The arc is the lesson, and it implicates Stanford and Milgram by family resemblance. The qualities that made these findings famous were the qualities that should have provoked suspicion. A real effect in human psychology is usually modest, noisy, and dependent on conditions, the sort of finding that needs careful statistics to see and resists a clean retelling. A dramatic effect, large and surprising and morally electrifying, the kind that fits on a stage and converts an audience in a single sitting, is precisely the kind that the crisis taught us to distrust, because that profile is the signature of a small sample, a flexible analysis, and a publication system hungry for astonishment. The dramatic demonstration persuaded by the same mechanism the dramatic workshop persuades, by the force of a vivid spectacle that overwhelms the slow questions, and the persuasion ran ahead of the proof in the laboratory exactly as it does in the training room. Stanford and Milgram were the most dramatic demonstrations of all, and they were sold to the public with the showmanship of a TED stage decades before the term existed, and the corrections that finally caught up with the power pose are the same corrections that have caught up with them.

The discipline’s response to the crisis shows, by contrast, what honest inquiry looks like and how far the dramatic studies fell short of it. The reformers pushed for preregistration, the practice of declaring in advance exactly what a study will measure and how, so that the flexibility Orne and the false-positive paper exposed cannot be used to fish for significance after the fact. They pushed for larger samples, for the publication of null results, for the sharing of data and materials, and above all for direct replication by independent teams as the test a finding must pass before it is believed. Against that standard the famous demonstrations look like relics of a prescientific age. None was preregistered. The samples were far too small to trust. None was built to be replicated, and the prison study could not be replicated even in principle, since the ethics rules written in part because of these experiments now forbid it. The dramatic paradigm flourished in the era before these safeguards existed, in a field that rewarded a startling story and asked few hard questions of it, and the safeguards arrived too late to stop the demonstrations from becoming the furniture of the culture.

Knowing a study is broken, it turns out, does little to dislodge it, and the persistence of the discredited demonstrations is itself a lesson in the power of the vivid. The Stanford Prison Experiment remains in introductory textbooks, often with only a cautious footnote about the controversy, because no footnote competes with the story. Milgram is taught every year to students who will remember the shock machine long after they have forgotten every careful qualification. The power-pose talk has been watched tens of millions of times and continues to circulate, the recantation of its own author reaching a fraction of the audience the original reached. These are intellectual

zombies, findings that have been killed in the literature and go on walking through the culture because the thing that made them famous, their dramatic charge, is untouched by their refutation. A correction is an argument, quiet and conditional, and an argument cannot overtake a spectacle in the memory of a public that met the spectacle first. The trainer who invokes the prison study today is usually repeating what the culture still tells him is true rather than lying outright, which is how a debunked science keeps licensing a discredited method long after both have been exposed.

This is the recursion stated as a general law. The evidence offered for the dramatic method is made of the same material as the dramatic method. Stanford was a staged exercise in coached cruelty with a foregone conclusion, dressed as research. Milgram was a piece of improvised moral theater whose star result was selected for shock value. The wider literature that backed the situationist worldview was inflated by the same practices, the small dramatic study and the suppression of dull contradictions, that the dramatic method uses to sell itself. When a trainer invokes what psychology has shown about the power of the situation, the trainer is not reaching past the exercise to an independent authority. That trainer points at a hall of mirrors, at earlier and more prestigious performances of the same trick, and mistakes the reflection for corroboration. The foundation does not support the building. The foundation is the building, poured a second time and labeled bedrock.

This explains a pattern that would otherwise look like coincidence, the resemblance between the great experiments and the exercises they authorized. The eye-color drill, the prison study, the privilege walk, the shock machine, the corporate workshop that breaks a person down, all share a single template, because they are the same invention appearing under different roofs. Each builds an artificial situation. All assign people to roles or castes. Each applies a controlled cruelty and watches for the reaction. The vivid reaction it provokes is then treated as a revelation about human nature, and the audience is persuaded by the force of the spectacle rather than the weight of the evidence. The laboratory version borrowed the prestige of science and lent it back to the classroom version, so that teacher and professor could each point to the other as proof. Pull the thread and the whole structure of authority unravels at once, because there was never an independent science underwriting the pedagogy. There was only the pedagogy, performed once in a university and once in a schoolroom, each performance citing the other, the method validating itself in an echo chamber it mistook for a body of evidence.

There is a human cost to a false foundation, beyond the waste of believing things that are not so. The situationist demonstrations taught two generations a bleak anthropology, the doctrine that every person is a latent guard, a potential torturer, an obedient instrument of atrocity waiting for the situation that will reveal him. That teaching corrodes rather than humbles, whatever its admirers suppose, because a person convinced he is a monster in waiting is primed to accept the humiliation that claims to

exorcise the monster, and a society convinced that decency is a thin veneer over universal cruelty will reach for harsh pedagogies to scrub the cruelty out. The discredited experiments did more than mislead a field. They supplied a flattering license for contempt, the belief that treating people as latent oppressors is simply facing a hard truth about them, when the truth the better evidence supports is less lurid and more demanding, that people are moral agents who usually resist cruelty when they can see it plainly and who deserve to be addressed as such. A pedagogy that begins by telling its subjects they are oppressors in embryo has insulted them on the authority of experiments that never earned the right to the insult.

The exercise at the center of this book belongs to the lineage now in ruins. When Jane Elliott divided her classroom by eye color, she was performing the same demonstration the laboratory was performing in the same years, the rapid manufacture of prejudice and cruelty by the arrangement of a situation, and she presented the result with the same confidence that it revealed a truth about human nature. Her classroom had even less claim to the authority of science than the laboratory did. It enrolled a few children at a time, kept no control group, gathered no systematic data, and was never replicated under conditions that could rule out the teacher's expectations shaping what she saw, which were the flaws that sank the prison study and the obedience studies. The drill was a piece of theater whose persuasive power lay in its vividness, a teacher watching children turn on one another and concluding that the situation had done it, and the conclusion drew its credibility from a surrounding science that has since been exposed as theater of the same kind. Remove the borrowed authority and the eye-color exercise is what it always was beneath the costume, a striking classroom performance mistaken by its author and its admirers for an experiment.

The collapse of the science strips the method of the one thing that let it override objection, its certificate of authority. A teacher or trainer who proposes to humiliate a roomful of people meets resistance, and the resistance is answered by an appeal to expertise, the assurance that psychology has established the necessity and the safety and the value of the ordeal, that the discomfort is the documented price of a documented benefit. That assurance was always the method's trump card, the move that converted a reasonable refusal into a sign of fragility or denial, and it rested entirely on the prestige of the famous experiments. With the experiments exposed, the certificate is void. The trainer can no longer say the science demands this, because the science demands nothing of the kind and has spent a decade retracting the claims that were taken to demand it. What remains is a person asking other people to submit to engineered degradation on no better authority than his own conviction that it is good for them, which is a request that deserves the answer such requests have always deserved from anyone who has not been talked out of his own judgment.

The conduct of the experimenters after exposure completes the rhyme with the programs of the previous chapter. Confronted with the archival evidence, Zimbardo defended the prison study and insisted it still revealed something true whatever the records showed. Confronted with the failed replications, the champion of the power pose kept the claim alive by narrowing it to a feeling and pressing on, while the lead author who recanted stood out as the exception that proved how rare the recantation is. The pattern is the one the drug warriors and the prison-tour promoters displayed when their numbers came in, the preference for the felt truth of a vivid demonstration over the cold verdict of the data, held now by credentialed scientists rather than police chiefs. A finding that feels revelatory is defended like a possession, and the impulse to protect it outlasts the evidence that should have ended it, in the laboratory as in the school district. The dramatic method corrupts the judgment of the people who practice it and the judgment of the people who study it, by the same mechanism, the conviction that something this striking has to be true.

Fairness requires marking the limits of this demolition, because it would be a cheap and false victory to claim more than the evidence allows. Nothing here shows that situations have no influence on behavior, which would be absurd; circumstances plainly shape what people do, and a later chapter rests part of its constructive case on a situational effect that does hold up under scrutiny. Nothing here shows that human beings are never cruel or never obedient, that the Holocaust had no psychology, or that prejudice is imaginary. The claim is narrower and it is exact. The specific dramatic demonstrations enlisted to license the pedagogy of staged cruelty were oversold, fragile, or staged, and the honest science that has replaced them describes people as reasoning agents who interpret a situation, resist manipulation, and frequently refuse, rather than as clay to be molded by a strong situation. That corrected picture does not support the exercise. It indicts the exercise, because a method that treats people as passive material to be reshaped by engineered humiliation is working from a model of the human being that the best current evidence rejects. The science did not merely fail to back the method. The science, corrected, describes the kind of creature on whom the method should not work and to whom it should not be done.

Honesty also requires turning the blade to face the one who wields it, because a careful reader will already have asked the obvious question. If dramatic findings from small samples are to be distrusted, what protects the findings this book itself relies on, the reactance that makes coercion backfire, the moral licensing behind the rented conscience, the stereotype threat invoked in an early chapter. The question is fair, and the answer is the standard applied without favor. Some of those supporting effects are themselves contested in the same reckoning described here, and where they are, this book does not rest its weight on them. Stereotype threat and moral licensing enter these pages as illustrations of a mechanism, suggestive rather than load-bearing, and if either

were to dissolve tomorrow the argument would stand, because it does not hang from a single laboratory result. What the case against the dramatic method actually rests on is the kind of evidence the dramatic demonstrations never had. It rests on the convergence of field trials across crime and drugs and prejudice gathered in an earlier chapter, programs measured in the world with control groups and counted outcomes, and on the contact finding examined in the chapter on what works, built from hundreds of studies. A claim earns trust in proportion to how it was assembled, and the findings this book credits were assembled the way the situationist legends were not, out of many measurements and replications rather than one unforgettable afternoon. The skepticism of this chapter is no weapon aimed only at the other side. Turned on the book's own sources, it leaves the argument standing, because the argument was built to survive it.

This is the turn that matters most, and it deserves to be put without hedging. For half a century the situationist studies were read as a warning about how easily we become cruel, and the reading licensed a pedagogy that manufactures cruelty in order to inoculate against it. The corrected science delivers close to the opposite message. People resist when the suffering is real and near, defy when a peer shows them it can be done, see through manipulation more often than the legend allowed, and need to be actively pushed, coached, and deceived before a fraction of them will harm a stranger. What the corrected science shows is a portrait of ordinarily decent people whose decency is sturdier than the legend permitted, not the latent monsters held in check by civilization that the old reading imagined. The practical lesson for anyone who wants to reduce cruelty is to strengthen the conditions under which people already refuse it, clear sight of the victim, the example of a dissenter, the freedom to say no, rather than to stage the cruelty and hope the staging teaches. The honest reading of the evidence points toward respect for the moral capacity of the person and away from the engineered degradation that assumes its absence.

This reading, the one that survives scrutiny, does more than discredit the method; it points toward the alternative the final chapters will defend. If people are moral agents who respond to the legitimacy of a request, to the visible humanity of others, and to the example of those around them, then the way to reduce cruelty and prejudice is to work with that agency rather than against it. The situational influence that does hold up is the one Gordon Allport described long before the dramatic studies, the finding that sustained cooperative contact among people of equal standing, working toward a shared goal, lowers the prejudice between them. That is a situational effect, and it treats the people inside it as partners rather than subjects, it persuades rather than coerces, and it has survived the scrutiny that destroyed the staged demonstrations. The honest science of situations, in other words, survives the loss of the dramatic version. It has been clarified, and the clarified version recommends contact, cooperation, and respect, the slow building of conditions in which decent impulses can operate, which is the opposite

of the engineered humiliation the discredited studies were used to justify.

Two condemnations now stand against the dramatic method, and they arrive from opposite directions and meet in the same place. The previous chapter showed that the method fails by its results, that when its programs are measured they do not work and often do harm. This chapter has shown that it fails by its foundations, that the famous experiments offered as its warrant were performances rather than proofs and have not survived examination. A practice discredited by what it does and by the evidence it came from has little left to stand on, and what remains is to ask why it persists and spreads regardless, and where it is spreading now. That question is no longer academic, because the logic of the method, the conviction that virtue can be manufactured by staging the right humiliation, has moved out of the classroom and the corporate retreat and into the politics of a frightened age, claimed by some as the cure for a society's sins and denounced by others as a weapon against its freedoms. The next chapter turns to that argument, to the vulnerable days in which a discredited method has become a public cause, and to what an honest reckoning with the evidence should mean for a republic arguing about how to teach its citizens to be good.

PART SIX

The Cure and the Refusal

CHAPTER FOURTEEN

The Vulnerable Days

The two chapters before this one put the dramatic method on trial as a matter of science and convicted it twice, once for what it does to the people who undergo it and once for the experiments it claimed as its warrant. This chapter changes the venue from the laboratory to the country, because the method did not stay in the schoolroom and the corporate retreat where the earlier chapters found it. It climbed into the machinery of the state and the large institution, and it made the climb at the worst imaginable moment, in a decade frightened enough to welcome it. A society reaches for coercive pedagogy when it is afraid, when the stakes feel total and the patience for slow persuasion runs out, and the years since 2020 have been such a time, a long season of fear and division in which each faction came to believe that the survival of the country hung on correcting the beliefs of the other. Those are the vulnerable days of the title, and they carry a particular danger. They are the conditions under which a free people will pick up the tools of compelled belief and persuade itself that the emergency has earned them.

Being concrete about what made these days so vulnerable is worth the space, because the fear was real and the temptations grew from real soil. The decade opened with a pandemic that confined people to their screens and frayed every nerve, and then, in the spring of 2020, the filmed killing of George Floyd under a police officer’s knee set off the largest protests the country had seen in generations and a sweeping reckoning with the racism still wired into its institutions. The grief and the anger were earned, the injustice they named was old and documented, and the desire to do something equal to the moment was honorable. Into that desire the method poured itself, because a society convulsed by a real wrong is impatient with slow remedies and hungry for a gesture large enough to match its feeling. The reaction, when it came, grew from a fear of its own, the sense among millions that a familiar country was being remade overnight by institutions they distrusted, that their children were being taught to despise themselves and their nation. Two earnest fears, each blind to the other’s sincerity, each certain that the emergency justified extraordinary measures, are the perfect climate for compelled belief, and the years that followed grew the crop the climate predicted.

Everything in this chapter depends on a distinction that the heat of the present works to erase, the distinction between a goal and a method. The goal of a fairer country, one in which a child's race or sex does not set the ceiling on the child's life, is worth defending, and this book has defended it on every page. The method this book attacks has a shape entirely separate from that goal or any other. It sorts people into camps by an accident of birth or identity, installs an authority that cannot be refused, licenses the humiliation of the marked group, stages that humiliation before an audience, and reads any refusal of the ritual as fresh proof of the guilt the ritual exists to expose. Such a shape can be loaded with any politics at all. It can be aimed at the privileged or the marginal, pointed left or right, justified by liberation or by order, and it will run the same way in every case, because it is a technology for manufacturing compelled belief, and a technology answers to whoever holds it. The chapter that follows watches the same instrument pass from one set of hands to another in the space of a few years, and the lesson is in the passing.

One clear way to see the method as an instrument of power is to look at the regimes that wielded it without embarrassment. The starkest case is the Maoist struggle session, the public denunciation ritual that reached its height during China's Cultural Revolution between 1966 and 1976. An accused person, often a teacher or an official or anyone marked as a class enemy, was brought before a crowd, made to bow, forced to confess crimes of thought, and subjected to hours of shouted accusation and humiliation, sometimes beaten, sometimes driven to death. The settings were schools and workplaces and auditoriums, the ordinary rooms of daily life turned into theaters of denunciation, and the genius of the thing was its conscription of the crowd. Students were turned against their teachers, friends pressed to betray friends, children coached to expose their parents, until the bonds that held a society together had been converted into instruments of its terror. The struggle session was thought reform by spectacle, the manufacture of a new belief through the staged degradation of the old, and it is the purest specimen the modern world has produced of the method this book describes, scaled up from a classroom exercise to an engine of national fear.

The second example wore the opposite politics and a far milder face, which is the point of pairing them. At the height of the American anticommunist panic, the University of California required its faculty to sign a loyalty oath, a sworn statement that the signer was not a communist and held no proscribed beliefs. The faculty, some nine hundred strong, balked, and a standoff ran through 1949 and 1950 until the Regents voted in 1950 to fire dozens of professors and other employees who would not sign, the oath soon hardening into a state statute that bound public workers across California. A court ordered the fired academics reinstated the next year, and the California Supreme Court made the reinstatement final in 1952, but by then the lesson had been delivered, that a person's livelihood could be made to depend on a public profession of correct

belief, and that the refusal to profess was treated as confession of the disloyalty the oath was meant to catch. One regime was murderous and one was bureaucratic, and the distance between them in scale and cruelty is enormous and must never be flattened into a glib equivalence. The instrument, though, was recognizably the same in both hands, a compelled public profession of belief, enforced by an authority that controlled a person’s standing or his life, with dissent read as guilt. The form is the danger, and the form does not care which cause it serves.

A loyalty oath carries a cruelty its designers rarely intend, which is that it catches the wrong people. The committed liar signs without a tremor, since the words cost him nothing and the job is worth a signature, and he goes on believing exactly what he believed before. The person it traps is the one with a conscience, the honest dissenter who cannot bring himself to swear to something he doubts, or the scrupulous believer who will not cheapen a real conviction by reciting it on command. An instrument meant to filter out the disloyal filters out the principled instead, rewarding the cynic who performs and punishing the honest soul who declines to, which is the same perverse selection the earlier chapters found in the laboratory, where the subjects who saw through Milgram’s deception were the ones who refused to play along. Compelled belief measures something other than belief. It measures the willingness to be compelled, and the people most willing are seldom the people a decent institution should most want to keep.

Carry that instrument into the present and the first place it appears is on the political left, in the institutions the earlier chapters already examined. The mandatory diversity workshop, the session that gathers employees to be told they are complicit by virtue of their race and invited to confess the privilege the trainer has assigned them, is the struggle session in a softened and bureaucratic register, the same staged contrition performed under fluorescent light with a human-resources facilitator in the warden’s chair. The academic version was more formal and more consequential. For roughly a decade it became standard for universities to require a diversity statement from anyone applying for a faculty job, a written profession of the candidate’s commitment to the doctrine and a record of the candidate’s service to it. The statement was frequently the first gate in the process, read before the scholarship, so that a brilliant chemist or historian could be cut from the pool before anyone examined the work, on the strength of an insufficiently fervent paragraph about equity. At one campus of the University of California, a review found that as many as three out of four applicants were eliminated on their diversity statements alone, the science and the teaching never reaching a human eye.

Look closely at what a diversity statement screens for and the objection sharpens past the question of compulsion. It does not test whether a candidate treats students fairly, since fairness shows in conduct over years and cannot be read off a paragraph

written to order. It tests fluency in a particular dialect, the ability to produce the expected vocabulary with the expected warmth, and that is a test the smooth performer passes and the awkward egalitarian fails. A chemist who has quietly mentored every struggling student of every background for thirty years, but who writes a flat and dutiful statement because the genre embarrasses her, scores below a fluent newcomer who has mentored no one and writes beautifully about equity. The instrument rewards the performance of virtue over the practice of it, which is the counterfeit this book has tracked from the beginning, the rented conscience installed now as a hiring filter. An institution that selects its faculty this way selects for the willingness and the skill to perform fairness on demand rather than for fairness itself, and it teaches everyone who wants a job that the performance is what the institution rewards.

The corporate wing of the movement supplied its own proof that the method was a posture rather than a conviction, and it supplied the proof with embarrassing speed. After the upheavals of 2020, companies rushed to mandate diversity training and to staff diversity offices, competing to announce their commitment in the moral language of the moment. When the political weather changed a few years later, the same companies cut the same programs with the same haste, retiring the offices and quietly deleting the pledges, and a tracker kept by a higher-education trade publication counted scores of colleges following the corporate lead and dismantling the apparatus they had built. A conviction does not evaporate the moment it stops being fashionable. A posture does, and the speed of the retreat revealed what the buildup had been, a response to social pressure rather than a considered judgment about what reduces injustice. The institutions had embraced the struggle session when embracing it was the admired thing, and they deserted it when deserting it became the admired thing, and at no point in either the embrace or the desertion is there much sign that anyone asked whether the method worked. Fashion explained the whole arc, the rise and the fall alike, which is the surest mark of a practice chosen for display.

The people who named this instrument most precisely were liberals inside the academy who recognized the shape of the thing, rather than its enemies on the right. A mathematician at Davis warned in 2019, in the pages of her professional society’s journal, that the diversity statement was an echo of the old loyalty oath, a political test enforced with real consequences for those who failed it. In a survey of faculty by a free-speech organization a few years later, a majority of professors who called themselves moderate, and a quarter of those who called themselves liberal, agreed that the statement functioned as an ideological litmus test. A liberal professor at Harvard’s law school said plainly that the requirement made him wince, and a celebrated psychologist at the same university called the practice a thinly disguised loyalty oath that ought to be abolished. These were not conservatives hunting for a grievance. They spoke as people sympathetic to the goal who could see that the method had curdled into

compulsion, that a profession of belief extracted as the price of a job is a profession worth nothing and a precedent worth fearing. The honesty did the cause credit, and it conceded the central charge of this book from inside the movement that had adopted the method, that the staged confession is the oppressor's tool no matter how just the cause that picks it up.

That instrument did not remain on the left, and the speed with which the other side seized it is the heart of the chapter's warning. The first move came in September of 2020, when an executive order from the federal government barred agencies, the military, and federal contractors from conducting training that promoted a list of forbidden ideas about race and sex, among them the notion that the country is racist at its root, that a person bears responsibility for wrongs committed in the past by others of his race, or that anyone should be made to feel discomfort or guilt on account of his race. Federal money was tied to compliance, contractors who violated the order could be barred from future work, and the government set up a hotline so that employees could report their colleagues, an apparatus a civil-rights legal group described as McCarthy-esque. The order was rescinded a few months later by a new administration, but it had already become a template. Researchers at a law-school project that tracked the campaign counted, by the time it closed its count at the end of 2024, more than eight hundred fifty such measures introduced across forty-nine states and the District of Columbia, most of them aimed at public schools, most of those aimed at restricting what teachers could say, and many of them lifting the forbidden-concepts language directly from the federal order that had inspired them. A prominent state law made it a form of actionable discrimination to subject a student or worker to any instruction that produced discomfort or guilt on account of race, the precise mirror of the harm the diversity workshop had been manufacturing on purpose.

The campaign on the other side carries a harm the workshops did not, and an honest account has to name it even while condemning the shared method. A mandatory training, for all its coercion, inflicts a bad pedagogy while leaving every fact still learnable, the history of the country available in every library. A law that bars a teacher from presenting the documented record of slavery, segregation, and their long aftermath reaches past method and attacks knowledge itself, telling a student there are things about his own nation he may not be taught. The vagueness compounds the injury. When a statute forbids instruction that might produce discomfort on account of race, and leaves discomfort undefined, the cautious teacher who cannot afford a lawsuit avoids the subject altogether, and a whole region of history falls into shadow, less because anyone banned the specific lesson than because no one could be sure which lesson was safe. On the goal, then, the two campaigns are not equals, and this book will not pretend they are. On the method they share, the compelled profession and the policed belief, the verdict is the same, because restricting by statute what a teacher may say is the same act of

compulsion as extracting a confession by workshop, enforced this time by the loss of a job or a grant rather than the loss of a colleague's good opinion.

The turn completed itself at the start of 2025, when the federal government changed hands again and the instrument was pointed back the other way with the full power of the state behind it. A cluster of executive orders abolished diversity offices and positions across the government, rescinded the affirmative-action obligations that had bound federal contractors since the 1960s, and required every contractor and grantee to certify that it ran no diversity program offending the administration's reading of the civil-rights laws. The attorney general was directed to hunt for private companies and universities whose diversity efforts might warrant investigation. That certification requirement is the diversity statement turned inside out, a compelled profession of innocence of the doctrine rather than of belief in it, sworn under penalty and tied to the flow of federal money, and a federal court recognized it as such within weeks, blocking much of the scheme on the ground that it compelled speech and punished a viewpoint the government disfavored. That injunction did not survive the appeal. In February of 2026, only weeks before this book went to press, the Fourth Circuit set it aside, holding that a challenge to the orders on their face was unlikely to succeed and sending the case back to the trial court, where narrower challenges to particular enforcement actions remain open. The instrument, for the moment, stands. The same constitutional objection that liberals had raised against the loyalty oath, and that conservatives had raised against the diversity statement, now lay against an order written to please the conservatives, because the objection was never about the content at all, only about the compulsion, and the compulsion had simply switched sides.

A threshold is crossed when the method moves from the institution to the state, and the move deserves to be marked, because the two differ in their power to harm. A coercive workshop run by a company can be quit, a job can be left, a campus can be avoided, and the worst a private orthodoxy can do is make a particular employer or school unbearable. The state is not an employer a person can leave. When the government conditions federal money on a profession of belief, threatens investigations of private companies for the content of their training, and reaches through its enforcement agencies into the practices of universities and contractors across the country, it brings to the manufacture of belief a power no workshop ever had, the power of law, taxation, and prosecution. The federal orders of 2025 did exactly this, pairing the certification demand with instructions to hunt out offenders and with a separate order stripping away a long-standing tool for proving discrimination, an escalation from private coercion to public force. That is the moment a bad pedagogy becomes a constitutional danger, when the body applying the compulsion is the one body a citizen cannot escape, and the only authority left to restrain it is the law the compeller is busy bending.

The damage of the state's intervention reaches well past the programs it formally ends, into the larger territory of what people do from fear. When an administration announces investigations of private companies for the content of their training and conditions federal money on a profession of innocence, the rational response of a cautious institution is to scrub anything that might draw scrutiny, including efforts that break no law and harm no one. Lawful programs vanish, less because a court found them illegal than because their sponsors cannot afford the risk of being made an example, and the chilling effect does the work a statute could not, manufacturing conformity through the threat of ruin. The first version of the order, five years earlier, had worked the same way from the other direction, when contractors canceled established training and trainers lost their livelihoods over an order that was later rescinded, the fear outrunning the law and reshaping behavior before any case was decided. A government learns that it need not win in court to win in practice. It need only make the cost of noncompliance high enough that people surrender in advance, which is how compelled belief operates even when its formal instruments are struck down.

Set the two campaigns beside each other and they resolve into mirror images, each the photographic negative of the other. One side built a regime in which a profession of faith in diversity was the price of entry to a profession; the other built a regime in which a profession of innocence of diversity was the price of a federal contract. Workshops on the one side were engineered to produce guilt by race; laws on the other made it illegal to produce guilt by race. One side screened out the candidate who would not affirm the doctrine; the other screened out the contractor who would not forswear it. Each measure was a loyalty oath, and the two oaths contradicted each other word for word while sharing a grammar down to the bone, the grammar of compelled belief enforced by the people who control the money and the jobs. Each side adopted the instrument in the sincere conviction that its own cause was the exception, the one good reason that justified coercion, and each side pointed at the other's coercion as the provocation that made its own necessary. The method is contagious in exactly this way. Beyond failing to persuade, it calls its own reflection into being, handing the other side both the weapon and the excuse.

The contagion works as a loop, and naming its turns matters, because each side experiences only its own half and calls that half a response. The struggle session and the loyalty-oath statement, spreading through the schools and the corporations, produced exactly the resentment the science predicted, and the resentment found a political voice that promised to smash the apparatus. That promise elected the people who wrote the concept bans, and the concept bans, with their vague threats and their policing of teachers, produced exactly the censorship that let the other side say, with justice, that the right had become the true enemy of free thought. No single cause drives a national

backlash, and the reaction against the workshops drew on much besides them, on partisan media and electoral calculation and the ordinary opportunism that turns any grievance to advantage. The honest claim is narrower and still damning. The coercive exercise was a real and avoidable contribution to the cycle, a genuine provocation the other side could point to and exploit, one excess feeding the next, the workshop lending fuel to the backlash, the backlash lending cover to the book ban, the book ban lending the workshop its air of vindication, and around the circle it went, every turn leaving both sides more certain that the other had started it and that coercion was the only language left. A country caught in that loop has stopped having an argument. It is conducting two purges that take turns, and the purges need each other, since the cruelty of each is the recruiting poster for the other.

The mirror is true as far as it goes, and it does not go all the way, so the image needs a correction before it misleads. A mirror shows two things of equal size, and the two campaigns are not of equal size, however identical their grammar. Hold the difference in plain view. The diversity statement compels a person to praise a doctrine he may not hold, and the cost of refusing is a job he can seek elsewhere, in a country whose libraries still hold every book. The concept ban reaches past the person into the shelf, empties the classroom of a documented history, and backs the emptying with the criminal law of the state, which a citizen cannot quit the way he quits an employer. One is a coerced sentence a person must speak. The other is a true sentence a person may no longer be taught. Both are compulsion and both are wrong, and a reader who takes only the symmetry from this chapter has taken half of it, because the method is shared and the damage is not. This book holds both truths at once, that the instrument is the same in either hand and that the hand reaching for the state’s full power to erase what may be known has gone further into the dark. The symmetry condemns the method. The asymmetry remembers that a banned book is a graver thing than a compelled compliment, and refusing to forget that completes the argument rather than retreating from it.

A reader keeping count will catch the book in an apparent inconsistency here, and the catch deserves an answer rather than a dodge. This book spends most of its length on the workshop and the classroom and far fewer pages on the loyalty oath and the concept ban, which sits oddly beside the admission that the ban is the graver wrong. The proportion is real, and I will not pretend it is balanced, because it is not, and it is not meant to be. This is a book about a teaching method, and the teaching method that has spread furthest, into the most schools and the most workplaces and the most ordinary lives, is the coercive exercise I came up through and can take apart from the inside. The page count follows the subject, not the severity. Where the method is thickest the analysis goes deepest, and the method is thickest in the schoolroom and the training session, which is why they fill these chapters while the state’s censorship, graver and

rarer and already opposed by half the country, takes fewer. The proportion of attention records where the particular machine this book studies has done its widest work, and a reader should not read it as a ranking of harm.

The reason a frightened age reaches for an instrument it would denounce as tyranny in its opponents' hands is not stupidity, and it is worth understanding rather than scorning. Fear collapses the distance between disagreement and threat, so that the person who holds the wrong view stops looking like a fellow citizen to be argued with and starts looking like a danger to be neutralized. When the stakes are felt as total, the survival of the republic or the soul of the nation, the slow and uncertain work of persuasion comes to seem like a luxury the emergency cannot afford, and compulsion comes to seem like responsibility. The presence of the other side's coercion finishes the argument, because nothing licenses a hard measure so thoroughly as the belief that one is only answering in kind. Each escalation is defensive in the mind of the person who commits it, and the sum of the defensive escalations is a ratchet that tightens in one direction only, toward a politics in which both sides agree that belief should be compelled and disagree only about which belief. That is the road by which republics lose themselves, through a thousand reasonable-seeming surrenders rather than a single thunderclap, each side proving the other's point.

A final irony attends the use of compulsion to spread a belief, which is that compulsion is the worst available method for the purpose. A confession extracted under threat changes no mind; it produces the words without the conviction, the nod in the room and the resentment in the parking lot, and it teaches the person compelled to hide what he thinks rather than to think differently. The belief does not spread. It goes underground, hardening in private into the opposite of what the ritual intended, so that the workshop meant to reduce prejudice often breeds a quiet, resentful version of it, and the loyalty oath meant to secure agreement secures only a population skilled at faking agreement. Persuasion has the reverse property. A person argued with as an equal, shown evidence, allowed to object and be answered, can come to believe a thing, and the belief he reaches that way is his own and will hold. Every regime of compelled belief forgets this, mistaking the silence it imposes for the assent it wants, and discovers too late that it has produced a nation of practiced liars rather than a nation of the persuaded.

A fair reader will object that this is a lazy equivalence, that to set the diversity statement next to the book ban is to pretend that two unequal things are the same, and the objection deserves a precise answer rather than a dismissal. The two sides are not equivalent in their goals, and nothing here says they are. The aim of opening an institution to people it once excluded is a worthy aim, and the aim of forbidding a teacher to discuss the documented history of race in his country is not its equal, because the second reaches past the method and attacks the knowledge itself. On the goals, the

ledger does not balance, and an honest writer should say so. The claim of this chapter is narrower and it survives the objection intact. It is a claim about the method, not the mission, and on the method the ledger does balance, because compelled belief is compelled belief whether it compels a confession of guilt or a renunciation of an idea, and the apparatus that enforces the one is the apparatus that enforces the other. To attack a coercive workshop is legitimate criticism. To answer it by empowering the state to ban concepts and demand loyalty certifications is a second helping of the method rather than a correction of it, and the people who cheer the second while condemning the first have objected to their enemy’s tactics and not to the tactics themselves.

There is a simple test that separates a principle from a method, and it clarifies the whole dispute. Ask whether you would accept the instrument in the hands of your enemy. A person who believes in free expression accepts that his opponent may speak, because the principle protects the speech he hates along with the speech he loves, and a person who believes in equal treatment accepts that the law binds his own side as it binds the other. A method passes no such test. The advocate of the mandatory diversity statement would be appalled to see a future administration require a mandatory statement of patriotic belief, and the advocate of the concept ban would be appalled to see a hostile legislature ban the concepts he holds dear, which tells each of them, if he is listening, that he has been defending a weapon and not a value. Whatever you would refuse to let your enemy hold, you should refuse to hold yourself, because in a country where power changes hands the enemy will eventually hold everything you build. The instruments of compelled belief fail this test on both sides at once, and that shared failure, not any tally of grievances, is what places them under the same condemnation.

The hardest application of the test is to one’s own side, and it is the application almost everyone dodges. Seeing the coercion in the enemy is easy, and being appalled by the loyalty oath the other tribe administers and the heretics it burns costs nothing. The discipline a free conscience requires is to turn the same eye inward, to ask whether the ritual one’s own side runs, the confession one’s own movement extracts, the dissenter one’s own faction punishes, is any less an instance of compelled belief because the belief is held to be true. A reader who finishes this chapter convinced only that the other side is the real menace has missed half its point. The menace is the method, and the method lives on the reader’s own side too, dressed in the reader’s own righteousness, and the test of whether a person loves liberty or merely loves winning is whether he can condemn the coercion that serves his own cause. Most people cannot, which is how the vulnerable days stay vulnerable, a whole country of citizens each vigilant against the other’s tyranny and blind to the instrument in his own hand.

What a free society does instead is harder and slower and less satisfying than a ritual, which is the whole reason the vulnerable days tempt people away from it. The liberal order rests on a wager that runs against the grain of fear, the wager that minds are

better changed by argument, evidence, and contact than by compelled confession, and that a belief worth holding is one a person has reasoned his way toward rather than recited under threat. It treats people as equal before the law rather than equally available for humiliation, it protects the right to be wrong because it knows that the power to punish wrong belief is the power that will someday be turned on the truth, and it shelters the heretic out of memory rather than sympathy for heresy, the memory that every orthodoxy now revered was once a minority opinion that a compelled consensus would have crushed. The norms the method tramples, due process and free expression and the refusal to assign guilt by category, protect the powerless rather than obstructing justice for them, because the person with the least power has the most to lose when belief becomes enforceable and category becomes destiny.

The principle the method violates is among the oldest in the liberal tradition, and the country once stated it with rare clarity. In the middle of the Second World War, when the temptation to compel a show of national unity ran at its height, the Supreme Court considered whether schoolchildren could be forced to salute the flag against their conscience, in the case of West Virginia State Board of Education versus Barnette, and it held that they could not. The opinion put the matter in words that ought to hang over every diversity seminar and every loyalty certification alike, that no official, high or petty, has the power to prescribe what shall be orthodox in politics or other matters of opinion, or to force a citizen to confess by word a belief he does not hold. That is the rule both of the present campaigns break. The compelled diversity statement forces the confession the Court forbade, the compelled loyalty certification forces its opposite, and each rests on the conviction the Court rejected, that an authority may decide what its people must profess. A nation that takes that ruling seriously cannot run either apparatus, because the apparatus is the machinery the ruling exists to forbid.

The people with the most to lose from a regime of compelled belief are the people such regimes always claim to protect. A minority is a minority because it is outnumbered and out-powered, and its security in a free country rests less on the goodwill of the majority, which can evaporate, than on rules that bind the majority whatever it happens to feel, the rules of equal treatment, due process, and protected dissent. Those rules are exactly what the method erodes. Every time an institution establishes that belief can be made a condition of employment, that guilt can be assigned by category, that the refusal of a ritual is evidence against the person who refuses, it strengthens a machinery that a hostile majority will one day aim at the minorities the machinery was built to defend. The struggle session is a luxury of the secure, a weapon that feels safe only to those who cannot imagine losing power. A group that knows it may be outvoted should be the last to want belief made enforceable, because the enforcement will not always run in its favor, and the protection it discards to win a present argument is the protection it will need most in the argument to come.

Events of 2025 supplied the argument's proof, faster than even a pessimist would have guessed. The apparatus of compelled belief that one side spent a decade constructing on the campuses and in the corporations gave that side no protection when power changed hands. Seized intact by the other side, it was aimed in the opposite direction, the machinery of the mandatory statement reborn as the machinery of the mandatory certification, the habit of conditioning a livelihood on a profession of belief simply inheriting a new set of approved beliefs. This is what a loaded weapon left on a table does. It waits for whoever enters the room next, and the next occupant is rarely the one who set it down. A movement that builds an engine of coercion in the confidence that it will always hold the controls has misunderstood the most reliable fact about power, that it changes hands, and that the tools built to enforce one orthodoxy are bequeathed, fully assembled, to the enforcers of its opposite. The only protection a free people has ever had against this is to refuse to build the engine, to lose the argument honestly rather than win it by a method that the next election hands to the enemy.

What is finally at stake in the vulnerable days is the difference between a republic and a regime, and the difference lies exactly here. A regime decides what its people must believe and uses its power to enforce the decision. A republic refuses that power on principle, holding that the government and its institutions have no business prescribing the inner convictions of free citizens, and that the freedom to be wrong is the price and the proof of the freedom to be right. Each side in the present struggle, reaching for the machinery of compelled belief, reaches past its immediate opponent for the republic's own restraint, trying to dissolve the rule that says no faction may use the state to manufacture conviction. The victory either side imagines, a country reeducated into its preferred orthodoxy, is a country that has stopped being a republic and become a regime with one party's name on the door. The deepest danger of the method is not its failure to make people good, real as that failure is. Wielded by the state in a frightened time, the method succeeds at something else, the slow conversion of a free country into one more place where power tells the mind what to hold.

Both of the regimes this chapter began with came to an end, and how they ended is its own instruction. The loyalty oaths were struck down in the courts, the panic that produced them burned out, and the country came to regard that whole episode as a stain, a season of cowardice it now teaches its children to recognize and refuse. China's convulsion ended with its instigator's death, and the party that had driven it eventually rendered a formal verdict against it, though the reckoning there remains guarded and incomplete. Coercive moral campaigns share this much of a life cycle, that they pass, and that the societies which ran them look back with a shame proportional to the damage done. The present campaigns will pass as well, and a later generation will study them with the same mixture of bewilderment and embarrassment, asking how sensible people convinced themselves that extracting a confession or banning a book was the road to a

just country. The only question still open is the size of the wreckage between now and then, how many careers ruined, how much trust destroyed, how much of the slow capital of a free society spent, before the shame arrives and the instruments are set down.

Both sides come away from this chapter with something to resent, which is the surest sign the chapter has been fair. The left is told that the method it embraced in good conscience is the oppressor's method, that it does not work, and that it armed the reaction that has now turned the method against it. The right is told that it was correct about the coercion of the mandatory workshop and the loyalty-oath statement, that its grievance was real and its critics were right to wince, and that it then answered a coercive method with a more powerful coercion of its own, betraying the equality before the law and the freedom of belief it claimed to be defending. Neither side is granted the comfort it wants, the comfort of a clean conscience and a guilty opponent. The goal of equality stands undamaged through all of it, defended here as it has been defended throughout, and the verdict falls only where it has always fallen in this book, on the method, condemned now in every hand that has reached for it.

This book attacks the method in every hand precisely because it wants the goal to survive, and the goal is now in danger from its supposed friends. Equality before the law, the open institution, the country that does not cap a child's life at the color of his skin, these are causes worth a life's work, and they have been set back, not advanced, by the people who yoked them to a coercive method. The struggle session left the workplace no fairer and made fairness look like a racket, handing the cause's enemies their most effective weapon, the true and damaging charge that the movement for equality had become a movement for compelled confession. Anyone who wants the goal should want the method gone more fiercely than any opponent does, because the method is the millstone around the goal's neck, the thing that lets a cynic dismiss a just cause by pointing at the ugly ritual performed in its name. Defending equality and refusing the struggle session are one project rather than two opposed ones, and the chapters that remain take up the constructive half of it, the search for the ways that work.

None of this means a free society does nothing about injustice, and the chapters ahead will insist it can do a great deal. The point is that the real work is harder and quieter than the ritual, and that is why institutions under pressure reach for the ritual instead. A struggle session can be scheduled for a Tuesday and finished by lunch, a diversity statement can be required with a line in a job posting, a concept can be banned with a single vote, and each produces a visible token of resolve that a frightened public can see. The things that change how people treat one another, sustained contact among equals, shared work toward a common end, institutions built so that fairness is the default rather than the performance, take years and show slowly and photograph badly. A society in a hurry chooses the gesture over the work every time, which is the deepest reason the vulnerable days produce so many rituals and so little change, and the reason

the constructive chapters that follow ask for a patience the frightened age would rather not spend.

Three condemnations of the dramatic method now stand together. The science showed that it fails by its results and collapses at its foundations, and this chapter has shown that its political career ends in a weapon passed back and forth between factions that have forgotten how to persuade. What remains is the question the whole book has been moving toward, the question a reader has every right to demand once the demolition is complete. If the staged confession is the wrong way to make a person better, and the evidence and the history and the politics all say that it is, then what is the right way. The next chapter turns from the autopsy to the cure, to the methods that have survived the scrutiny that destroyed the dramatic ones, the patient and unglamorous work of contact and cooperation and honest instruction by which human beings are brought closer to one another, and which the vulnerable days have made it more urgent, not less, to defend.

CHAPTER FIFTEEN

What Works

A book that spends fourteen chapters tearing something down owes its reader a different chapter at the end, the one that says what to build in the cleared ground. The demolition is finished. The dramatic method has been convicted by its results, by the collapse of the experiments it claimed as proof, and by the political career that turned it into a weapon, and a reader who has followed the argument this far has every right to ask the hard question that an autopsy cannot answer. If staging a cruelty is the wrong way to teach a person to be fair, what is the right way. There is good news, and it is that the question has answers, that decades of careful work have identified methods which reduce prejudice and open institutions and survive the scrutiny that destroyed the demonstrations. The better news is that those methods turn out to be the photographic negative of the one this book has spent its length attacking, so that the cure and the disease can be told apart at a glance, by everything they reverse.

This contrast deserves a single frame before the evidence fills it in, because the frame is the argument. The dramatic method coerces; what works engages. Where the dramatic method shames, what works respects. The dramatic method assigns guilt by category; what works builds cooperation among equals. Where one stages a catharsis and calls the feeling a result, the other changes a structure and measures the outcome over years. The dramatic method treats the person as clay to be remolded by a strong situation; what works treats the person as an agent to be persuaded, included, and given something useful to do. Every term inverts, and the inversion follows from something deeper than style. It is the difference between a method built on contempt for the person and a method built on respect for him, and the evidence in this chapter is, at bottom, evidence that respect outperforms contempt at the task contempt claims as its own.

The foundation was laid in 1954, in a book by the Harvard psychologist Gordon Allport called *The Nature of Prejudice*, and it has held up better than almost anything published in social psychology since. Allport proposed what came to be called the contact hypothesis, the idea that prejudice between groups tends to diminish when members of those groups come into contact with one another. The crucial part, the part

that careless summaries drop, is that he did not claim mere proximity would do it. He specified the conditions under which contact helps, and the conditions are exacting. The groups must meet as equals rather than in a hierarchy that confirms the stereotype. They must work toward a common goal rather than compete. The cooperation must be real, structured into the task so that the members need one another. And the surrounding authority, the law or the institution or the custom, must sanction the contact rather than undermine it. Set those conditions and contact reduces prejudice; omit them and contact can harden it, which is why a desegregation that throws hostile groups together without changing the structure around them often makes matters worse. The conditions are the whole point, and every one of them is the opposite of the staged exercise, where status is rigged to humiliate, where the goal is exposure rather than common work, where the design induces hostility rather than cooperation, and where the authority licenses the cruelty instead of restraining it.

The reason the conditions matter so much becomes clear once you ask what contact changes inside a person. Prejudice lives partly on abstraction, on a category called them that is filled with imagined traits and never corrected by a real face, and contact under the right conditions replaces the abstraction with particulars, the specific coworker or classmate who refuses to behave the way the category said he would. The mechanism the later research identified is twofold. Sustained contact reduces the anxiety that makes a stranger feel like a threat, and it builds the capacity to take the other person’s point of view, the two changes together doing the work no lecture about tolerance can do. This also explains the failures, which confirm the theory rather than contradicting it. Contact between groups of unequal status, in conditions of competition or fear, gives the prejudice fresh evidence rather than correcting it, because a hostile encounter with one member of a despised group is read as proof of what the whole group is like. The staged exercise engineers precisely the toxic version. It manufactures inequality, competition, and fear in a single afternoon, which is why, measured against the conditions that make contact heal, it is built to make contact harm.

For half a century after Allport, the contact hypothesis accumulated studies, and in 2006 two researchers, Thomas Pettigrew and Linda Tropp, gathered the whole literature and submitted it to the kind of test the previous chapters found so often missing. Their meta-analysis pooled more than seven hundred independent samples drawn from over five hundred studies, comprising a quarter of a million people across thirty-eight nations, and the verdict was clear and stable. Contact typically reduces prejudice, with an inverse relationship between contact and prejudice appearing in ninety-four percent of the samples. The effect was modest in size, a correlation of roughly point two one, the honest magnitude of a real social phenomenon rather than the inflated figure of a staged demonstration. What matters most for this book is the detail that distinguishes a genuine finding from a fragile one. The more rigorous the study, the larger the effect it found, the

exact reverse of the replication crisis, where the flashy effects shrank toward nothing under careful testing while this quiet one grew. The contact effect was not an artifact of selection or publication bias, which the authors tested for directly. It generalized beyond the specific people in the room to the outgroup as a whole, it appeared across an enormous range of groups and settings, and it reached well past race to prejudice against the old, the disabled, the mentally ill, and the gay. Allport's conditions were found to strengthen the effect without being strictly necessary to it, which is to say that contact helps somewhat on its own and helps a great deal when it is structured the way he prescribed.

Two features of that meta-analysis deserve more weight than the modest effect size might suggest. The first is the direction of the relationship between rigor and result, which runs opposite to the pattern that wrecked so much of the field. When a finding is an artifact of loose method, tightening the method shrinks it toward zero, as happened to power posing and ego depletion under careful replication. When a finding is real, tightening the method tends to sharpen it, and in Pettigrew and Tropp's data that is broadly what appeared, the better-controlled studies producing effects no weaker than the loose ones, the mark of a phenomenon that exists rather than a fashion that does not. The second feature is generalization. The benefit of contact spread beyond the particular individuals encountered to the outgroup as a whole, and later work found that even indirect contact, knowing that a member of one's own group is friends with a member of the other, can lower prejudice at a distance. A finding that holds up under scrutiny and reaches beyond its immediate occasion is the kind of finding a policy can be built on, and the contact hypothesis has accordingly guided the design of everything from school desegregation to the reconciliation programs that follow civil wars.

Honesty about the contact finding requires reporting the strongest challenge to it, and the challenge comes from the same researchers this book has leaned on against the diversity industry. In 2019 Elizabeth Levy Paluck, with Seth Green and Donald Green, returned to the contact literature and applied the strictest filter available, counting only studies that randomly assigned people to contact or its absence and then measured the result after a delay rather than on the spot. Out of the hundreds of studies in the field, twenty-seven met that bar. Within that rigorous core the news was mixed in a way the book is obliged to report. Contact still tended to reduce prejudice, which confirms the older finding rather than overturning it, and the conclusion that contact typically helps came through intact. Yet the effect on racial and ethnic prejudice specifically ran markedly weaker than the headline figure from the larger pool, the authors found the same hints of publication bias that shadow every literature, and the most telling feature was an absence. The rigorous studies almost never examined adults, and almost never examined adults' racial prejudice, which is the exact target the corporate and academic diversity programs claim to address. The strongest evidence for contact comes from the

young and from prejudices other than race; the place where the industry spends its money, the grown adult and his racial attitudes, is close to an evidentiary blank.

This cuts in the book’s favor rather than against it, and the point can be made without flinching. The diversity workshop is sold to adults as a cure for racial prejudice, and the best-controlled research cannot yet show that even genuine, structured, equal-status contact reliably moves that needle in that population, which leaves the staged humiliation, a far cruder instrument aimed at the same target, with no evidentiary leg whatever. The honest claim for contact is therefore narrower than an enthusiast would like. It helps, reliably, against a wide range of prejudices and across many groups, most clearly among the young and most strongly under the conditions Allport named, and it remains the best-supported tool on offer even though the rigorous evidence for shifting an adult’s racial views is thinner than anyone studying the question would wish. That is a claim built to survive the same scrutiny the earlier chapters turned on the dramatic studies, and surviving that scrutiny, rather than winning a headline, is the only standing this book has tried to earn for anything it recommends.

The benefits of contact reach past the narrowing of prejudice into territory the dramatic method never even claimed. The same body of research found that contact across a hostile divide tends to increase trust and to make forgiveness for past wrongs more possible, which is why the technique has become a staple of peacebuilding in places scarred by sectarian violence. After the worst of the conflict in Northern Ireland, programs that brought Protestant and Catholic young people into shared projects drew directly on the contact literature, on the finding that working together toward something both sides want erodes the certainty that the other side is monstrous. A few joint projects cannot heal a civil war, and to suggest otherwise would be the kind of overclaim this book exists to refuse. What the evidence shows is narrower and sturdier, that the mechanism scales from a classroom to a society without changing its character, that the same thing which makes two children stop despising each other across a jigsaw table is what begins, slowly and partially, to let two communities coexist, and that the mechanism is cooperation among equals rather than the spectacle of mutual accusation a divided society produces on its own.

The most elegant demonstration of structured contact was built in a school district in crisis rather than in a laboratory, and it offers a contrast this book cannot resist. In 1971, the schools of Austin, Texas, recently desegregated after a long history of separation, erupted into the hostility that unstructured contact predicts, with fistfights in the corridors and suspicion hardening along racial lines. A social psychologist at the University of Texas named Elliot Aronson was called in, and he made a diagnosis that located the problem in the structure rather than the souls of the children. The ordinary classroom, he saw, is built on competition, a contest for the teacher’s attention and approval in which children raise their status by putting other children down, and a

competitive room full of newly mixed groups is an engine for manufacturing exactly the contempt everyone wished to cure. His answer was to rebuild the structure around cooperation. In what he called the jigsaw classroom, the lesson was divided into pieces and each child was given one piece to master and teach to the others, so that no child could learn the whole and pass the test without depending on every other child, including the ones across the racial line he had arrived distrusting. The interdependence came as a feature of the assignment rather than a sermon about brotherhood, cooperation made structural in the way Allport had specified, and the results bore him out. Children in jigsaw classrooms came to like their group-mates more across racial and ethnic divisions, showed measurable gains in empathy and self-esteem, and the minority students improved academically without any cost to the performance of the majority.

The man who built the jigsaw classroom shared a first name with the woman who built the eye-color exercise, and the coincidence frames the choice the era faced. Confronted with the same problem, the hostility of children newly required to share a country, one teacher manufactured a cruelty and called it a lesson, and another built a cooperation and measured the result. Jane Elliott's exercise produced a vivid day and a television career and no evidence of lasting good. Elliot Aronson's classroom produced less drama and more children who learned and liked one another, in findings that have been replicated for fifty years. The difference between them is the difference this whole book has been drawing, between the method that treats the child as raw material for a demonstration and the method that treats the child as a person with work to do and something to contribute, and only one of the two has ever been shown to leave the room better than it found it.

The mechanism inside the jigsaw classroom shows, in detail, how equal status is engineered rather than wished for. In an ordinary class the low-status child, the one the others have already learned to dismiss, can be ignored at no cost, and the dismissal confirms itself. In a jigsaw group he holds a piece of the lesson the others need in order to pass the test, which converts him from someone they can scorn into someone they must listen to, and the listening, repeated over weeks, does what no instruction to be kind could accomplish. The structure makes the despised child valuable and the contemptuous child dependent, and the attitudes follow the structure. Set this beside the eye-color exercise one more time and the contrast completes itself. Both teachers wanted children to understand something about division, and each chose a lesson to engineer. Elliott's exercise taught that status is arbitrary and cruelty reversible, a lesson whose natural conclusion is cynicism, the wheel turning so that today's victim becomes tomorrow's tyrant. Aronson's classroom taught that the children needed one another, a lesson whose natural conclusion is solidarity, and the records show that the second lesson is the one that held and helped.

A word of honesty is owed about a famous study often cited alongside the jigsaw, because the standard of this book applies to its own side. The Robbers Cave experiment of the 1950s, in which the psychologist Muzafer Sherif set two groups of boys at a summer camp against each other and then reconciled them through shared tasks that required cooperation, is the usual textbook source for the power of common goals. The trouble is that Sherif’s study has the same defects the earlier chapters found in the prison and the shock machine. He ran more than one version and reported the one that worked, an earlier attempt having failed when the boys refused to turn on each other, and the experimenters shaped the conflict they claimed to observe. The honest course is to hold the demonstration at arm’s length while keeping the finding, because the mechanism it dramatized, that cooperation toward a goal both groups need will reduce the hostility between them, does not rest on Sherif’s theatrics. It rests on the meta-analytic weight of hundreds of sounder studies, on the jigsaw results, and on the workplace evidence that comes next, and a finding survives the fall of any single demonstration when the broader record carries it.

That workplace evidence is the most consequential, because it concerns the institutions that spend the most money on the dramatic method and get the least for it. The sociologists Frank Dobbin and Alexandra Kalev, working from three decades of data on more than eight hundred American firms, sorted the common diversity interventions into those that move the numbers and those that do not, and the sorting follows the logic of everything above. The failures are the coercive tools, the ones built to control and to punish. Mandatory diversity training, the compulsory workshop, tends to produce anger and resistance and, in many cases, more animosity toward other groups than before, because forcing a person to attend a session on his own bias activates the bias and provokes the rebellion that any coerced adult feels. Hiring tests and bias-policing grievance systems fare no better, for the same reason, that they treat the manager as a suspect to be constrained. The successes are the engaging tools. Programs that work apply three principles, and the three are the contact hypothesis in business clothes: engage managers in solving the problem rather than scolding them as the cause of it, expose them to women and minority colleagues through real working contact, and build a social accountability that makes fair-mindedness a thing one is seen to practice. The interventions that follow these principles are the ones that succeed, mentoring programs that pair senior people with talent they would not otherwise have met, self-managed teams that put people of different backgrounds to work as equals on a shared task, diversity task forces that give managers ownership of the outcome, and targeted recruitment that widens the pool. Mentoring alone, on their numbers, raised the representation of several underrepresented groups in management by margins between nine and twenty-five percent. The lesson is stark and it rhymes with the science. Force fails and engagement works, because a manager, like a child and like a canvassed voter,

is an agent who resists being coerced and responds to being included.

Why the coercive tools backfire and the engaging ones succeed follows the same psychology the rest of the chapter describes, applied to adults with budgets and titles. A mandatory training session casts the manager as a defendant, a suspect whose hidden bigotry must be corrected under compulsion, and an adult so addressed does what adults do when accused and ordered, resenting the charge and resisting the correction, sometimes leaving the room more hostile than he entered it. A diversity task force casts the same manager as part of the solution, a person of standing asked to help fix a problem the organization takes seriously, and that role recruits his self-respect instead of insulting it, because most people want to see themselves as fair and will act to protect that self-image once it is engaged. Mentoring and the cross-group team work by the contact mechanism directly, putting the manager in sustained, equal, cooperative contact with the people the abstraction told him to discount. The detail that ought to humble every diversity department is that some of the most effective tools were never designed for diversity at all. Self-managed teams and cross-training were built to get work done, and they improved fairness as a byproduct, which suggests that the surest route to an equal workplace may be less a program aimed at bias than a structure that puts different people to work together as equals.

There is a quieter family of methods that works by changing the process rather than the person, and it deserves its place because it sidesteps the whole question of belief. If bias enters a decision through the discretion of the person making it, then narrowing the discretion narrows the bias, and you need not have reformed anyone’s heart to have improved the outcome. The cleanest illustration is the symphony orchestra. When the economists Claudia Goldin and Cecilia Rouse studied what happened after orchestras began holding auditions behind a screen that hid the musician’s identity from the jury, they found that the screen raised the chances a woman would advance and be hired, and they estimated that this single change of procedure accounted for a substantial part of the rise in women’s presence in the major orchestras, which had stood below five percent in 1970 and climbed toward a quarter by the time they wrote. They were careful about the estimate, noting that some of it carried wide uncertainty and one part ran the other way, and that care is itself the point, the calibrated claim of an honest study rather than the sweeping promise of a demonstration. The principle generalizes, and a body of work on what one researcher called designing for equality has built on it, the insight that a structured interview beats an unstructured one, that a clear and shared set of criteria beats a vague impression, that changing the default beats exhorting people to overcome themselves. Design treats bias as a problem to be engineered around rather than a sin to be confessed, and it has the great advantage of working on the people who would never attend the workshop and would lie on the loyalty oath.

The design approach has a champion in the economist Iris Bohnet, whose work gathers the scattered findings into a single rule, that fixing the environment is easier than fixing the person. Behavioral science spent decades trying to talk people out of their biases and produced meager results, because a bias operating below awareness is not reachable by an appeal to awareness. What the environment can be made to do, the person often cannot. A hiring committee that evaluates candidates against a fixed list of criteria, agreed in advance, makes fewer biased choices than one that forms a general impression and rationalizes it afterward. A panel that compares several candidates side by side judges them more on merit than one that considers each in isolation, where stereotypes rush in to fill the gaps. Removing a name or a photograph from the early stage of a review, where it can be done without deceiving anyone, strips out the cue that triggers the bias before it fires. Each of these is a small piece of plumbing, unglamorous and effective, and together they make the case that the most reliable way to reduce discrimination is to redesign the moment of decision so the bias has less room to operate, leaving the harder question of the heart for another day, or for no day at all.

The last method is the one the struggle session most directly betrays, which is persuasion, and the strongest evidence for it arrived through a scandal. A celebrated 2014 study had claimed that a brief conversation with a gay canvasser durably changed voters’ minds on same-sex marriage, and it turned out to have been built on fabricated data, exposed by two young researchers who had tried to extend it and found the numbers impossible. Rather than abandon the question, those researchers, David Broockman and Joshua Kalla, ran the real experiment, and in 2016 they reported it in Science. Canvassers went door to door in South Florida and held a single conversation of about ten minutes with each voter, a conversation built on invitation rather than confrontation, asking the voter to recall a time he had been judged unfairly and to consider the experience of a transgender person in that light. That one exchange measurably reduced prejudice against transgender people, and the reduction was still detectable three months later, a larger shift than the average American’s decline in anti-gay feeling across more than a decade. The finding rebukes the whole premise of the dramatic method. Prejudice that the method insists can be dislodged only by shock and shame turned out to yield to a respectful conversation that treated the listener as a thinking adult, asked nothing of him but his attention, and trusted him to reach the conclusion himself.

The scandal behind that experiment ties the constructive chapter back to the destructive ones. The fabricated 2014 study had been a sensation, the apparent proof that a single gay canvasser could permanently change a stranger’s mind, and it collapsed when the two researchers trying to build on it could not reconcile its data with reality and traced the impossibility to invented numbers. They could have walked away from a burned field. Instead they treated the fraud as a question still worth answering honestly,

designed the real experiment with the safeguards the fake one lacked, and let the chips fall, which is how science is supposed to correct itself and so rarely does. What they found vindicated the method the fraud had only pretended to test. The reason the doorstep conversation works is that it does the opposite of an accusation at every turn. It declines to tell the voter he is a bigot and asks him instead to remember a time he was judged unfairly. It makes no demand that he adopt a position and invites him instead to imagine another person’s situation, drawing the conclusion out of his own reflection. The voter is treated as a competent moral reasoner who can be trusted with a hard question, and being treated that way, he reasons his way somewhere better and stays there, which is the whole case for persuasion over compulsion compressed into ten minutes at a door.

Persuasion carries a limit the other methods do not, and it should be admitted squarely, the limit of scale. A conversation that changes one mind in ten minutes cannot be broadcast to a million minds at once, and the research is clear that the mass versions, the television spot and the mailer, do little of what the doorstep does. This looks at first like a fatal weakness, until one remembers that the mass ritual scales beautifully and accomplishes nothing, so that the real choice lies between a method that reaches few and works and a method that reaches many and fails. The honest response to the limit is to take the depth seriously and propagate it deliberately, training canvassers in the technique, building organizations that hold thousands of these conversations one at a time, treating the slow accumulation of changed minds as the actual shape of persuasion rather than waiting for a shortcut that does not exist. The work that lasts has always been retail rather than wholesale, a matter of one person taking another seriously, and the discovery that this old truth can be measured and taught is among the more hopeful findings in the literature, even if it disappoints anyone hunting for a way to fix a nation by Friday.

Stand back from the five findings and a single shape emerges, the shape this chapter promised, the exact inverse of the dramatic method’s signature. What works arranges contact among equals rather than rigging a hierarchy to humiliate, the condition Allport named and Pettigrew and Tropp confirmed across a quarter of a million lives. It structures that contact into cooperation, in Aronson’s jigsaw classroom, where a child passes the test only by depending on the others. The engaging tools recruit the people they mean to change rather than scolding them, in the task forces and mentoring and cross-group teams that moved Dobbin and Kalev’s numbers when the training did not. Design routes around the bias that exhortation cannot reach, in the orchestra screen that hid the player from the jury. And persuasion replaces compulsion, in the ten minutes at a Florida doorstep that left a voter changed three months on. What works treats the person throughout as a reasoning agent with dignity and something to offer, never as the soft clay the staged exercise assumes. The methods differ in their settings, the school and the

firm and the street, but they agree in their anthropology, and the anthropology is the one the corrected science of the previous chapters insisted on, the human being who resists being pushed and responds to being respected.

Underneath the five methods lies a single claim about human nature, and it is the claim the corrected science of the earlier chapters forced upon us. The staged exercise and the experiments that licensed it assumed a soft and passive creature, a person who is whatever the situation makes him, who can be reshaped by a strong enough scene and will absorb the lesson the scene was built to deliver. The evidence says no such creature exists. A real person reasons about his situation, resists being pushed, sees through manipulation, and clings to the conclusions he has reached himself, which is why the push and the shame produce resentment while the invitation and the structure produce change. Every method that works in this chapter is built for that real person and not the imaginary clay. Contact gives him corrected evidence and lets him update. Cooperation gives him a reason to value the people he was taught to dismiss. Engagement enlists him in the fix rather than casting him as the flaw. Design routes around the bias he cannot introspect. Persuasion hands him the question and trusts his answer. The methods succeed because their model of the person is true, and the dramatic method fails because its model of the person is false, and no amount of theatrical force can make a false model come right.

If all this is true, and the evidence says it is, an obvious question presses, the question of why institutions keep reaching for the workshop and the mandate instead. The answer was given in the previous chapter and is worth restating here, because it explains the persistence of a failure in the face of a known cure. The working methods are slow, undramatic, and hard to display. A jigsaw classroom changes children over a semester and produces no footage of catharsis. A mentoring program shows its results in a promotion rate three years out. The screen at an audition prevents a bias no one will ever see prevented. A respectful conversation at a doorstep converts one voter at a time. None of these yields the cleansing spectacle, the visible token of institutional resolve, the photographable moment of moral seriousness that a frightened organization under public pressure wants to produce. The struggle session offers all of that by lunchtime. That choice, between the gesture and the work, is decided again and again by which one looks better on a Tuesday, and the working methods lose that contest precisely because they are working methods, busy with results instead of appearances.

The deeper reason the ritual keeps winning is that it offers an institution something the working methods cannot, which is absolution without change. A company that runs a mandatory training has done something visible and finite; it can point to the completed sessions, report the attendance, and assure itself and its critics that it has addressed the problem, all without altering a single structure that produces the problem. The working methods grant no such release. They require the slow rebuilding of how hiring is done

and how teams are formed and how managers are held to account, work that is never finished and never photogenic and never lets anyone declare victory. An organization frightened of looking complicit will nearly always prefer the gesture that ends the discomfort to the labor that ends the injustice, and the dramatic method is, above all, a machine for ending the discomfort of the people in charge. This is the same substitution the book has traced from the beginning, the performance of virtue standing in for its practice, and it explains why a method that does not work is so much easier to sell than the methods that do.

A danger shadows every method in this chapter, the danger that an institution will perform the cheap counterfeit of it and claim the credit due the real thing. Contact has its counterfeit in mere proximity, the seating of different people in the same building with nothing structured to make them equals or partners, which produces friction and gets called integration. Cooperation has its counterfeit in the token team assembled for a photograph and disbanded before it does any work. The accountability that makes managers fair-minded has its counterfeit in the committee that meets, files a report, and changes nothing. Each genuine method demands sustained effort, money, and the patience to wait years for a result, and each can be hollowed into a gesture that demands none of those and delivers none of the benefit. The effort is the price of the result rather than a defect of the working methods to be engineered away, the thing that distinguishes a practice that changes an institution from a ritual that only reassures it. An organization serious about the goal can tell its own counterfeits by a single test, whether it is willing to spend years and see little to show on camera, because the methods that work ask exactly that and the methods that fail promise to spare it.

Honesty requires marking the limits of the cure as plainly as the failures of the disease, and this is where the chapter parts company most sharply with the thing it replaces. None of these methods is magic, and offered as magic they would become one more overclaiming demonstration. Contact reduces prejudice on average and modestly, and the effect on entrenched racial prejudice specifically runs weaker than the headline figure, while contact gone wrong, the hostile or unequal kind, can deepen the division it was meant to heal. Structural fixes can be gamed, a screen worked around, a task force turned into a paper exercise. Persuasion is labor, one conversation and one voter at a time, and it does not scale into a slogan. Mentoring and engagement move the numbers by meaningful margins and not by miracles. The proper claim for all of it is the modest and durable one the evidence supports, that these methods help, reliably, in bounded ways, when done with care, which is the only kind of claim an honest account of human improvement can make. That method promised transformation in a morning and delivered harm. The methods in this chapter promise improvement over time and deliver it, and the difference between those two promises is the difference between a confidence trick and a science.

The honest catalog of limits runs further than a single paragraph can hold, and the book’s credibility depends on laying it out rather than hurrying past it. Contact reduces prejudice on average, which means it fails in particular cases, and the averages conceal a weaker effect on the most entrenched racial hostilities than on prejudices more lightly held. The conditions that make contact work, genuine equality and real cooperation and institutional support, are difficult to arrange and easy to counterfeit, and a desegregation that produces only proximity can leave both groups worse off. Structural fixes solve the problems they are aimed at and leave untouched the ones they are not, so a screen at an audition does nothing about who was trained to audition in the first place. None of these methods reaches the deeper machinery, the housing patterns and school-funding formulas and inherited wealth that shape who ever arrives at the door of the firm or the conservatory, and a chapter on what reduces prejudice should not be mistaken for a chapter on what produces justice, which is a larger matter requiring larger tools. The proper spirit is the one the working methods model and the dramatic method abandoned, a spirit that claims only what it can show, welcomes the correction of its own errors, and prefers a modest truth to a thrilling lie. That spirit is the real alternative this book has argued for, more than any single program, and it is the thing a frightened age most needs and least wants.

I want to push the honesty one step past where it is comfortable, to the hardest case of all, because a critic will go there if I do not. The problem the whole argument cares about most, the settled racial prejudice of an adult, is the problem for which the rigorous evidence is thinnest, close to a blank when the standard is the randomized trial rather than the hopeful study. Contact helps adults less than it helps children, persuasion is measured one trial at a time and its firmest results sit in domains other than race, and a reader is entitled to the blunt summary that no one has yet shown, to the standard a drug would have to meet, a method that reliably softens a grown person’s racial hostility. I could hide that behind the gentler findings, and I will not, because hiding it is the move the dramatic method made for decades. Here is why the admission does not sink the book. Not knowing with certainty what cures a thing is a better place to stand than knowing with certainty what worsens it, and the dramatic method is the second of those, a treatment measured and found to harm. The case in these pages rests on two findings, that the method fails, which the evidence shows, and that the honest alternatives do no damage while they help where they can, which the evidence also shows, and neither depends on the cure being sure. A doctor who cannot yet cure a disease still does right to stop prescribing the pill that deepens it and to reach instead for the regimen that helps where it can and harms nowhere. That is the whole of the claim, and it holds whether or not the cure is ever found.

Gather these findings and point them at the question the schoolteacher in Riceville was trying to answer, and a real curriculum appears, one that asks nothing she could not

have done. A teacher who wants children to understand prejudice can build the cooperative structures that made the jigsaw work, assigning the interdependent tasks that turn a classroom from a contest into a common effort. She can teach the documented history of injustice fully and honestly, the thing the concept bans would forbid, because understanding what was done requires knowing it, and knowing it differs from being made to feel guilty for it. She can lead the perspective-taking conversation the canvassing study validated, inviting a child to recall his own experience of unfairness and to imagine another's, rather than assigning him a caste and a feeling. None of this manufactures a cruelty, stages a humiliation, or compels a confession, and all of it has evidence behind it that the eye-color exercise never had. The alternative to the bad lesson was always a better lesson, the one the research had been quietly assembling in the decades the dramatic method spent on the stage, available to any teacher willing to trade a vivid morning for a durable result.

One more thing should be said before the turn, about the goal these methods serve. The people who built the jigsaw classroom, ran the canvassing experiment, and sorted the corporate data were friends of equality, not skeptics hunting for an excuse to do nothing, and they were trying to find what would move the goal rather than merely perform devotion to it. What they found is that the end the dramatic method claims and betrays is reachable by other means, slower and surer and kinder than the staged confession. To choose contact over the struggle session, structure over the loyalty oath, and persuasion over the compelled creed is to choose the goal over the gesture, the result over the performance, the person over the spectacle made of him. That choice is open to any institution and any teacher willing to make it, and the evidence that it works has been sitting in the journals for half a century, waiting for the frightened to grow brave enough to prefer a quiet success to a loud one.

That difference is the note on which to leave the argument and turn back toward where the book began. The cure for the pedagogy of the oppressor is the patient refusal of ritual altogether rather than a better ritual of humiliation, the slow and respectful work of putting people of different kinds together as equals, giving them something worth doing in common, building the structures that make fairness the path of least resistance, and trusting them with the honest conversation that treats them as the reasoning adults they are. It is harder than a staged cruelty and it photographs worse, and it is the only thing that has ever been shown to work. What remains is to return to the schoolteacher in the small Iowa town who started all of this, and to the words an honest observer of her exercise might finally say, the words that give the last chapter its title.

CHAPTER SIXTEEN

Not My Thing

Fifty-three years later, the room is the clearest thing I own. I can still see the construction-paper band on my arm and the single letter printed on it, still feel the particular heat of being eight years old and sorted and watched by people who were sure they were doing me good. The book you have nearly finished began in that room, with a paper armband and an adult who had mistaken cruelty for instruction, and it has traveled a long way since, through the famous original in Iowa and the family of imitators, through Freire and Illich, through the collapse of the experiments and the politics of a frightened decade and the quiet methods that work. It ends where it started, because the answer was in the room the whole time, smaller than any of the arguments I have spent fifteen chapters assembling, contained in three words an eight-year-old said to his teacher when she told him to pick up the whip. Not my thing.

I had no theory then, only a refusal. When the roles reversed on the second day and the teacher pressed me to give the orders the brown bands had given me, to feel for myself the pleasure of command she was certain would complete the lesson, I declined, and the words I found were the only ones I had, that it was not for me, not my thing. There was no argument beneath them, only the instinct that what she was offering was not a thing I wanted to become. The whole of this book is the argument I could not make at eight, the brief filed half a century late on behalf of a child’s plain refusal. Everything in it, the inversion and the iatrogenic harm and the absent subject and the evidence and the history, is an elaboration of three words a third grader already knew were true.

I should tell the day plainly, because the memory has the grain of something that happened rather than something I have read about myself. We were lined up and sorted without explanation, sent to one side of the room or the other the way the rules of a rigged game are always withheld from the people about to lose it. A brown-eyed boy, a friend the week before, ordered me to fetch him water, and when I stayed in my seat he told me he would have to whip me, and I told him to try. He went to the teacher, who announced that I would be punished with the shame of history, a phrase that meant nothing to me at eight and means too much to me now. Around me the other blue bands

brayed like donkeys on command and ground through forced jumping jacks while the older grades looked on, because the thing did not stay in our room. It followed us into the cafeteria and onto the playground, where children who had never met us read our rank off our arms. I have spent a lifetime learning that the cruelty was the point and the audience was the design, and the body learned it first, in a single morning, with a paper band cutting slightly into a thin arm.

The sentence the adults kept repeating, the one I have never been able to put down, was that now we knew how it felt to be Black. There were one or two Black children in the entire school and none in that room, which was the condition the lesson required and never confessed, because a single Black child standing among us would have shattered the pretense in a second, the game ceasing to be a game for the one person present to whom it was not theoretical. The presumption in that sentence still astonishes me. A white teacher in an all-white room told a circle of white eight-year-olds that a morning with a colored armband had handed them the knowledge of a life they would walk out of at three o’clock and never reenter. It was a theft dressed as empathy, the conversion of someone else’s unchosen and unending experience into a costume the comfortable could put on for a lesson and take off for lunch, and the person whose life was being borrowed was not in the room to say whether the loan was fair.

The teacher’s phrase has aged into an irony she could not have intended. She told me I would be punished with the shame of history, meaning the weight of past wrongs, and the exercise she was defending has itself become a piece of history now, a documented episode a later generation studies the way I am studying it, as a specimen of how good intentions go wrong. There is a deeper irony underneath that one. The thing the method always substitutes for is real history, the honest account of what was done to actual people, slavery and segregation and their long machinery, taught plainly enough that a student understands it. That history needs no armband and no manufactured shame. It needs only to be told, and the telling is the thing the exercise skips on its way to the performance, and the thing, in our own decade, that the concept bans would forbid a teacher to tell at all. Between a method that fakes the feeling of history and a politics that would erase the facts of it, the actual past keeps getting lost, and the student is left with everything except the knowledge that would have done him good.

The refusal cost me, in the small currency a school deals in. I was the only child in the room punished for any of it, singled out for declining what everyone else performed. My mother taught down the hall, and my refusal embarrassed her in front of the colleagues whose lesson I had spoiled, so she took the television away for a week to teach me what losing a privilege felt like, never seeing that the stripping of a privilege was the engine of the exercise she was punishing me for resisting. The cost was real and it was tiny, a week without a screen, and the machine runs on costs of exactly that size, calibrated to be just large enough to buy the compliance of a child who would rather

keep his television than keep his refusal. I kept the refusal. With the set gone I read instead, and the freedom I found in those books has never run out, which is the first thing the punishment taught me, that the price of refusing is smaller than it looks and the reward is larger.

I want to be fair to my mother, because she was not a cruel woman and the bind she stood in was real. She taught fourth grade in that same building, among the colleagues who had designed and defended the exercise, and her son had stood up in the middle of it and refused, which embarrassed her in the small and total way that a professional is embarrassed before her peers. What she did next she did out of the only playbook she had. She took a privilege away to teach me the cost of stepping out of line, and she could not see, because the frame would not let her see it, that she had just run a miniature of the same exercise on her own child at her own kitchen table, sorting me into the role of the punished and stripping a privilege to make a point. She meant it as discipline and love, and it was both. The tragedy of the method is that it recruits decent people to administer it, and my mother was one of them, a good teacher carrying out a bad lesson because the people she trusted had told her it was the right one, which is the most ordinary way that harm gets done, by kind hands that have been handed the wrong tool.

There were two refusals, and the second matters more than the first. On the first day I would not obey, which is the refusal everyone understands, the marked child declining to grovel. On the second day I would not command, and that is the rarer thing. The method’s cleverest trap is the whip rather than the lash, the offered pleasure of the top rather than the humiliation of the bottom, because most people will resist being degraded and far fewer will refuse the invitation to degrade. The teacher was sure the lesson would land the moment I tasted power, and the whole genre depends on that certainty, on the favored child accepting the role and discovering that he enjoys it. Declining the whip is the harder refusal and the more important one, because it is the refusal of the corruption the apparatus exists to perform, the conversion of an ordinary child into a willing instrument of a cruelty he has been told is just.

I want to be precise about what the refusal was and was not, because the word can be cheapened into mere contrarianism, the reflex of the person who says no to everything to feel like somebody. That was not it. None of that was a refusal to learn that racism is real, which it is, or that it is cruel, which it is, or that a child should grow up to oppose it, which I have. The refusal was aimed at the method, the specific demand that I perform a feeling on command and enact a cruelty to prove I understood one, and the distinction is the whole hinge of this book. A person can hold the truth about injustice with his entire heart and still decline to be marched through a degradation ceremony in its name, because the ceremony does not teach the truth and was never the only way to learn it. Refusing the rigged game is not the same as refusing the lesson the game pretends to carry, and the proof is that I left that room believing more firmly, not less, that ranking

people by an accident of birth is a lie, having felt in my own arm exactly how arbitrary and how enforced such a ranking is. The refusal kept the truth and discarded the cruelty, which is the only sorting worth doing.

There is a harder version of that objection, and I would rather raise it than have it raised against me. A skeptic can grant everything just said and still ask whether what I have dressed as conscience was only temperament, the stubbornness of a particular boy who did not like being told what to feel, and whether a man who grew up to write books has spent a lifetime flattering a childhood streak by calling it principle. I cannot fully answer that, and I have stopped trying, because the honest truth is that the conscience and the stubbornness do not come apart cleanly, and at eight they did not come apart at all. The argument does not need them separated. What it needs from the refusal is small. It needs the refusal to have been possible, a child handed the armband able to decline it, and nothing more, not nobility and not rarity and not a finer character. I am also aware that I am the one who lived to tell it, the survivor narrating his own small stand while the children who complied are not here to narrate theirs. My refusal is offered here for one purpose only, as evidence that the ritual can be refused at all, and the most stubborn child in the room supplies that evidence as well as the most virtuous would.

The small no has a long and honorable lineage, unknown to me at eight. It is the same act, scaled down to a classroom, that a court named when it ruled that no official may force a schoolchild to salute a flag against his conscience, the principle that the state cannot compel a citizen to confess a belief by word or gesture. It is the act of the conscientious objector, the dissenter, the person who declines the loyalty oath, the one who will not sign. The content varies and the stakes vary, from a paper armband to a prison sentence, but the structure holds constant, a single person refusing to let an authority use his mouth or his hands to perform a belief the authority has chosen for him. That refusal is the atom of conscience, the smallest indivisible unit of a free person, and every larger freedom is built out of it. A society keeps its liberty to exactly the degree that it keeps producing people who can say no to a ritual, and loses its liberty when the rituals grow large enough and frightening enough that the supply of such people runs dry.

I have wondered for years about the other children, the ones who stayed in the game, and the wondering carries no judgment of them. They were eight. They did what eight-year-olds do when an adored teacher gives them the rules, which is to follow the rules and try to win the game in front of them. The blue bands who brayed and dropped into jumping jacks were obedient children rather than bad ones, which is what the room was built to produce. The brown bands who gave the orders were no budding tyrants, only children handed a script that rewarded a little cruelty and praised them for performing it. That is the quiet horror of the thing, that it works on the good and the ordinary, that it can take a roomful of decent kids and, for an afternoon, sort them into

the obedient and the imperious and call the result a lesson. I do not think most of them remember it, and I am not sure that is a mercy. The exercise taught them, briefly and without their knowledge, how readily they would wear whichever band they were handed, which is a thing about themselves they were better off learning some gentler way, or not learning like that at all.

A fair objection deserves an answer here, the objection that moral education is supposed to be uncomfortable, that you cannot teach a child about cruelty without disturbing him, and that my refusal sounds like a demand to be kept comfortable. The objection misses a distinction the whole book turns on. Discomfort that arises when a person meets a hard truth is the honest friction of learning, and no one should be spared it; the discomfort of reading what was done in the slave markets, of sitting with the facts of a lynching, of facing one's own small complicities, is the price of an education and worth paying. The discomfort the exercise inflicts is a different thing entirely, a suffering manufactured by an authority and assigned by category and staged for an audience, pain as performance rather than pain as the cost of knowledge. I am not asking to be spared the truth. I am refusing to be conscripted into a spectacle. The honest teacher disturbs a student by showing him something real; the coercive one disturbs him by doing something to him, and the difference between being shown and being processed is the difference between an education and an ordeal.

A coercive ritual cannot run on people who decline it. It needs participants, the marked who submit and the favored who enforce and the audience that watches and believes, and a single refusal anywhere in that machinery jams it. The child who will not command has stepped out of the role written for him and become something the exercise cannot use, a present and unsilenced person where the design called for compliant material. This is why the refusal is the thing every version of the method most works to prevent, with its pressure and its punishment and its insistence that opting out is itself a kind of failure. The performance requires a full cast. One person who folds his arms and says the words breaks the spell for everyone in the room, because he proves in front of them that the thing can be refused at all.

The deepest reason a refusal is dangerous to the ritual returns to the thing my own classroom taught, that the exercise runs only in sealed conditions, with its true subject absent or silenced. My room worked because no Black child was in it to say the game was not a game. The whole genre depends on that kind of sealing, on a controlled space where the script holds because no one inside it will break the frame. A refusal breaks the frame from within. The child who folds his arms is a subject who will not stay silent, a piece of reality that has gotten into the sealed room and will not behave, and his mere presence, unbought and unbowed, reintroduces the one thing the exercise was built to exclude, a person who is not playing the assigned part. This is why the pressure to participate falls heaviest on the one who hesitates, and why opting out is treated as an

offense rather than a neutral choice. The seal must hold, and a single refusal is a crack in it, and the people running the ritual sense, even when they will not say it, that cracks are how sealed things break.

The woman who built the original is alive as I write this, in her nineties, and still publicly defends the exercise she first ran the morning after King was killed. She defends it the way a mother defends a child, granting that the day was hard and the children felt betrayed and insisting that the rest of the year redeemed it, without conceding that the method itself might be the flaw. Her sincerity has never been in doubt. She had a genuine awakening, she still speaks of King's death with undimmed force after nearly sixty years, and the goal she names is one I share, that there is a single human race and the rankings we build on top of it are lies. What she has not said, in fifty-eight years of public defense, is the thing the eight-year-old said without effort. She has not said that it was not the way. To concede it would mean losing the child she has defended for a lifetime, and so she defends the method against the evidence in exactly the manner the method trains everyone to defend it, as an article of faith immune to result.

I find I cannot be cruel about her, because the trap she is in is a human one and I can feel its jaws. Imagine having built your life's meaning on a single morning in 1968, having carried it into a thousand rooms and onto television and into the books written about you, having been told for half a century by admirers that you are a hero of conscience. Imagine then being asked, in your nineties, to consider that the morning was a mistake, that the method you gave your life to was the wrong tool for the right job, that the children were harmed more than helped. Almost no one could say yes to that, and the cost of saying it climbs with every passing year and every lecture, until at last the admission would be less a correction than a self-erasure. So she defends it, and I understand why, and I do not for that reason excuse it, because the defense has a cost paid by other people, the children still being marched through the exercise in classrooms that take her word for its worth. Sympathy for the person and judgment of the method sit together without conflict. I can grieve the trap and still say the thing she has not, that it was not the way, the same three syllables a child found in an afternoon.

It helps to remember why the original spread so far, because the reason indicts the method twice over. The exercise became famous because it was filmed. A documentary crew captured a session in 1970, and the footage of small children turning on one another at a teacher's word was riveting the way a car crash is riveting, and public television carried it, and a network revisited it years later when the original students gathered as adults to testify to its power. The camera made the career. What the camera loved was the drama, the tears and the cruelty and the visible transformation, the same theatrical charge this book has identified as the disease rather than the cure. A quiet method that works produces nothing worth filming, and a dramatic method that fails

produces an unforgettable hour of television, and the second got copied into a thousand classrooms while the first sat in the journals, which is the whole tragedy of the genre compressed into a question of what makes good footage. The exercise was sold by the same quality that made it harmful, its capacity to manufacture a spectacle, and a culture that mistakes spectacle for proof bought it by the million.

My small classroom turned out to hold the whole species, every argument this book makes already present in miniature. The inversion was there in the rigged game itself, the assignment of a caste by the color of an eye, enforced by an authority I could not appeal, which is the oppressor's structure whatever banner it marches under. Iatrogenic harm, the cure that injures, was there in the result, because the exercise meant to teach my town empathy produced instead a father's slur and a hallway of older children treating every labeled third grader as prey. The absence of the subject was there in the precondition no one admitted, a room that needed no Black child in it and a teacher standing guard to keep the fiction intact, a lesson about people that required those people to be missing. Studying a thousand cases was never necessary to find the pattern. I had worn it on my arm.

Set the schoolteacher's lifetime beside the child's afternoon and the contrast holds a quiet lesson about what certainty costs. She has had fifty-eight years and a thousand audiences and every honor a movement can bestow, and in all that time she has gone without ever saying three plain syllables about her own work. I had an afternoon and a confiscated television, and the syllables came to me before I understood them. The difference is neither intelligence nor virtue; she is formidable and sincere and I was a frightened third grader. The difference is that I had nothing invested in the exercise and she had everything, and a position defended long enough and publicly enough becomes a cage its holder can no longer leave. There is a warning in that for all of us, for me as much as for her, that the surest way to lose the ability to say it was not the way is to spend years insisting that it was. The child could refuse because he was free of the thing. The adult cannot concede because she is captive to it, and the captivity is self-built, one defense at a time.

I held this for half a century as a private grievance before I understood it was anything more, the way a person carries an old wound that no one else remembers and that seems to belong to him alone. What changed was the recognition rather than the wound, late and sudden, that my single classroom had been one specimen of a species, that the paper armband in Lincoln was a copy of an exercise in Iowa that was itself the seed of a method which spread into a thousand schools and then into the training rooms of corporations and governments and the politics of a whole country. The grievance turned out to be evidence. The thing done to me had been done, in a hundred forms, to millions, and is being done still, and the private ache I had nursed so long was a small clean sample of a large and respectable-looking error. That is why I have finally written

it down, less to settle a personal score, which would not be worth a book, than because the specimen in my memory turned out to explain the species, and because the species is loose again in a frightened time that has forgotten what my eight-year-old self knew in an afternoon.

The honest place to stand on all of this is uncomfortable, and my own family marks out its edges. My mother defended the method out of decency and embarrassment, a good woman enforcing a bad lesson because her colleagues had staked themselves on it. The parents who finally ended it defended their children out of the ugliest motive available, the refusal to let white sons and daughters feel for a single day what they were content to let others feel for a lifetime. Neither the method's friends nor its enemies come out of the story clean, which is the book in a sentence. I am for the goal my mother served and against the method she defended, and I hold that second position without joining the people who use the method's failures as a weapon against the goal. The boy in the room had the stance exactly right before he had the words for it. Refuse the rigged game and keep the aim it counterfeits.

There is a straight line from that confiscated television to the page you are reading, and I have come to see it as the only justice the episode ever produced. Sent to the books because the screen was gone, I found in them the one room no authority could sort me in, where the only rank was between a true sentence and a false one and I got to decide which was which. I became a reader that week in a way I had not been before, and the reader became, in time, a writer, a person whose whole work is the placing of one honest word after another in defiance of whoever would prefer a useful lie. This book is the late fruit of a small punishment. The teacher who tried to teach me obedience to a rigged hierarchy taught me instead, by way of my mother's discipline, that the inside of my own mind was a place no rigged hierarchy could enter, and I have been writing from inside that room ever since. She would not count it a success. I count it the only one the exercise ever gave me.

The refusal became, over a life, a part of who I am, and I will not pretend that has been simple. To be the person who says it is not my thing is to be, often enough, the person outside the circle, the one who would not sign, the colleague who declined the exercise everyone else found moving, and there is a loneliness in that I have felt in many rooms since the first one. The instinct that served an eight-year-old has cost the adult invitations and approvals and the easy warmth of agreeing with the people around him. I would not trade it. The same refusal that left me outside the circle left me in possession of myself, free to write a sentence I believe instead of one I was assigned, and that possession has been the ground of every worthwhile thing I have done. A person learns, if he is lucky, that the approval he loses by refusing was never worth what it cost to keep, and that the self he protects by refusing is the only thing he was ever going to get to keep anyway. The boy paid a week of television. The man has paid more, and the

price has never once exceeded the worth of the thing it bought.

I will not pretend the refusal is a revolution, because mine was not. My folded arms did not stop the second day. The exercise ran on schedule the next morning and ran over me. The thing that ended it in my town was the parents' fury, and their fury was poisoned at the root. One child saying no does not bring a machine down, and a book that has spent its length refusing to overclaim will not start in its final pages. The refusal is the seed and not the harvest, the proof and not the cure. What it demonstrates, to the refuser and to everyone watching, is the single fact the method most needs hidden, that the ritual is refusable, that consent can be withheld, that the person in the armband was never the clay the exercise required him to be.

I have thought about why those three flat words were the right ones, and I think it is their plainness that matters. They are not a slogan or a manifesto or a denunciation, nothing a movement could chant or a critic could quote, only the smallest possible refusal, the verbal shrug of a person declining to be drafted, and their smallness is their power, because they make no claim to heroism and demand no audience and start no fight. A child can say them. A tired adult can say them in a meeting without making a scene. They travel where grander resistance cannot, into the ordinary rooms where the armbands are handed out, and they do the one necessary thing, which is to withhold the yes the ritual cannot proceed without. Not my thing. It is the most modest sentence in the language, and it has stopped more cruelty than all the thunderous ones combined, because it is the one a frightened person can bring himself to say.

So the book ends where it began, in a small room in the spring of 1973, with a child who looked at the adult in charge and declined the rigged game, whatever it cost him. That refusal is the thing the entire machine, in every form it has worn from an Iowa classroom to a corporate seminar to a government's loyalty oath, is built to prevent. It is also the only thing that has ever stopped it. And it asks nothing rare of anyone, no courage beyond the reach of an ordinary person on an ordinary day, only the willingness to look at whoever is handing out the armbands and say the three words a frightened eight-year-old found when he had nothing else to give. Not my thing.

CODA

After the Armband

An indictment closes by naming the crime, and I have named it across sixteen chapters, the manufacture of empathy through manufactured suffering, the pedagogy of the oppressor dressed in the clothes of the oppressed. A book that ended on the indictment alone would have done only half its work, because tearing down a bad method is the start of an obligation rather than the discharge of one. The question that survives the last chapter is the one a reader is right to press. If not the armband, then what, and the honest answer comes in two parts, a refusal and a labor, and the book fails if you carry away only the first.

I press the two halves together because each fails without the other, and the failures are mirror images. A refusal with no labor behind it curdles into a posture, the satisfied no of a person who has declined the workshop and gone home pleased with himself, having confused keeping his hands clean with doing some good. A labor pursued through the bad method is the error this whole book has anatomized, the sincere worker reaching for the struggle session because it feels like action and yields a result he can show. The first leaves the injustice untouched and feels righteous about it; the second attacks the injustice with a tool that deepens it. Only the two together, the no to the rigged game and the yes to the real work, add up to a person who is both free and useful, which is the only kind of person worth being in a time like this one.

The refusal you already have, because it costs nothing but nerve and is available in any room. When the next ritual is handed to you, and it will be, as a session you must attend or a statement you must sign or a creed you must perform to keep your standing, the free act is the oldest one, to decline the rigged game, to refuse both the grovel and the whip. I have spent a book defending that small word and will not take any of it back.

The labor is the harder half, and it is the half a frightened decade would rather skip. Having refused the counterfeit, you still owe the world the thing the counterfeit was a forgery of, the real reduction of real cruelty, the slow and unglamorous work the chapter on what works laid out in detail. You do not get to refuse the struggle session and call

yourself finished. The goal the method betrayed is real, and it waits for you on the far side of your refusal, asking to be served by means that respect the people you serve it among. Refusing a bad tool is not the same as abandoning the job, and a person who refuses the armband and then does nothing has kept his hands clean and left the work undone.

Two readers are owed a sentence before the end, because the present argument is loud with people ready to misread it in opposite directions. The aim the method counterfeited, the dismantling of the cruelties that rank human beings by the accident of their birth, is not in question anywhere in these pages, and that apparatus is no abstraction to me; it does measurable harm to actual lives, and tearing it down is worth a life's work. To the reader who has taken sixteen chapters of criticism as permission to abandon the goal, who reads an attack on the method as cover for an attack on equality itself: the failures of a teaching method are no argument against the thing it failed to teach, and to use the book that way is to do with it what the bad teachers did with their good intentions, turn a tool toward the opposite of its purpose. To the reader on the other side, who defends the method because the goal is sacred and any criticism feels like treason: conceding that the method fails is the precondition of ever reaching the goal rather than a surrender to the method's enemies, because you cannot arrive somewhere on a vehicle that carries you backward. The goal deserves better than the struggle session; it deserves the things that work.

The hardest version of the refusal, and the one this frightened decade most needs, is the refusal aimed at your own side. Folding your arms against the armband the other tribe hands out is easy, and refusing the loyalty oath your enemies administer feels righteous in the doing. The discipline is to refuse the ritual when it is your own people running it, when the compelled belief is one you happen to hold and the degradation is aimed at someone you are glad to see degraded. That is the moment the refusal costs something real, the social cost of breaking ranks with the people whose approval you want, and it is the moment that proves whether you love the principle or only your team. A person who will refuse only the other side's coercion has refused nothing at all. He has merely chosen a side, and the armband he declines to wear he is happy to pin on someone else the moment his faction holds the pins. The three words mean nothing if they are said only to opponents. They count when you say them to your friends. This asks no heroism, and I will not dress it as such. It is the smallest and oldest act of freedom there is, the withholding of consent, and it is available in any room on any day to anyone willing to pay its modest price.

The whole of it comes down to a sentence a frightened child said because he had nothing else to say, and I give it to you now as the only instruction this book contains. When they hand you the armband, hand it back. Refuse the rigged game, and then go and do the real work the game was pretending to do. That is the beginning and the end

of what I know. Not my thing, said first, and then the harder words that have to follow it, our work now, the slow and ordinary and unglamorous building of a fairer place by means that leave everyone in the room still standing, still themselves, still free.

APPENDIX

Where to Start

The book's argument is negative for most of its length, a case against a method, and a reader who accepts it is owed a positive answer in a form he can act on. What follows gathers, by the place a reader is most likely to stand, what the evidence assembled in these pages supports doing. It resists the shape of a checklist, because the methods that work are practices and dispositions rather than steps, and a numbered procedure would falsify them the way a recipe falsifies a craft. It also resists the promise the dramatic method made, the promise of a result by Friday. Everything here is slow, asks more than a ritual asks, and delivers a gain that arrives in years rather than in an afternoon. The one thing it offers in exchange is that, unlike the staged cruelty, it works.

For the teacher, and the parent. The lesson that works is built on cooperation rather than contest. Divide the material so that each child holds a piece the others need and cannot pass the test without, the structure Elliot Aronson called the jigsaw classroom, which made newly desegregated children in Austin come to like one another across the lines they had arrived distrusting, the minority students gaining academically and the majority losing nothing. Teach the documented history of injustice fully and plainly, because understanding what was done requires knowing it, and knowing it is a different act from being made to feel guilty for it. When the talk turns to another group's experience, lead the perspective-taking conversation the canvassing research validated, inviting a child to recall a time he was treated unfairly and to imagine another's, rather than assigning him a caste and a feeling and watching him perform it. None of this stages a humiliation, compels a confession, or manufactures a cruelty, and all of it rests on evidence the eye-color exercise never had. The rule beneath the practice is plain: engineer the structure that makes children need one another, and trust the attitudes to follow the structure, because they do.

For the manager or employer. The interventions that move the numbers, in the three decades of firm data Frank Dobbin and Alexandra Kalev examined, are the ones that engage rather than coerce. Drop the mandatory training, which casts the manager as

a suspect and provokes the resentment any accused adult feels, and reach instead for the tools that recruit his self-respect, a diversity task force that gives him ownership of the outcome, mentoring that pairs senior people with talent they would not otherwise have met, self-managed and cross-group teams that put different people to work as equals on a shared task, and targeted recruitment that widens the pool. Mentoring alone, on their figures, raised the representation of several underrepresented groups in management by margins between nine and twenty-five percent. Then change the process where bias enters through discretion, the lesson of the orchestra screen that raised a woman’s chances once the jury could not see who was playing, the structured interview that beats the general impression, the criteria fixed and agreed in advance, the side-by-side comparison that leaves a stereotype less room than the candidate considered alone. Iris Bohnet’s rule holds across all of it, that fixing the environment is easier than fixing the person, because a bias operating below awareness cannot be reached by an appeal to awareness. Measure what you have changed by who now gets hired and promoted. Attendance at a session is the certificate the earlier chapters warned against, and only the hiring shows that anything moved.

For the citizen and the neighbor. The method here is the one the struggle session most directly betrays, the respectful conversation, and its strongest evidence came from a doorstep. When David Broockman and Joshua Kalla sent canvassers through South Florida to talk with voters for about ten minutes each, the exchange that worked declined to call anyone a bigot and asked instead that he recall a time he had been judged unfairly and consider a transgender person’s experience in that light, and the prejudice it reduced was still lower three months on. Do that. Ask the person to remember his own experience of unfairness rather than informing him of his sins, invite him to imagine another’s situation rather than demanding he adopt your conclusion, and trust him to reason his way somewhere better, because treated as a competent adult he tends to. The work is retail rather than wholesale, one conversation at a time, and it will not scale into a slogan or a broadcast, which is the price of its being the kind of thing that changes a mind and holds.

The common test. One question sorts the real version of these methods from the cheap counterfeit that claims their credit. The genuine article asks for sustained effort, money, and the patience to wait years for a result that will never photograph well; the counterfeit offers proximity in place of structured contact, the team assembled for a photograph in place of the team that does the work, the committee that files a report in place of the accountability that changes a decision. An institution serious about the goal can know its own counterfeits by asking whether it is willing to spend years and show little on camera, because the methods that work ask exactly that, and the methods that fail promise to spare it. The spirit underneath all of them is the one the dramatic method abandoned, a respect for the person as a reasoning agent who resists being pushed and

responds to being included, and it is that respect, more than any single program, that the evidence in this book finally recommends.

Acknowledgments

A book of this kind is built on the labor of others, and the debts are mostly to the dead and the documented. It owes its evidence to the researchers who ran the studies and reported them honestly, including the ones whose findings undercut their own hopes. The cases come from the journalists and court reporters who wrote down what happened in classrooms and training rooms when no one came out looking good. The argument owes most to Paulo Freire and Ivan Illich, who supplied the two ideas the book leans on hardest, and who would likely have disagreed with some of what I have done with them.

I owe a particular debt to the witnesses, the former students and the parents and the teachers who went on the record about exercises that flattered no one involved, and to the scholars who returned to the famous experiments and tested whether they were true. The book could not stand without that work.

The errors that remain are mine. So is the argument. I wrote this book the way the book asks everyone to act, by my own hand and on my own responsibility, and I am content to be judged on it.

A Note on Method and Sources

A book that accuses a popular practice of doing harm takes on a burden of proof, and I have tried to carry it honestly. The rule throughout was simple. A claim that an exercise hurts people has to be shown, by evidence a skeptic could check, rather than felt or asserted because it suits the argument. Where the record was thin, I have said so. Where I reconstruct my own classroom, I rely on memory and mark it as memory.

The cases come from the public record. The mock auctions, the Gestapo role-plays, the scared-straight tours, and the diversity-training scandals are documented in news reporting, school-board minutes, and court filings, and I have tried to date and locate each one rather than pass along a rumor. For the famous exercises I went to the primary material where I could reach it, the documentaries, the contemporaneous coverage, and the biographies, and I have preferred a named source on the record to an anonymous one.

For the empirical claims I leaned on the strongest evidence available, which means systematic reviews and meta-analyses ahead of single striking studies. One dramatic experiment makes a good story and a poor proof, and the history told in these pages is in part the history of dramatic experiments that did not hold up. When the better-built research and the memorable demonstration disagreed, I sided with the research, even where the demonstration was the more famous of the two.

A word on fairness. This book is hard on a method and loyal to a goal, and I have worked to keep the two apart on every page. Attacking the way a thing is taught is a separate act from opposing what it hopes to achieve, and a reader who comes away thinking I have argued against equality has misread me, in some cases on purpose. I have tried to cite no source I believe to be arguing in bad faith, on either side of the present quarrel, and to grant the people I criticize the most generous reading their own words will bear. The goal of a society without caste is one I share. The argument is only that you cannot reach it by building caste in a classroom.

Notes on Sources

This section points to the sources behind the book's load-bearing claims, so that a reader who wants to check the record can find it. Full publication details for the works named here appear in the Bibliography that follows.

The account of the eye-color exercise rests on the contemporaneous record and the primary footage. The children's own compositions ran in the *Riceville Recorder* in April of 1968. Footage was filmed for the 1970 documentary *The Eye of the Storm* and revisited in the 1985 *Frontline* film *A Class Divided*, and the fullest narrative treatments are William Peters's two books and Stephen Bloom's later biography. That its creator still defends the exercise is a matter of the public record as well, in her continuing lectures and interviews through 2025, in which she has again described it as an exercise in empathy. The simulation game treated in the chapter on *StarPower* follows its designer's own published account from 1969.

The classroom scandals are drawn from the public record, and I have dated and located each rather than pass along a rumor. In Bronxville, a private school there ran mock auctions in two fifth-grade classes in March of 2019; the New York Attorney General investigated, found that the re-enactments had seriously harmed the students, and reported that the school had dismissed the teacher. The Watertown episode, at a Watertown, New York elementary school in the spring of 2019, had a fourth-grade teacher direct white pupils to bid on their Black classmates and instruct the children sold to take the buyer's surname and to call him master; the teacher was placed on leave, and the family's account was carried by regional and national news. Washington's case, in a Capitol Hill elementary school in December of 2021, had a staff member direct a third-grade library class to act out scenes from the Holocaust, cast a Jewish child as Hitler, and, by a parent's account reported at the time, tell the children not to tell anyone; the district called it an unapproved lesson and placed the staff member on leave, and the episode was first reported by the Washington Post and widely covered. The earlier auctions near Norfolk and Columbus in 2011, the auction in Maplewood in 2017, and the multi-day Nazi-and-Jew simulation in a South Carolina classroom in 2011 are documented in contemporaneous local and regional reporting. Throughout, I have used

the public description of what happened and left the individual teachers unnamed, since the argument is about the practice rather than the person.

On the question of simulating atrocity, the position attributed to the United States Holocaust Memorial Museum is taken from its published guidelines for teachers, which advise against simulation exercises on the ground that they leave students with a false sense of having understood the event. For the scared-straight programs, the verdict follows the Campbell systematic review by Petrosino and colleagues, which found that the programs increased rather than reduced offending. The reading of drug-resistance education follows the meta-analytic literature on the original DARE curriculum, beginning with Ennett and colleagues, and the federal reviews that came after. Diversity training is assessed through the work of Frank Dobbin and Alexandra Kalev, whose analysis of three decades of data from more than eight hundred firms is reported in their 2016 article and their later book, through the 2016 task-force report of the Equal Employment Opportunity Commission, and through the reviews of prejudice-reduction and implicit-bias research by Elizabeth Levy Paluck and Donald Green and their colleagues. The estimate of the industry’s size is the author’s own, drawn from published market figures that vary by source and offered as an order of magnitude rather than a precise sum.

The legal and policy history can be checked against the documents themselves. A citizen’s right not to be compelled to a profession of belief was settled in West Virginia State Board of Education versus Barnette in 1943. The loyalty-oath fights of the early 1950s, including the reinstatement of the non-signing University of California faculty, ran through the California courts and were made final in 1952. In September of 2020 the first federal order restricting workplace training on race and sex was issued, and it was revoked the following January; the two orders that revived the approach were issued in January of 2025; and the litigation over them reached the Fourth Circuit, which in February of 2026 set aside the nationwide injunction and held a challenge to the orders on their face unlikely to succeed, while leaving narrower challenges open. The count of state and local measures restricting instruction on race comes from the law-school tracking project that compiled them through the end of 2024.

Reassessment of the Stanford Prison Experiment follows the archival work of Thibault Le Texier, published in 2018 as a book and in 2019 in the journal *American Psychologist*, together with the broader reckoning over the reliability of the famous mid-century social-psychology demonstrations. The book’s framing ideas come from Paulo Freire’s *Pedagogy of the Oppressed*, first published in 1968, and from Ivan Illich, whose *Medical Nemesis* of 1975 supplied the language of iatrogenesis and counterproductivity and whose *Deschooling Society* preceded it. Where the book reconstructs the author’s own grade-school morning, it relies on memory and says so, and nothing in the argument depends on it.

List of Terms

A glossary of the central terms this book uses, several of them coined or repurposed in its pages. The definitions are short by design; most are argued at greater length in the chapters where they appear.

Coercive moral pedagogy. This is the practice the book names and indicts: the attempt to teach empathy or virtue by forcing a person through manufactured suffering, ranking, or humiliation. The family includes the eye-color exercise, the privilege walk, the poverty simulation, the mock auction, and the mandatory sensitivity training.

The dramatic method. A shared technique beneath that family, staging a small theater of injustice and casting the participants in it, on the theory that living a degradation for an afternoon teaches what explanation cannot.

The five-part anatomy. The structure common to every member of the genre, set out in Chapter Three and named here in its five parts. First, a *caste by accident*: rank assigned by a trait the participant did not choose and cannot change for the duration, an eye color or an armband or a card drawn at random, mimicking the arbitrariness of real caste while lacking its permanence. Second, an *unrefusable authority*: the instructor or facilitator runs the exercise and no participant is free to disobey, so consent is absent. Third, *licensed cruelty*: permission granted by that authority for the favored group to demean the disfavored one. Fourth, a *staged audience*: the classmates, colleagues, or cameras for whom the humiliation is performed, since the lesson is built to be witnessed. Fifth, the *reversal*: roles flip or the exercise ends on schedule, revealing the cruelty as a performance the authority could switch off, the one feature real caste never has.

The inversion, or the pedagogy of the oppressor. First of the book's three central claims, developed in Chapter Seven against Paulo Freire: a method built from caste, command, licensed cruelty, and spectacle stays the oppressor's method in any hands, and trains submission to cruelty whatever the cause it serves.

Iatrogenic empathy. Second central claim, developed in Chapter Eight from Ivan Illich's idea of iatrogenic harm, an injury caused by the treatment itself. It holds that the empathy exercise frequently produces the reverse of empathy, hardening attitudes and leaving participants worse than it found them.

The absence of the subject. Third central claim, developed in Chapter Nine. These rituals require the people they claim to be about to be missing or silenced, run only in sealed

conditions, and collapse on contact with the real world.

The sealed room. Shorthand for those required conditions: a controlled space, a compliant audience, and no one present who lives the burden being simulated.

The dramatic-demonstration paradigm. A style of mid-twentieth-century social psychology, associated with the era's most famous experiments, that sought to reveal human nature through a single vivid staged scene. Chapter Thirteen traces its loss of standing in the replication crisis and the reinvestigation of its signature studies.

Contact hypothesis. Gordon Allport's 1954 account, later confirmed by large meta-analysis, that contact between groups lowers prejudice when four conditions hold: equal status within the encounter, shared goals, cooperation rather than competition, and the support of legitimate authority. This is the evidence-based alternative the book sets against the dramatic method.

The jigsaw classroom. Elliot Aronson's cooperative teaching design, in which each student holds one necessary piece of a shared lesson and the group succeeds only by depending on one another. Offered here as the photographic negative of the dramatic method, lowering prejudice through cooperation in place of staged cruelty.

Struggle session. A Maoist public ritual in which a designated target was surrounded, denounced, and made to confess before a crowd. The book invokes it as the political ancestor of the coerced public confession, a comparison meant to name a shared mechanism rather than to equate a classroom with the Cultural Revolution.

Loyalty oath. A compelled declaration of belief or allegiance, such as the one imposed on University of California faculty from 1949 to 1952. The book treats the forced oath and the forced confession of feeling as two forms of one wrong, an institution prescribing what a person must profess.

The privilege walk. Participants step forward or back in answer to statements about their advantages, a widely used exercise that turns a structural analysis into a physical tableau of rank. Chapter Five examines how it humiliates without instructing and stages the poor while excluding them.

Reactance, or the backfire effect. A documented tendency of people to resist attempts to coerce their beliefs, so that heavy-handed persuasion often strengthens the attitude it meant to dissolve. This is one mechanism behind the iatrogenic result.

Bibliography

This is a selected bibliography of the principal published sources behind the argument, organized by chapter. It is not exhaustive. Many of the individual cases, the mock auctions, role-plays, and training scandals, are documented in news reporting, school-district records, and court filings that are dated and located in the text itself rather than listed again here.

Overture and Chapter One: The Armbands

The Overture and the opening chapter rest on the author's own recollection of an eye-color exercise conducted in a Lincoln, Nebraska, public school in 1973, together with period context for the city and its schools. They cite no secondary literature directly.

Chapter Two: April 5, 1968

Stephen G. Bloom, *Blue Eyes, Brown Eyes: A Cautionary Tale of Race and Brutality* (Oakland: University of California Press, 2021).

William Peters, *A Class Divided: Then and Now*, expanded edition (New Haven: Yale University Press, 1987).

Todd M. Mealy, *Shades of Brown: The Official Biography of Jane Elliott and the Blue Eyes, Brown Eyes Exercise* (Mechanicsburg, PA: Sunbury Press, 2023).

The Eye of the Storm, directed by William Peters (ABC News, 1970).

"A Class Divided," *Frontline* (PBS), 1985.

Chapter Three: The Anatomy of the Machine

Paulo Freire, *Pedagogy of the Oppressed*, translated by Myra Bergman Ramos (New York: Herder and Herder, 1970; written in Portuguese and first published in Spanish, 1968).

Ivan Illich, *Deschooling Society* (New York: Harper and Row, 1971).

Chapter Four: StarPower

R. Garry Shirts, *StarPower* (Del Mar, CA: Simile II, 1969).

Chapter Five: The Privilege Walk and the Tunnel

Peggy McIntosh, “White Privilege and Male Privilege,” working paper (Wellesley College Center for Research on Women, 1988), widely excerpted as “White Privilege: Unpacking the Invisible Knapsack” (1989).

Chapter Six: The Auction

Simone A. Schweber, “Simulating Survival,” *Curriculum Inquiry* 33, no. 2 (2003): 139-188.

United States Holocaust Memorial Museum, *Guidelines for Teaching About the Holocaust* (Washington, DC: United States Holocaust Memorial Museum).

Chapter Seven: The Pedagogy of the Oppressor

Hannah Arendt, *The Origins of Totalitarianism* (New York: Harcourt, Brace, 1951).

Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Viking, 1963).

Paulo Freire, *Pedagogy of the Oppressed* (see Chapter Three).

Chapter Eight: Iatrogenic Empathy

Ivan Illich, *Medical Nemesis: The Expropriation of Health* (London: Calder and Boyars, 1975; New York: Pantheon, 1976).

Jack W. Brehm, *A Theory of Psychological Reactance* (New York: Academic Press, 1966).

Lisa Legault, Jennifer N. Gutsell, and Michael Inzlicht, “Ironic Effects of Antiprejudice Messages: How Motivational Interventions Can Reduce (but Also Increase) Prejudice,” *Psychological Science* 22, no. 12 (2011): 1472-1477.

June P. Tangney, Jeffrey Stuewig, and Andres G. Martinez, “Two Faces of Shame: The Roles of Shame and Guilt in Predicting Recidivism,” *Psychological Science* 25 (2014): 799-805.

Chapter Nine: The Absence of the Subject

Arielle M. Silverman, Jason D. Gwinn, and Leaf Van Boven, “Stumbling in Their Shoes: Disability Simulations Reduce Judged Capabilities of Disabled People,” *Social Psychological and Personality Science* 6, no. 4 (2015): 464-471.

Arielle M. Silverman, “The Perils of Playing Blind: Problems with Blindness Simulation and a Better Way to Teach about Blindness,” *Journal of Blindness Innovation and Research* 5, no. 2 (2015).

Chapter Ten: From Stunt to Standard

U.S. Equal Employment Opportunity Commission, Select Task Force on the Study of Harassment in the Workplace, *Report of the Co-Chairs Chai R. Feldblum and Victoria A. Lipnic* (Washington, DC: EEOC, 2016).

Frank Dobbin and Alexandra Kalev, *Getting to Diversity* and “Why Diversity Programs Fail” (see Chapter Eleven).

Chapter Eleven: The Rented Conscience

Frank Dobbin and Alexandra Kalev, “Why Diversity Programs Fail,” *Harvard Business Review* 94, no. 7-8 (2016): 52-60.

Frank Dobbin and Alexandra Kalev, *Getting to Diversity: What Works and What Doesn’t* (Cambridge, MA: Belknap Press of Harvard University Press, 2022).

Benoît Monin and Dale T. Miller, “Moral Credentials and the Expression of Prejudice,” *Journal of Personality and Social Psychology* 81, no. 1 (2001): 33-43.

David Boles, *The Rented Crowd* (David Boles Books, 2026).

Chapter Twelve: What the Evidence Says

Anthony Petrosino, Carolyn Turpin-Petrosino, Meghan E. Hollis-Peel, and Julia G. Lavenberg, “‘Scared Straight’ and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency,” *Campbell Systematic Reviews* (2013).

Susan T. Ennett, Nancy S. Tobler, Christopher L. Ringwalt, and Robert L. Flewelling, “How Effective Is Drug Abuse Resistance Education? A Meta-Analysis of Project DARE Outcome Evaluations,” *American Journal of Public Health* 84, no. 9 (1994): 1394-1401.

Elizabeth Levy Paluck and Donald P. Green, “Prejudice Reduction: What Works? A Review and Assessment of Research and Practice,” *Annual Review of Psychology* 60 (2009): 339-367.

Elizabeth Levy Paluck, Roni Porat, Chelsey S. Clark, and Donald P. Green, “Prejudice Reduction: Progress and Challenges,” *Annual Review of Psychology* 72 (2021): 533-560.

Chapter Thirteen: The Collapse of the Dramatic Paradigm

Craig Haney, Curtis Banks, and Philip Zimbardo, “Interpersonal Dynamics in a Simulated Prison,” *International Journal of Criminology and Penology* 1 (1973): 69-97.

Thibault Le Texier, “Debunking the Stanford Prison Experiment,” *American Psychologist* 74, no. 7 (2019): 823-839.

Open Science Collaboration, “Estimating the Reproducibility of Psychological Science,” *Science* 349, no. 6251 (2015): aac4716.

Stanley Milgram, *Obedience to Authority: An Experimental View* (New York: Harper and Row, 1974).

Gina Perry, *Behind the Shock Machine: The Untold Story of the Notorious Milgram Psychology Experiments* (Brunswick: Scribe, 2012).

Dana R. Carney, Amy J. C. Cuddy, and Andy J. Yap, “Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance,” *Psychological Science* 21, no. 10 (2010): 1363-1368.

Eva Ranehill, Anna Dreber, Magnus Johannesson, Susanne Leiberg, Sunhae Sul, and Roberto A. Weber, “Assessing the Robustness of Power Posing,” *Psychological Science* 26, no. 5 (2015): 653-656.

Thibault Le Texier, *Histoire d’un mensonge: Enquête sur l’expérience de Stanford* (Paris: La Découverte, 2018).

Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn, “False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant,” *Psychological Science* 22, no. 11 (2011): 1359-1366.

S. Alexander Haslam and Stephen D. Reicher, “Contesting the ‘Nature’ of Conformity: What Milgram and Zimbardo’s Studies Really Show,” *PLoS Biology* 10, no. 11 (2012): e1001426.

Stephen Reicher and S. Alexander Haslam, “Rethinking the Psychology of Tyranny: The BBC Prison Study,” *British Journal of Social Psychology* 45, no. 1 (2006): 1-40.

Bettina Stangneth, *Eichmann Before Jerusalem: The Unexamined Life of a Mass Murderer*, translated by Ruth Martin (New York: Alfred A. Knopf, 2014).

Martin S. Hagger et al., “A Multilab Preregistered Replication of the Ego-Depletion Effect,” *Perspectives on Psychological Science* 11, no. 4 (2016): 546-573.

Chapter Fourteen: The Vulnerable Days

Executive Order 13950, “Combating Race and Sex Stereotyping,” *Federal Register*, September 22, 2020.

Executive Order 14151, “Ending Radical and Wasteful Government DEI Programs and Preferencing” (2025).

Executive Order 14173, “Ending Illegal Discrimination and Restoring Merit-Based Opportunity” (2025).

Abigail Thompson, “The University’s New Loyalty Oath,” *Notices of the American Mathematical Society* 66, no. 11 (2019).

CRT Forward Tracking Project, University of California, Los Angeles, School of Law.

West Virginia State Board of Education v. Barnette, 319 U.S. 624 (1943).

Chapter Fifteen: What Works

Gordon W. Allport, *The Nature of Prejudice* (Cambridge, MA: Addison-Wesley, 1954).

Thomas F. Pettigrew and Linda R. Tropp, “A Meta-Analytic Test of Intergroup Contact Theory,” *Journal of Personality and Social Psychology* 90, no. 5 (2006): 751-783.

Elliot Aronson and Diane Bridgeman, “Jigsaw Groups and the Desegregated Classroom: In Pursuit of Common Goals,” *Personality and Social Psychology Bulletin* 5, no. 4 (1979): 438-446.

Claudia Goldin and Cecilia Rouse, “Orchestrating Impartiality: The Impact of ‘Blind’ Auditions on Female Musicians,” *American Economic Review* 90, no. 4 (2000): 715-741.

Iris Bohnet, *What Works: Gender Equality by Design* (Cambridge, MA: Belknap Press of Harvard University Press, 2016).

David Broockman and Joshua Kalla, “Durably Reducing Transphobia: A Field Experiment on Door-to-Door Canvassing,” *Science* 352, no. 6282 (2016): 220-224.

Elizabeth Levy Paluck, Seth A. Green, and Donald P. Green, “The Contact Hypothesis Re-evaluated,” *Behavioural Public Policy* 3, no. 2 (2019): 129-158.

Muzafer Sherif, O. J. Harvey, B. Jack White, William R. Hood, and Carolyn W. Sherif, *The Robbers Cave Experiment: Intergroup Conflict and Cooperation* (Norman: University of Oklahoma Book Exchange, 1961).

Gina Perry, *The Lost Boys: Inside Muzafer Sherif's Robbers Cave Experiment* (Brunswick: Scribe, 2018).

Questions for Discussion

These questions are offered for readers, classrooms, and reading groups who want to argue with the book rather than only absorb it. They follow its line of thought from the personal opening to the closing refusal, and none of them has a settled answer the author would insist on.

The book opens with an eight-year-old refusing to play a game everyone else accepted. What did that refusal cost him, and what would compliance have cost instead?

Coercive moral pedagogy is defined here as the attempt to teach empathy by manufacturing suffering. Where have you met an exercise that fits that description, in a school, a workplace, or a training room, and how did it feel from the inside?

The five-part anatomy holds that every such exercise assigns caste by accident, enforces it through an authority no one can refuse, licenses cruelty, stages an audience, and ends in a reversal. Does that structure describe anything you have witnessed, or does it strain to fit cases it was never built for? Where it does fit, which element surfaced first?

Good teaching often asks students to feel discomfort, to sit with a hard text or a disturbing photograph. Where is the line between a discomfort that teaches and the manufactured suffering the book indicts, and does the book’s net risk catching honest teaching along with the abuse it means to stop?

The hardest claim in the book is that a just cause does not change what a coercive method teaches. Do you accept that a method can carry its own lesson regardless of its goal, or does intention weigh more heavily than the argument allows?

Jane Elliott ran her exercise for decades and still defends it past the age of ninety, conceding only that the day was hard and never that the method was wrong. Having read the case against the technique, can you hold your judgment of the woman apart from your judgment of the thing she built?

A central argument is that the eye-color exercise needed a room with no Black child in it. Does the absence of the real subject weaken these exercises as much as the book

claims, and can you imagine one of this kind that would survive its subject’s presence?

Iatrogenic harm means an injury caused by the treatment itself. The book contends that the empathy industry often hardens the attitudes it sets out to soften. If that is so, why have the exercises stayed popular for so long?

Consider the scared-straight programs, which raised the offending they were meant to prevent. What does it say about good intentions that a frightening, memorable lesson can fail in the exact direction it was designed to move?

The mock slave auction and the Gestapo role-play appear here as the genre at its worst, the point where the subject is present and the ritual comes apart. What separates those cases from the milder ones, and is the difference a matter of degree or of kind?

Chapter Thirteen argues that the famous experiments which once made this style of demonstration look scientific have lost their footing. How much should the unraveling of the Stanford prison study change the way a reader trusts the dramatic classroom exercise?

Convergence is the spine of the evidence chapter, the claim that programs as unlike one another as scared-straight tours, drug education, and workplace training all fail in the same direction. How much weight should a reader put on agreement across fields that otherwise share little, and could that agreement be an artifact of who studies these programs and how?

The author holds that he is for the goal of equality and against the method, and that the two stances sit together without strain. Did the book convince you that a person can attack a practice while keeping faith with its aim, or did that line blur as you read?

Chapter Fourteen charges that the reluctance of the method’s defenders to police their own pedagogy handed their opponents a weapon. Is that a fair charge, and what would honest self-policing have looked like?

Set the mandatory diversity statement beside the classroom book ban, as Chapter Fourteen does, and the book calls them mirror images that share one method while granting that their goals are unequal. Does the mirror hold, or does treating the two as a matched pair flatter one side and slander the other? Where would you place the boundary between a shared method and unequal ends?

The remedy the book proposes, contact on equal terms, cooperation toward shared goals, and lasting structural change, is deliberately undramatic. Why might a quiet, slow cure be a hard sell beside one vivid afternoon of role-play?

Elliot Aronson’s jigsaw classroom enters as the photographic negative of Elliott’s exercise, born in the same era and built on the opposite principle. If cooperation does lower prejudice where staged humiliation does not, what accounts for the difference, and how settled do you find the evidence on either side?

Suppose you had to teach the reality of slavery, or the fact of unearned advantage, to a room of children or new employees, with the dramatic shortcut off the table. What would you do, hour by hour, and how would you know whether it had worked?

Think about the staged audience in the anatomy. How does performing a lesson for spectators change what the lesson deposits in the people performing it, and where do you see that same dynamic outside any classroom?

The book closes where it began, on a refusal it calls the first and last defense against every pedagogy of the oppressor. Is the freedom to decline a rigged game a sufficient defense on its own, or only the beginning of one?

Having finished the book, would you answer differently now if you were handed an armband and told to wear it? What would you do, and what do you think it would cost you?

About the Author

David Boles is a writer, dramatist, editor, and publisher who has worked across medicine, law, and dramatic literature. He holds an MFA from the Oscar Hammerstein II Center for Theatre Studies at Columbia University and has belonged to the Dramatists Guild of America since 1984, along with the Authors Guild and PEN America. In 1975 he founded David Boles Books, and he has since written and published widely in nonfiction, drama, and fiction, including the Ideas and Inquiry line to which this book belongs. He lives and works in the New York area, in close collaboration with his wife, Janna Sweenie, an educator and performer in American Sign Language. His books and ongoing writing can be found at BolesBooks.com.

Also by the Author

The Source Code of the Self: How a Playwright Wrote the Operating System of Popular Psychology

The Rented Crowd: Counterfeit Applause and the Manufacture of the Public Mind, from Nero's Payroll to the Bot Farm

The Off Switch: Money, Property, and the Power to Switch Off a Life

The Nearest Hand: Silver, Stimulus, Silicon, and the Five-Century Story of Who Gets Paid First

The Brittle Self: A History of the Fragile Mind, from the Glass King to the Simulation

Beyond the Burial Tree: The Seizure of the Native Dead and the Unfinished Return

In My Mind I'm Standing Up: A History of Recantation and the Coerced Word

For the complete and current catalog, in print and in audio, visit BolesBooks.com.